

# The option value of new therapeutics

Anup Malani and Feifang Hu\*  
MacClean Center Conference, University of Chicago

October 25, 2004

## Abstract

Drug control agencies such as the U.S. FDA typically decide whether to approve a new therapeutic based upon the average treatment effect of the new therapeutic. Unless the mean effect of the new therapeutic is greater than a conventional treatment or placebo, the agency will not permit the new therapeutic's use. In the presence of ex post heterogeneity in treatment effects, however, average treatment effects fail to account for the full value of a new therapeutic. Even if the new therapeutic is no better than placebo for the average patient, there may be a number of patients who are better off under the new therapeutic than any conventional treatment. If this is the case, prohibiting use of the new therapeutic imposes a welfare loss on these patients. This paper examines the major influences on this loss, which we call the option value of the new therapeutic. It proposes simple estimators for this value and calculates the option value of drugs examined in the Cardiac Arrhythmia Suppression Trial (CAST) I. Finally, the paper examines drawbacks of relying on option value to make drug approval decisions.

The U.S. Food and Drug Agency (FDA) employs a simple decision-rule when deciding whether to approve a new drug for use by physicians: the average treatment effect of the new drug must be superior to the average effect of placebo. If there is ex post heterogeneity in treatment effects, however, this decision rule may reject drugs that have positive social value. If a drug does not ultimately have the same effect on each and every patient, then there may be some patients who are better off with the new drug than with placebo or conventional treatment, even though the average patient may be better off with placebo. We call this lost social value the option value of the drug because, if the drug is not approved, doctors will not have the option of using it on patients for whom other treatments are inferior.

---

\*University of Virginia Law School and Health Evaluations Sciences Department, University of Virginia School of Medicine; Department of Statistics, University of Virginia. We thank Maria Fitzpatrick for research assistance. The Cardiac Arrhythmia Suppression Trial (CAST) is conducted and supported by the National Heart, Lung, and Blood Institute (NHLBI) in collaboration with CAST Investigators. This Manuscript was not prepared in collaboration with investigators of the CAST and does not necessarily reflect the opinions or views of the CAST or the NHLBI.

A simple example with binary outcomes illustrates the cost of the FDA's decision rule. Consider a hypothetical ailment, from which all patients spontaneously recover with probability  $p_0$ . Suppose there is a new treatment that will cure a patient with probability  $p_1$ . If  $p_1 < p_0$  and the new drug is not approved, a patient's probability of recovery is simply  $p_0$ . If the drug were approved despite the fact that  $p_1 < p_0$ , a patient's probability of cure would be  $p_0 + (1 - p_0)p_1 > p_0$ . Every patient would be better off by the probability of not recovering spontaneously times the probability of recovery with the new drug.

The option value of treatment is not a novel concept, certainly not to economists (Heckman et al., 1997; Aakvik et al., 2000; Heckman, 2001). While it is familiar to the biostatistics and medical literatures, neither appears to discuss it directly (see, e.g., Plewis, 2002). Whereas virtually all relevant studies in these fields appreciate that treatment effects vary by clinical group, we have yet to find a paper that examines the implications of heterogeneity in treatment effects even after conditioning on all clinical variables to the investigator. In econometrics parlance, it is as if treatment outcomes were regressed upon available clinical variables and treatment dosage, and the literature had not focused on the positive implications of variance in the residual and in coefficient estimates.

Practitioners — namely the FDA and the pharmaceutical industry — are clearly aware of the implications of treatment heterogeneity. After all, the FDA only requires that a new drug be demonstrated superior to placebo for some identifiable patient subpopulation and pharmaceutical companies employ careful screening processes to ensure that the subjects enrolled in a clinical trial of a new drug are likely to respond positively (relative to placebo) to the drug (Abboud 2004). These adjustments are limited in how effectively they capture the option value of certain drugs, however, because those patient characteristics observable to investigators may not fully capture all the variance in treatment effects. Practicing doctors have more information on treatment effects because they spend more time with the patient or because investigators are constrained by trial protocol.

This paper is a first attempt at a systematic analysis of the option value of a drug and the cost of average treatment effect decision rules for drug approval or insurance coverage. Our main analytic findings are as follows:

1. The option value of a drug is a product of heterogeneity in treatment effects. This heterogeneity can be broken down into three parts: heterogeneity that can be predicted with information (on the patient) available to the investigator and the practicing doctor, with information available only to the doctor, and with information available to neither. Each set of information yields conceptually distinct option values (though there may be interactions between information sets).
2. Decision rules, such as the FDA's, that examine only the average treatment effect of drugs cost society the option value of a new drug when the average value of the new drug is lower than the average value of the relevant

benchmark. This lost value rises in the heterogeneity of treatment effects. Where the benchmark is the best conventional treatment currently available (as opposed to merely placebo), as it often is in Europe, the lost option value is a function of the order in which drugs are introduced, i.e., the timing of innovation.

3. Data from cross-over clinical trials are sufficient to produce point estimates of the full option value of a new drug. Data from parallel-arm trials can provide bounds on the full option value or point estimates of the option value due to heterogeneity predictable with information available to the research investigator. They cannot provide point estimates of the remaining option value of the drug.
4. Decision rules that rely on the full option value of a new drug will overestimate the actual or "realized" value of adding that drug to practicing doctors' portfolios. There are two reasons. First, practicing doctors lack the information to realize option value due to heterogeneity which can only be predicted with information that neither they nor investigators have. Second, practicing doctors may not act optimally even on the information they do have. Indeed, average treatment effect decision rules can be rationalized on the assumption that doctors ignore all information they have about the heterogeneity of treatment effects. Since this assumption is extreme, it is likely that average treatment effects underestimate the realized option value of a new drug.

The remainder of the paper may be outlined as follows. Part I of this paper defines the option value of drugs and characterizes the cost of decision-rules that rely on average treatment effect rather than the option value of a drug. Part 2 provides estimators for the option value of a drug using data from ordinary clinical trials. Employing data from the Cardiac Arrhythmia Suppression Trial (CAST) I, Part 3 demonstrates significant ex post heterogeneity of treatment effects of two drugs (encainide and flecainide) for the suppression of arrhythmia and thus the potential for significant option value with these drugs. It estimates this option value using two of the estimators from Part 2. Finally, Part 3 discusses some hazards of a decision-rule for drug approval which relies upon the option value of drugs.

## 1 Option value of drugs

### 1.1 Definition of lost option value

The value of a drug to an individual patient is the improvement in health outcomes it produces for that patient given that the drug is used when (and only when) it offers a superior outcome to alternative remedies. The drug has no value if it is not used. If the drug is used when a better, mutually exclusive remedy is available for the patient, however, use of the drug prevents the patient

from experiencing a superior outcome and is thus a social waste. The full value of the drug across all patients is the sum of improvements in health outcomes it offers patients for whom the drug is superior to alternatives. We call this the option value of the drug because this is the value that would be lost if doctors were able optimally to prescribe therapies for patients but were prohibited from prescribing this one particular drug.<sup>1</sup>

In mathematical terms, the option value of a drug  $n$  in a static setting can be written as

$$\begin{aligned} V_n &= \int \cdots \int (y_n - \max\{y_0, \dots, y_{n-1}\}) f(y_1, \dots, y_n | y_n \geq \max\{y_0, \dots, y_{n-1}\}, \Omega) \\ &= E(y_n - \max\{y_0, \dots, y_{n-1}\} | y_n \geq \max\{y_0, \dots, y_{n-1}\}, \Omega) \end{aligned} \quad (1)$$

where  $f(y_1, \dots, y_n)$  gives the population distribution of outcomes  $\{y_0, \dots, y_n\}$  for drugs  $k = 1, \dots, n$ ; drugs  $k = 1, \dots, n - 1$  are alternative therapies to drug  $n$ ; and  $\Omega$  is all information available on each patient in the population. For simplicity, (1) assumes the patient may only take one treatment and all treatments are mutually exclusive. We subtract the outcome under the next best treatment for each individual because it is a gain which could be achieved even without drug  $n$ .

The option value of a new drug is mainly a function of heterogeneity of treatment effects. The uncertain component of that heterogeneity can be fully represented by the information set  $\Omega$ . That set has three components  $(X, Z, W)$ . The first component  $X$  is information in the possession of investigators and practicing doctors; the second component  $Z$  is in the possession of only doctors; and the third component  $W$  is in the possession of neither party. Conditioning on  $X$  gives the heterogeneity that can be predicted by information in the possession of research investigators and the maximum option value that can be extracted by these investigators. Conditioning on  $X$  and  $Z$  gives the heterogeneity that can be predicted by information in the possession of practicing doctors and the maximum option value that can be extracted by these doctors. Finally, the difference in option value when conditioning on  $\Omega$  and when conditioning on  $(X, Z)$  gives the option value that exists, but cannot be realized because no one has sufficient information to extract it. In other words, there are times when a patient will do better on the new drug than any alternative but, because no one can predict this with available information, one cannot recommend the use of the new drug over alternatives. This categorization of the main information sources of option value is useful in accounting for the lost option value from an average treatment rule and for accounting for the lost option value generated by each of the estimators in Section 2.

Drug control agencies such as the FDA and the European Agency for the Evaluation of Medicinal Products require that a drug be proven effective and

---

<sup>1</sup>The option value of a drug includes its "real option value" as economists use that term. For example, a physician may learn about the value of conventional (or no) treatment to a given patient over time. The physician's task is to employ this information to choose the optimal time to switch to the new treatment.

safe before it can be prescribed by doctors.<sup>2</sup> In the U.S., efficacy requires demonstration that a drug is more effective than a placebo control. In Europe, efficacy requires that the drug be superior to either placebo or a conventional control. Which benchmark is appropriate depends on which is ethically permissible to use. In general, placebo-controlled trials are permitted in Europe only for minor ailments or where there is no established conventional treatment for an ailment (Jost, 2001).

Because health insurance is a significant hurdle to access to a new drug and insurance companies are increasingly turning to the use of formularies to control costs, the standards for inclusion of drugs in formularies are an addition hurdle to realizing the full value of a drug.<sup>3</sup> Formularies mainly rely upon comparisons of cost-effectiveness, and not merely effectiveness. On the theory that drugs introduced earlier in time have lower costs than drugs introduced later (a good example being generic forms of drugs no longer under patent), employment of a cost-effectiveness standard is similar to employment of a standard that requires a drug to be more effective than a conventional control (American Medical Association – Council on Ethical and Judicial Affairs, 1998; Harris, 2004).

Whatever the benchmark for approval or coverage of a drug, drug control agencies and pharmacy benefit managers at insurance companies primarily consider the average effect of a drug relative to the average effect of the benchmark when making access decisions. This may be due to the fact that full distributions of treatment effects are difficult to estimate or that averages are simpler statistics to understand. As such, we can write the placebo benchmark and conventional treatment benchmark decision rules for access to a new drug as  $1[E(y_n) > E(y_0)]$  and  $1[E(y_n) \geq \max\{E(y_0), \dots, E(y_{n-1})\}]$ , respectively, where  $1[\cdot]$  is an indicator of whether the drug is approved. This implies that the option value lost under the placebo rule and the conventional treatment

---

<sup>2</sup>Technically, in the U.S., approval is required before any the drug's can make any specific health-related claims about the drug. Doctors may prescribe any substance for purposes of therapy subject only to medical malpractice liability and sanctions by licensing authorities. In practice, however, drug control agency approval is viewed a prerequisite for both sale and prescription for medical purposes.

In Europe, approval is possible through three channels. First, a drug company may file with the AEMP. Second, a drug company may file with a European Union (EU) member country's drug control agency and send notice to all other EU nations. Other EU members, however, may formally oppose the application and block approval of the drug in their jurisdictions. Third, the company can file for approval with a specific member's agency, but only will only be permitted to sell it's drug in that jurisdiction.

<sup>3</sup>Technically, a formulary does not prohibit access to a drug. At most it requires the patient to pay out of pocket for the drug even though she has already paid her insurance premiums. Moreover, many insurance companies permit the patient's physician to override the formulary or the patient to petition the insurance company for an exception to its formulary policy. Mitchell (1999). In practice, however, absence of coverage for a treatment leads to lower use of that treatment (cite).

rules are

$$\begin{aligned}
LV_n &= E(y_n - \max\{y_0, \dots, y_{n-1}\} | y_n \geq \max\{y_0, \dots, y_{n-1}\}, \Omega) & (P) \\
&\quad - E(y_n - \max\{y_0, \dots, y_{n-1}\} | E(y_n) > E(y_0), X) \\
LV_n &= E(y_n - \max\{y_0, \dots, y_{n-1}\} | y_n \geq \max\{y_0, \dots, y_{n-1}\}, \Omega) & (CT) \\
&\quad - E(y_n - \max\{y_0, \dots, y_{n-1}\} | E(y_n) \geq \max\{E(y_0), \dots, E(y_{n-1})\}, X)
\end{aligned}$$

respectively. Note that there are two sources of leakage in value: the failure to consider heterogeneity in treatment effects (due to the average effect rule) and lack of information unavailable to investigators (during implementation of the average effect rule).

## 1.2 Determinants of lost option value

In order for there to be lost option value due to an average effect rule, three conditions must be met. First, there must be ex post heterogeneity in treatment outcomes. If everyone has the same response to treatment, the average treatment effect captures all information about each patient's response to treatment. Second, the minimum of the support treatment effects of the benchmark treatment must be less than the maximum of the support for treatment effects of the new drug. Otherwise, the worst possible outcome with the benchmark drug is better than the best possible outcome with the new drug. Third, and most importantly, the average treatment effect of the new drug must be worse than the average effect of the benchmark treatment. If the average treatment effect of the new drug is better, then it will be approved by the drug control agency (or covered by the insurance company), and practicing doctors will be able to tap its option value for their patients.

A L'Abbe plot provides a simple illustration of the lost option value of treatment due to average-effect decision rules.<sup>4</sup> See Figure 1. The x- and y-axes of this plot gives health outcomes of individuals in the control and new drug arms, respectively, of trials. The option value of the new drug is equal to sum of the vertical distances from each point above the 45° degree line to that line. An average treatment effect rule requires a plot of the average outcome in each group (indicated by the large dot). If this point lies below the 45° line, then the rule rejects the new drug. A drug experiences positive lost option value if a plot of the average outcome in each arm lies below the 45° but there are individual outcomes above that line.<sup>5</sup>

If the benchmark for an average effect rule is conventional treatment, the lost option value due to that rule is a function of the timing of innovation. If drugs

---

<sup>4</sup>A L'Abbe plot is used in systematic analyses with each point representing the mean outcome from the two arms of a trial. Our use is a slight modification of the plot.

<sup>5</sup>An intuitive, but not so useful observation is that the lost option value falls in covariance between the new drug and the benchmark. This is easily derived from the definition of the covariance of  $(y_1, y_2)$  in the case where there is only one alternative therapy ( $k = 1$ ) to the new drug ( $k = 2$ ):  $E(\Delta_2 | \Delta_2 > 0) = -\left(\frac{1}{\mu_1}\right) \sigma_{12} + \frac{1}{\mu_1} E(y_1 y_2) - \mu_1 + E(\Delta_1 | \Delta_1 > 0)$ , where  $\Delta_2 = y_2 - y_1 = -\Delta_1$ .

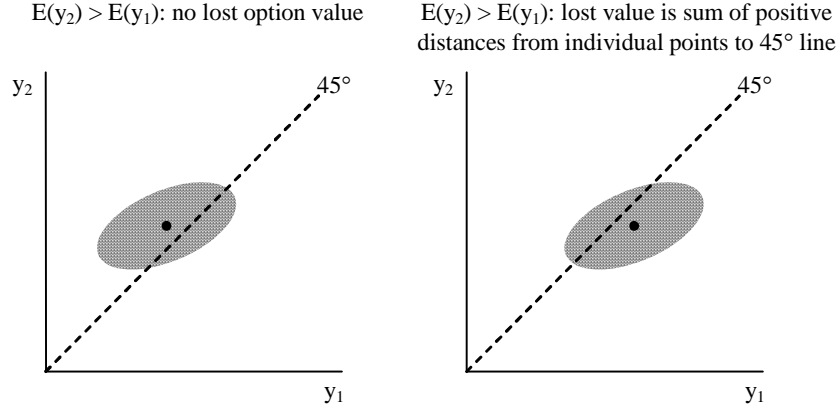


Figure 1: Illustration of average effect rule, option value, and positive lost option value. (Large dot is treatment effect for the average patient.)

are invented in the order of their average treatment effects, each drug will be approved. The reason is that the average effect of each new drug will be greater than the average effect of all conventional controls. When the benchmark is placebo, timing is irrelevant. The effect of placebo, i.e., natural progression, does not change over time.

### 1.3 Current methods of extracting option value

The current regulatory regime in the U.S. has three methods of extracting option value. The first is to use a placebo benchmark for drug approval. Presumably the conventional benchmark is more difficult to meet than the placebo benchmark, otherwise no conventional treatment would have been approved in the first case. Thus the placebo benchmark is less likely to reject any given drug and thus lose out on option value.

The second solution is to measure more patient covariates in clinical trials conducted to support an application for new drug approval. The goal is for the information set  $X$  to cannibalize as much of  $Z$  as possible so that the investigator can determine the option value of information typically available only to physicians. An example of this is the open titration period before randomization, a period which allows investigators to observe patient reaction over time to determine the optimal dosage for each patient. Gathering more data has much potential, but the additional data is generally underutilized because investigators still report only the average treatment effects conditional on a larger  $X$ . (This is likely because that is the only estimator the FDA will recognize.) The estimators proposed in the next section attempt to correct this suboptimal use of available information.

The third solution is an ad hoc approach devised by pharmaceutical compa-

nies to increase the probability of drug approval: select subjects for enrollment in trials if they are likely to respond positively to the new drug and poorly to placebo (Abboud, 2004). This approach is implemented by means of inclusion and exclusion criteria. It has significant information requirements because companies must determine the option value to patients before a clinical trial to support a new drug application is ever run. Therefore, this approach is either very costly, because companies must conduct mini-trials before actual trials to predict the option value of their new drug, or companies must focus on selecting patients who are unlikely to perform well on placebo rather than well on the new drug, because information to make such predictions are currently available.<sup>6</sup>

## 2 Calculating the option value of drugs

In this section we propose three estimators for the option value of a drug. The first estimator assumes the investigator has data from a cross-over trial and yields an estimate of the full option value of a drug, i.e., conditional on  $\Omega$ . The second and third estimators are appropriate for data from parallel-arm trials. The second estimator provides bounds on the full option value of a drug and third produces a point estimate of the option value given only information available to investigators, i.e., conditional only on  $X$ . The limitation of the second estimator is that bounds may be so wide as to be uninformative. The limitation of the third estimator is that it cannot extract the option value from information set  $Z$  available to the practicing doctor, let alone the full option value of the drug.

Because our purpose here is merely illustration, we assume there are only two drugs available. The new drug has index 2 and the control has index 1. (The control may be placebo or conventional control that is better than placebo for all subjects.) A limitation of our assumption is that, when there are more than two drugs, our estimators only provide an estimate of the option value of drug 2 relative to drug 1, not to all other drugs. Comparing a new drug to more than one other drug is a straightforward extension of our first estimator. The same is true for the second estimator, though the more alternative drugs there are the wider are the bounds generated. Comparing more than one drug is a non-trivial extension of the third estimator and we leave the problem to a future draft or research.

---

<sup>6</sup>Finally, the third approach may be a high risk solution. If pharmaceutical companies use inclusion criteria (as opposed to exclusion criteria), they may gamble a drug's approval on a small subset of all patients who draw positive option value. The reason is that, even if there are  $m$  discrete patient strata for which the new drug has positive option value, inclusion criteria as typically implemented only draws from one or a small subset of these strata. If the selected strata have higher variance than the average strata, then the probability of type I error (rejection of the new drug even though it has positive option value) rises.



## 2.1 Cross-over trials

In these trials we see each subject taking both treatment 2 and 1. An estimator for the option value of 2 relative to 1 is

$$\hat{V}_{21} = \frac{\sum_{i: y_{2i} > y_{1i}} (y_{2i} - y_{1i})}{N_{y_{2i} > y_{1i}}}$$

where  $i$  indexes subjects in the trial and  $N_{y_{2i} > y_{1i}}$  is the number of subjects who do better on drug 2 than 1.

Let  $(y_{1i}, y_{2i})$ ,  $i = 1, \dots, n$  be  $n$  independent and identical distributed samples. Suppose the joint density (or probability) of  $(Y_1, Y_2)$  is  $f(y_1, y_2)$ . We do not assume  $Y_1$  and  $Y_2$  are independent, because the responses of the same patient to the two drugs could be dependent.

As we defined in Section 1, the theoretical option value is

$$\begin{aligned} V_{21} &= E(Y_2 - Y_1 | Y_2 > Y_1) = \frac{E(Y_2 - Y_1)I(Y_2 - Y_1 > 0)}{EI(Y_2 - Y_1 > 0)} \\ &= \frac{\int_{-\infty}^{\infty} \int_{-\infty}^{y_2} (y_2 - y_1) f(y_1, y_2) dy_1 dy_2}{\int_{-\infty}^{\infty} \int_{-\infty}^{y_2} f(y_1, y_2) dy_1 dy_2}. \end{aligned}$$

Now we want to study the properties of the estimate  $\hat{V}_{21}$ . Theorem 1 shows that  $\hat{V}_{21}$  is a consistent estimate of  $V_{21}$ , and Theorem 2 provides the asymptotic normality of  $\hat{V}_{21}$ . Based on the result of Theorem 2, we can construct confidence interval of the theoretical option value  $V_{21}$ .

**Theorem 1** *Suppose that  $E[I(Y_2 > Y_1)] > 0$  and  $E(Y_2 - Y_1)I(Y_2 > Y_1)$  exists, where  $I$  is the indication function. Then  $\hat{V}_{21}$  converges to  $V_{21}$  in probability as  $n \rightarrow \infty$ .*

**Proof.** First, the estimate  $\hat{V}_{21}$  can be rewritten as

$$\hat{V}_{21} = \frac{\sum_{i=1}^n (y_{2i} - y_{1i}) I(y_{2i} - y_{1i} > 0)}{\sum_{i=1}^n I(y_{2i} - y_{1i} > 0)}.$$

By the weak law of large numbers, we have

$$n^{-1} \sum_{i=1}^n I(y_{2i} - y_{1i} > 0) \rightarrow EI(Y_2 - Y_1 > 0)$$

and

$$n^{-1} \sum_{i=1}^n (y_{2i} - y_{1i}) I(y_{2i} - y_{1i} > 0) \rightarrow E(Y_2 - Y_1)I(Y_2 - Y_1 > 0)$$

in probability. Therefore, from Slutsky's theorem, we have that  $\hat{V}_{21} \rightarrow V_{21}$  in probability. ■

**Theorem 2** Suppose that  $E[I(Y_2 > Y_1)] > 0$  and  $E(Y_2 - Y_1)^2 I(Y_2 > Y_1) < \infty$ . Then

$$\sqrt{n}(\hat{V}_{21} - V_{21}) \rightarrow N(0, \sigma^2),$$

in distribution, where

$$\begin{aligned} \sigma^2 &= \{E[I(Y_2 - Y_1 > 0)]\}^{-3} \times \\ &\quad \{EI(Y_2 - Y_1 > 0)E(Y_2 - Y_1)^2 I(Y_2 - Y_1 > 0) - \\ &\quad [E(Y_2 - Y_1)I(Y_2 - Y_1 > 0)]^2\}. \end{aligned}$$

**Proof.** Consider the difference of  $\hat{V}_{21}$  and  $V_{21}$ .

$$\begin{aligned} \hat{V}_{21} - V_{21} &= \frac{\{n^{-1} \sum_{i=1}^n (y_{2i} - y_{1i}) I(y_{2i} - y_{1i} > 0) EI(Y_2 - Y_1 > 0) - \\ &\quad n^{-1} \sum_{i=1}^n I(y_{2i} - y_{1i} > 0) E(Y_2 - Y_1)^2 I(Y_2 > Y_1)\}}{n^{-1} \sum_{i=1}^n I(y_{2i} - y_{1i} > 0) EI(Y_2 - Y_1 > 0)} \quad (2) \end{aligned}$$

$$\begin{aligned} &= \frac{\{EI(Y_2 - Y_1 > 0)[n^{-1} \sum_{i=1}^n (y_{2i} - y_{1i}) I(y_{2i} - y_{1i} > 0)] - \\ &\quad - E(Y_2 - Y_1)^2 I(Y_2 > Y_1)\} - \{E(Y_2 - Y_1)^2 I(Y_2 > Y_1) \times \\ &\quad [n^{-1} \sum_{i=1}^n I(y_{2i} - y_{1i} > 0) - EI(Y_2 - Y_1 > 0)]\}}{n^{-1} \sum_{i=1}^n I(y_{2i} - y_{1i} > 0) EI(Y_2 - Y_1 > 0)} \quad (3) \end{aligned}$$

By the central limit theorem, we have

$$\sqrt{n}[n^{-1} \sum_{i=1}^n (y_{2i} - y_{1i}) I(y_{2i} - y_{1i} > 0) - E(Y_2 - Y_1)I(Y_2 > Y_1)] \rightarrow N(0, \sigma_1^2)$$

in distribution, where  $\sigma_1^2 = \text{Var}((Y_2 - Y_1)I(Y_2 - Y_1 > 0))$ . Similarly we have

$$\sqrt{n}[n^{-1} \sum_{i=1}^n I(y_{2i} - y_{1i} > 0) - EI(Y_2 > Y_1)] \rightarrow N(0, \sigma_2^2)$$

in distribution, where  $\sigma_2^2 = \text{Var}(I(Y_2 - Y_1 > 0))$ .

Therefore, by Slutsky's theorem, we have

$$\sqrt{n}(\hat{V}_{21} - V_{21}) \rightarrow N(0, \sigma^2),$$

in distribution, where

$$\begin{aligned} \sigma^2 &= (EI(Y_2 - Y_1 > 0))^{-3} \{EI(Y_2 - Y_1 > 0)E(Y_2 - Y_1)^2 I(Y_2 - Y_1 > 0) \\ &\quad - [E(Y_2 - Y_1)I(Y_2 - Y_1 > 0)]^2\} \end{aligned}$$

■

Now we can estimate  $\sigma^2$  by

$$\begin{aligned} \hat{\sigma}^2 &= [n^{-1} \sum_{i=1}^n I(y_{2i} - y_{1i} > 0)]^{-3} \times \\ &\quad \{n^{-1} [\sum_{i=1}^n I(y_{2i} - y_{1i} > 0)] n^{-1} [\sum_{i=1}^n (y_{2i} - y_{1i})^2 I(y_{2i} - y_{1i} > 0)] - \\ &\quad (n^{-1} [\sum_{i=1}^n (y_{2i} - y_{1i}) I(y_{2i} - y_{1i} > 0)])^2\}. \end{aligned}$$

It is easy to show that  $\hat{\sigma}^2$  is a consistent estimate of  $\sigma^2$ . From Theorem 2, we can then construct 95% confidence interval of  $V_{21}$  as

$$[\hat{V}_{21} - 1.96\hat{\sigma}/\sqrt{n}, \hat{V}_{21} + 1.96\hat{\sigma}/\sqrt{n}].$$

Based on this confidence interval, we can also calculate a requisite sample size for a given accuracy level.<sup>7</sup>

## 2.2 Parallel-arm trials

**Observe no covariate.** If the investigator observes treatment outcomes from a parallel-arm trial, but no clinical covariates, then the investigator can construct bounds on the option value. Suppose the investigator has data from a parallel arm trial of drug 2 and drug 1. Estimators for the upper and lower bounds of the option value of drug 2 relative to drug 1 are

$$\begin{aligned}\hat{V}_{21} &= \frac{\sum_{i: y_{2i} > \min_j \{y_{1j}\}} (y_{2i} - \min_j \{y_{1j}\})}{N_{y_{2i} > \min_j \{y_{1j}\}}} \\ \hat{V}_{21} &= \frac{\sum_{i: y_{2i} > \max_j \{y_{1j}\}} (y_{2i} - \max_j \{y_{1j}\})}{N_{y_{2i} > \max_j \{y_{1j}\}}}\end{aligned}$$

where  $i$  and  $j$  index subjects in the trial and, e.g.,  $N_{y_{2i} > \min_j \{y_{1j}\}}$  indicates the number of subjects in arm 2 that manifest outcomes greater than the minimum outcome of subjects in arm 1.

As this is a work-in-progress, we have not completed our proof of the informativeness of this estimator. One cannot derive the asymptotic variance without knowing (or assuming) the distribution of the outcomes. One can, however, estimate the asymptotic variance employing bootstrap methods.

**Observe covariates.** Suppose the investigator observes not just  $y_{ki}$  for individual  $i$  in arm  $k$ , but also covariates  $x_{ki}$ , for  $k = 1, 2$ . Let  $\hat{\beta}_k$  be a consistent estimate of  $\beta_k$  from the regression equation  $y_{ki} = \beta_k x_{ki} + \varepsilon_{ki}$ , which is estimated on the sample of all subjects given drug  $k$  in the clinical trial, for  $k = 1, 2$ . Our proposed estimator for the option value of drug 2 is

$$\hat{V}_{21} = \frac{\sum_{i: y_{2i} > \hat{y}_{1i}} (\hat{y}_{2i} - \hat{y}_{1i})}{N_{\hat{y}_{2i} > \hat{y}_{1i}}}$$

where  $\hat{y}_{2i} = \hat{\beta}_2 x_{2i}$  and  $\hat{y}_{1i} = \hat{\beta}_1 x_{1i}$  (not  $\hat{\beta}_1 x_{1i}!$ ). This estimator provides an estimate of the option value of drug 2 conditional on covariates  $X$  observed by the investigator:

$$V_{21} = E(Y_2 - Y_1 | Y_2 > Y_1, X) \quad (4)$$

It does not provide an estimate of the option value due to information  $Z$  available only to the treating doctor. As such it is an underestimate of the full option value.

---

<sup>7</sup>This estimator requires a larger sample than the average effect estimator. It is not difficult to determine how much larger a sample is required to obtain any given level of confidence for the estimator above. We will do this in the next draft.

We will demonstrate the asymptotic properties of this estimator in the next draft.<sup>8</sup>

### 3 Application to CAST I

This section applies the two estimators of option value for data from parallel-arm trials to the Cardiac Arrhythmia Suppression Trial (CAST) I. This was a trial of encainide and flecainide for the suppression of ventricular arrhythmia in patients who had recently suffered a myocardial infarction. This trial is interesting from the perspective of option values because the patient population given active treatment experienced a higher mortality rate than the patient population on placebo.<sup>9</sup> As a result, practicing doctors are now discouraged from attempting to suppress arrhythmia following myocardial infarction with encainide and flecainide (CAST Investigators, 1989).<sup>10</sup> Nevertheless, we find that antiarrhythmia treatment may be beneficial for certain, identifiable patients.

The CAST I trial is described in detail elsewhere (CAST Investigators, 1989). Here we only provide a sketch of its design. Patients were eligible if they had suffered a myocardial infarction between six days to two years prior to the trial and a Holter recording revealed arrhythmia to the extent of six or more ventricular premature depolarizations per hour. Patients with severe symptoms due to arrhythmia or contraindications to the drugs being tested were excluded.

Patients first went through a two-week open-label titration phase during which investigators determined an appropriate dose for each active treatment (encainide, flecainide, moricizine) for each patient. Patients whose arrhythmias were successfully suppressed were randomly assigned to one of three treatment groups (encainide, flecainide, moricizine) depending upon which drug was successful at suppressing their arrhythmia. Once in this treatment group, subjects were randomized in a blinded fashion to the corresponding active treatment or placebo with equal probability.

A total of 2371 patients were enrolled in CAST I. Of these, 1498 experienced full suppression and were randomized to the encainide or flecainide treatment groups. (An additional 309 experienced full suppression and were randomized to moricizine. We ignore these patients in our analysis because they were incorporated into CAST II, a trial of moricizine versus placebo, when CAST I was aborted.) Of patients in the encainide group, 432 were randomized to active

---

<sup>8</sup>It is difficult to consider this problem, because the  $\hat{\beta}$  depends on the data. Also I feel this estimate does not estimate the value of  $V_{21}$  for the 0 and 1 response. This is because: when the first drug has success probability 0.9 and the second drug has success probability 0.8. then  $\hat{y}_{1i} = 0.9$ , and  $y_{2i} = 1$  with probability 0.8. In this case,  $N_{y_{2i} > \hat{y}_{1i}}$  is about  $0.8n$ , which is very different from  $N_{y_{2i} > y_{1i}}$  (the value of interest). This needs more thought.

<sup>9</sup>Indeed, the CAST I trial was aborted after two years because of these results. Subjects from CAST I who were initially randomized to moricizine, however, were transferred into the CAST II trial of moricizine versus placebo.

<sup>10</sup>Indeed, they are discouraged from using any drug for this purpose because of a similar experience with moricizine in the CAST II trial (Epstein et al. 1993).

treatment and 425 to placebo. In the flecainide group, 323 were randomized to active drug and 318 to placebo. Summary statistics for those given active treatment (encainide or flecainide) and those given placebo are reported in Table 1.

The primary end point for the trial was death from arrhythmia. These deaths can be categorized into three causes: non-cardiac death, cardiac death due to arrhythmia, and cardiac death not due to arrhythmia. Table 2 provides the frequency of each type of death by treatment group. Table 3 provides the Kaplan-Meier survival estimates for subjects given an active drug and those given placebo at 3, 6, 9, 12 and 15 months.

Death rates are clearly higher in the active drug group and the difference is statistically significant at the 1% level. (The same is true for death rates in the encainide group.) A logit regression of whether a subject died before censoring on all the covariates summarized in Table 1 confirms this result. The odds ratio with active treatment is 1.78 (with a standard error of 0.10). (With encainide and flecainide separately the ratios are 1.80 (0.08) and 1.74 (0.09), respectively.) The finding is also confirmed with a Cox proportional hazards model and the same covariates. The hazard ratio with active treatment is 1.84 (0.06). (For encainide and flecainide, the ratios are 1.79 (0.15) and 1.94 (0.08), respectively.) Clearly average effect estimators suggest that active drug is worse than placebo.

The logit regression reveals, however, significant heterogeneity in outcomes and in residuals. For example, non-standardized Pearson residuals, i.e.,  $y_i - p_i$ , where  $y_i$  is a 0-1 indicator for death of subject  $i$  and  $p_i$  is the predicted probability of death for subject  $i$  from the logit model, have a large standard deviation (0.23) relative to the standard deviation of actual outcomes (0.1). Moreover, the McFadden  $R^2$  statistic for the logit model is only 0.22.

The third and fourth columns of Table 4 provides estimates of bounds on the option value of active drug, as well as encainide and flecainide in particular. Bounds are not particularly informative in the case of binary outcomes, however, because the upper bound, which assumes that placebo always results in death, is simply the survival probability with the new drug. The lower bound is always zero because it assumes that placebo always cures the patient.

Implementation of the third estimator provides more meaningful information in the case of binary outcomes. The results are given in columns five through seven of Table 4. Column five indicates the number of subjects who are predicted to have superior outcomes in each treatment group.<sup>11</sup> Column six gives the average gain in survival probability among those expected to gain from treatment. Column seven gives the total gain averaged among all subjects in the population. It suggests that the option value from information available

---

<sup>11</sup>The fact that the number of subjects who gain from encainide is greater than the number who would benefit from any active drug is a product of variation in estimates of the effect of covariates across samples. The estimate of option value for the active drug is based on coefficient estimates employing first the active drug and then placebo. The estimate of option value for encainide (flecainide) is based on coefficient estimates employing first active-encainide (active-flecainide) sample and then the corresponding placebo sample.

to CAST Investigators was approximately a 1% gain in survival for all patients who suffer a myocardial infarction.

## 4 Limitations of option value approach

Our analysis thus far has focused on the cost — in terms of lost option value — of average treatment effect-based drug approval or coverage rules. In this section we attempt to resurrect those rules by pointing out some of the benefits of those rules and the costs of decision criteria that depend on option value as defined in (1). The benefits of average effect rules are straightforward: they are familiar and relatively simple to implement. As we shall see, they make few demands on doctors in terms of information gathering and optimal decisionmaking.

There are two difficulties with switching to an option-value based decision-making. First, doctors may not possess sufficient information to extract the full option value of a new drug. In particular, doctors do not have access to the information set  $W$ . Recall that this is, by definition, information that predicts heterogeneity in treatment outcomes but is not available to doctors. Any option value from such heterogeneity is left on the table.<sup>12</sup>

A second difficulty is that doctors may not have the skill or resources to allocate every patient to the optimal treatment even with full information  $\Omega$ . Evidence comes from, for example, a 1999 Institute of Medicine study that suggests that as many as 98,000 patients die each year due to medical errors in hospitals. It is unclear whether the prevalence of medical errors means that the FDA should rely on average treatment effect-based rules rather than option value-based rules. Average effect rule takes a rather pessimistic view of decisionmaking by practicing physicians. To see this, note that the difference-in-means estimator can be written

$$E(y_1 - y_2) = E(y_2 - y_1 | y_2 > y_1) - E(y_1 - y_2 | y_1 > y_2)$$

The estimator deducts from the option value of the new drug the lost option value of the conventional treatment. In other words, the estimator assumes doctors are completely unable to distinguish when to use the new drug versus conventional treatment and will always use the new drug if available. Even if medical errors are widespread, surely they are not this widespread.

## 5 Conclusion

Our discussion suggests that, whereas average effect estimators underestimate the actual or realized value of adding a new drug to physicians' portfolios, option value estimators may overestimate this value. A question that naturally follows

---

<sup>12</sup>One solution is to redefine the option value of new drug to exclude the value from  $W$ . The problem is that this new definition implies that the first two estimators in Section 2 are overestimates of the option value a new drug because they capture all heterogeneity in treatment effects.

is whether there is a superior, third alternative. We conclude this draft with a proposal for a new trial design intended to more accurately estimate the realized option value of a new drug.

Ordinary clinical trials randomize patients across a new drug and some control. The implication is that patients only experience one treatment, hence no option value. Moreover, they are not subject to (imperfect) physician decision-making, which may reduce that option value. A better approach would expose subjects to both: different portfolios and a physician’s discretion. Instead of randomizing subjects to treatments, we propose that doctors be randomized to different drug portfolios and that investigators track health outcomes among all patients of participating doctors.

More precisely, we propose that individual medical practices be randomized into two groups, one which is given access to the new drug and one which is not. Both groups are permitted to use all conventional treatments. Patients of doctors in the first, treatment group will be given the new drug if their doctor prescribes it. Patients of doctors in the second control group will not. In order to secure the full benefits of randomization, doctors in the treatment group must not be permitted to give the new drug to doctors outside their practice. Moreover, patients must not be told about the new drug lest they self-select into practices in the treatment group.

Assuming randomization is adequate, investigators will observe outcomes  $\{y_1, \dots, y_n\}$  among patients of doctors in the treatment group and  $\{y_{n+1}, \dots, y_m\}$  among patients of doctors in the control group. Assume for simplicity that there are only two treatments: the new drug and one conventional treatment. The expected value of outcomes in the first group is a consistent estimate of  $E(Y_1|g(y_1, y_2, X, Z) < 0) + E(Y_2|g(y_1, y_2, X, Z) > 0)$ , where  $g(y_1, y_2, X, Z) > 0$  embodies the decisionmaking criteria doctors use to assign patients to the new drug. The expected value of outcomes in the second group is a consistent estimate of  $E(Y_1)$ . The difference yields a consistent estimator of  $E(Y_2 - Y_1|g(y_1, y_2, X, Z) > 0)$ , which is the realized option value of the new drug.

There are obvious difficulties with this randomized-doctors design. It is more complicated and expensive to implement than randomized-patient designs. It may be worth the cost, however, in cases where preliminary investigation suggests that the mean effect of a new drug is inferior to the regulatory baseline and that treatment heterogeneity is significant.

## References

1. Aakvik, Arild, James J. Heckman, Edward J. Vytlaçil (2000). Treatment Effects for Discrete Outcomes when Responses to Treatment Vary Among Observationally Identical Persons: An Application to Norwegian Vocational Rehabilitation Programs. NBER Working Paper 262.
2. Abboud, Leila (2004). Drug Makers Seek to Bar ‘Placebo Responders’ From Trials. Wall Street Journal, June 18, 2004: B1, B5.

3. Cardiac Arrhythmia Suppression Trial Investigators (1989). Preliminary Report: Effect of Encainide and Flecainide on Mortality in a Randomized Trial of Arrhythmia Suppression after Myocardial Infarction. *New England Journal of Medicine*, 321(6): 406-412.
4. Echt, Debra S., Philip R. Liebson, L. Brent Mitchell, Robert W. Peters, Dulce Obias-Manno, Allan H. Barker, Daniel Arensberg, Andrea Baker, Lawrence Friedman, H. Leon Greene, Melissa L. Huther, David W. Richardson, and the CAST Investigators (1991). Mortality and Morbidity in Patients Receiving Encainide, Flecainide, or Placebo. *New England Journal of Medicine*, 324(12): 781-788.
5. Epstein, Andrew E., J. Thomas Bigger, Jr., D. George Wyse, Donald Romhilt, Robin A. Reynolds-Haertle, Alfred P. Hallstrom, and the CAST Investigators (1991). Events in the Cardiac Arrhythmia Suppression Trial (CAST): Mortality in the Entire Population Enrolled. *Journal of the American College of Cardiology*, 18(1): 14-19.
6. Epstein, Andrew E., Alfred P. Hallstrom, William J. Rogers, Philip R. Liebson, A. Allen Seals, Jeffrey L. Anderson, Jerome D. Cohen, Robert J. Capone, and D. George Wyse (1993). Mortality Following Ventricular Arrhythmia Suppression by Encainide, Flecainide, and Moricizine After Myocardial Infarction: The Original Design Concept of the Cardiac Arrhythmia Suppression Trial (CAST). *Journal of the American Medical Association*, 270(20): 2451-2454.
7. Harris, Gardner (2004). 2 Cancer Drugs, No Comparative Data. *New York Times*, Feb. 26, 2004: <http://www.nytimes.com/2004/02/26/business/26drug.html?pagewanted=print&position=>.
8. Heckman, James (2001). Accounting for Heterogeneity, Diversity and General Equilibrium in Evaluating Social Programmes. *Economic Journal*, 111: F654-F699.
9. Heckman, James J., Jeffrey Smith, and Nancy Clements (1997). Making the Most Out of Programme Evaluations and Social Experiments: Accounting for Heterogeneity in Programme Impacts. *Review of Economic Studies*, 64(4): 487-535.
10. Horowitz, Joel L., and Charles F. Manski (2000). Nonparametric Analysis of Randomized Experiments with Missing Covariate and Outcome Data. *Journal of the American Statistical Association*, 95(449): 77-83.
11. Jost, Timothy S. (2000). The Globalization of Health Law: The Case of Permissibility of Placebo-Based Research. *American Journal of Law and Medicine*, 26: 175-186.
12. Plewis, Ian (2002). Modelling impact heterogeneity. *Journal of the Royal Statistical Society Series A*, 165(Part 1): 31-38.



Table 1: Summary statistics for CAST I Trial.

	Active drug	Plac- ebo		Active drug	Plac- ebo
Age (yrs)	60.8	61.3	<b>History after MI (share)</b>		
	9.3	9.3	Thrombolytic therapy	0.28	0.25
Male (share)	0.81	0.83		0.45	0.43
	0.39	0.37	PTCA	0.19	0.18
Ejection fraction	0.40	0.39		0.39	0.39
	0.10	0.09	CABG	0.18	0.18
VPD/hr	135.5	130.8		0.39	0.39
	262.2	252.8			
			<b>Baseline examination</b>		
<b>Baseline ECG</b>			Sitting HR (beats/min)	73.6	73.1
PR Interval (s)	0.17	0.17		12.2	12.3
	0.03	0.03	Systolic BP (mm Hg)	115.4	115.4
	734	728		7.4	7.4
QRS Interval (s)	0.09	0.09	Diastolic BP (mm Hg)	76.8	76.1
	0.02	0.02		10.2	10.1
	754	742	Smoker (share)	0.40	0.39
				0.49	0.49
<b>History before MI (share)</b>			<b>Concurrent drugs (share)</b>		
Congestive heart failure	0.14	0.11	Beta-adrenergic blocker	0.31	0.33
	0.34	0.32		0.46	0.47
Angina	0.46	0.45	Calcium channel blocker	0.52	0.50
	0.50	0.50		0.50	0.50
Hypertension	0.29	0.32	Digitalis	0.21	0.18
	0.46	0.47		0.41	0.39
Diabetes	0.20	0.21	Nitrate	1.52	1.56
	0.40	0.40		0.50	0.50
Cardiac arrest	0.02	0.02	Diuretic	1.69	1.69
	0.15	0.15		0.46	0.46
VT	0.03	0.02			
	0.16	0.15			
Prior MI	0.35	0.37	Died before censoring (share)	0.08	0.05
	0.48	0.48		0.27	0.21
CABG	0.13	0.14			
	0.34	0.34			
PTCA	0.04	0.04			
	0.20	0.21			

Notes. First and second elements of each cell are the mean and standard deviation of the row variable. Unless otherwise indicated (by the third element of a cell), sample size for the placebo group is 743 and for active treatment is 755.

Table 2: Health outcomes (mortality) prior to censoring in CAST I.

	Encainide group		Flecainide group		Both groups		Total
	Active drug	Pla-cebo	Active drug	Pla-cebo	Active drug	Pla-cebo	
<b>Cause</b>							
Non-cardiac death	2	5	1	1	3	6	9
Death not due to arrhythmia	11	6	4	4	15	10	25
Death due to arrhythmia	29	12	15	7	44	19	63
<b>Total deaths</b>	42	23*	20	12	62	35*	97

Notes. For the total deaths category, an (\*) indicates that the difference in deaths with placebo and active drug is statistically significant at the 1% level.

Table 3: Kaplan-Meier survival estimates in active drug and placebo groups.

	Time (days)	Patients at risk	Deaths	Survivor function	Std Error	[95% Conf. Int.]	
Placebo	0	743	0	1	.	.	.
	91	641	7	0.9899	0.0038	0.979	0.9952
	182	537	7	0.9784	0.0057	0.9638	0.9872
	273	432	6	0.9662	0.0075	0.9477	0.9782
	364	321	7	0.9477	0.0101	0.9238	0.9643
	455	202	3	0.9377	0.0116	0.9106	0.9568
Active drug	0	755	0	1	.	.	.
	91	640	24	0.9668	0.0067	0.9509	0.9776
	182	528	18	0.938	0.0093	0.9169	0.9539
	273	415	6	0.9262	0.0104	0.9029	0.944
	364	303	8	0.9059	0.0124	0.8784	0.9274
	455	200	3	0.8947	0.0138	0.864	0.9188

Table 4: Estimates of option value of encainide, flecainide, and either drug.

Treatment	Second estimator			Third estimator		
	Sample size (subj. on active drug only)	Positive increment in survival probability		No. of subj. with better outcome on treatment	Average increment in survival probability	
		Upper bound	Lower bound		Subj. with positive gain	All subj.
Active drug	755	0.92	0	392	0.04	0.0141
Encainide	432	0.90	0	442	0.05	0.0189
Flecainide	323	0.94	0	181	0.40	0.0093