Title

Estimating Causal Effects of New Treatments Despite Self-Selection: The Case of Experimental Medical Treatments

Author

Hazlett, Chad

Chad Hazlett*

Estimating Causal Effects of New Treatments Despite Self-Selection: The Case of Experimental Medical Treatments

Abstract: Providing terminally ill patients with access to experimental treatments, as allowed by recent "right to try" laws and "expanded access" programs, poses a variety of ethical questions. While practitioners and investigators may assume it is impossible to learn the effects of these treatment without randomized trials, this paper describes a simple tool to estimate the effects of these experimental treatments on those who take them, despite the problem of selection into treatment, and without assumptions about the selection process. The key assumption is that the average outcome, such as survival, would remain stable over time in the absence of the new treatment. Such an assumption is unprovable, but can often be credibly judged by reference to historical data and by experts familiar with the disease and its treatment. Further, where this assumption may be violated, the result can be adjusted to account for a hypothesized change in the non-treatment outcome, or to conduct a sensitivity analysis. The method is simple to understand and implement, requiring just four numbers to form a point estimate. Such an approach can be used not only to learn which experimental treatments are promising, but also to warn us when treatments are actually harmful – especially when they might otherwise appear to be beneficial, as illustrated by example here. While this note focuses on experimental medical treatments as a motivating case, more generally this approach can be employed where a new treatment becomes available or has a large increase in uptake, where selection bias is a concern, and where an assumption on the change in average non-treatment outcome over time can credibly be imposed.

Keywords: Non-randomized trials, Observational studies, Clinical trials

1 Introduction

On 30th May 2018, the United States established a federal "right to try" law, allowing terminally ill patients to access experimental medical treatments that have cleared Phase 1 testing but were not yet approved by the Food and Drug Administration (FDA).[1] Such laws extend pre-existing methods of gaining access to unapproved treatments through "compassionate use" or "expanded access" programs, but sidestepping FDA petition and oversight procedures. Numerous ethical objections have been made to these laws, pointing to risks of dangerous side-effects that could shorten or worsen lives, raising false hopes among vulnerable patients, enabling quackery, undermining the FDA, and imposing a financial burden on desperate patients and their families.

Can we – and should we – learn anything about the efficacy and safety of drugs from those taking such experimental treatments? The first reaction of clinicians and statisticians alike may be an emphatic "no": without randomization into treatment and control groups, individuals will self-select into treatment, thus little can be learned from the observational results. Indeed, one criticism of these laws have been that they "will only make it more difficult to know if medication is effective or safe [1].

---

1 This law federally enshrines analogous rights previously recognized by 40 US states.

*Corresponding author: Chad Hazlett, Departments of Statistics and Political Science, University of California Los Angeles, Los Angeles, United States, e-mail: chazlett@ucla.edu, URL: http://www.chadhazlett.com

Yet, failing to carefully assess the effects of these treatment would both forgo the opportunity to learn which are promising, and perhaps worse, could greatly amplify their harm. Clinicians and investigators will surely make inferences about the effects of new treatments by comparing, at least casually, those who receive them and those who do not. Estimates based on comparisons of this type are not just biased, they are dangerous. As illustrated below, through the effects of self-selection, treatments that are actually harmful may appear to be beneficial in such naive comparisons, perversely encouraging more patients to take them.

We thus have a responsibility to learn the benefits and harms of such treatments, and to avoid that worst errors that naive comparisons would generate. Fortunately, a very simple technique outlined here allows valid estimates of treatment effects under such circumstances in which individuals self-select into treatment in unknown ways. Of course, such inference is not free. The critical assumption required is that in the absence of the new treatment of interest, the average outcome would remain stable or change by a specified amount across two time periods (one before the treatment is available, and one after). While not true in every case, this assumption is straightforward to understand and can often be assessed by experts familiar with treatment of the disease in question.

While this note takes experimental medical treatments as an illustration and motivation, the method described here applies to a wide variety of circumstances where a treatment becomes newly available, individuals or units opt into taking that treatment, and the assumption of stability in average non-treatment outcomes over time is reasonable.[2] The method may also be useful even in cases where a randomized trial is possible or has been conducted, but that trial had strict eligibility criteria or low willingness to consent to randomization, resulting in an estimate that may not generalize well to the population that would actually elect to take the treatment.

2 Proposed approach

Perhaps surprisingly, we can make inferences about how those who take a new treatment benefit from it even when patients select into taking treatment in unknown ways. Consider person i, with some observed outcome Y_i, such as whether or not they survive at one year post diagnosis. Using the potential outcomes framework [2] and assuming no interference, we consider not only the observed outcome for person i, but also two potential outcomes – the outcome she would have experienced had she been treated, $Y_i(1)$, and the outcome she would have experienced had she not been treated at $Y_i(0)$.

We next consider how the average outcome people would experience without the treatment, $\mathbb{E}[Y_i(0)]$ (or simply $\mathbb{E}[Y(0)]$) changes from the time period before the new treatment becomes available to the time period after,

$$\mathbb{E}[Y(0)|T = 1] - \mathbb{E}[Y(0)|T = 0] = \delta \tag{1}$$

where $T = 0$ designates a time period before a treatment (D) is available, $T = 1$ is a time period after it is made available. Note that $T \in \{0, 1\}$ designates time periods or windows – perhaps a few years wide – and not single points in time. This is important, as Y_i may take time to measure (e. g. the proportion surviving for one year post-diagnosis), and treatments may take time to take effect.

The core assumption required is on the value of δ. For simplicity of exposition and because it is likely the primary use case, we consider first the assumption that $\delta = 0$, which we call the *stability of the average non-treatment outcome* assumption. However, alternative assumed values of δ can be employed. Within this setting, besides the practical assumption that at least some eligible individuals take the new treatment,

---

2 Some examples of other cases where this approach may be applicable and where randomization is difficult or impossible include: What is the effect of a drug newly being used off-label by select physicians to treat a disease? What is the effect of warnings sent to a patient's doctor by a health monitoring system? Outside of medicine, what is the effect of an early release program from jail, assigned by judges, on recidivism of those who are released? What is the effect of legal aid, given to those who most need it, on legal outcomes? What is the effect of a television program on behaviors of those who choose to watch it, or the effect of advertisements on those who receive them?

$Pr(D = 1|T = 1) > 0$, only an assumption on δ is required to identify the average treatment effect among the treated. No assumption is required on the treatment assignment mechanism.

Showing this identification result is straightforward: we know the average non-treatment outcome among the whole group in period one, because it is equal to the mean observed (non-treatment) outcome in period zero (shifted by δ if $\delta \neq 0$). This group average is, in turn, a weighted combination of two other averages: the average non-treatment outcome among the untreated, which we observe, and the average non-treatment outcome among the treated, for which we can solve. That is, the average (non-treatment) outcome among the non-treated, combined with knowledge of the average non-treatment outcome among the whole group, tells us what the average non-treatment outcome among treated must be. Formally, by applying the law of iterated expectations,

$$\begin{aligned}
\mathbb{E}[Y(0)|T = 0] &= \mathbb{E}[Y(0)|T = 1] - \delta \\
&= \mathbb{E}[Y(0)|D = 1, T = 1]Pr(D = 1|T = 1) \\
&\quad + \mathbb{E}[Y(0)|D = 0, T = 1]Pr(D = 0|T = 1) - \delta
\end{aligned}$$

which we can re-arrange to identify $\mathbb{E}[Y(0)|D = 1, T = 1]$ in terms of observables,

$$\begin{aligned}
\mathbb{E}[Y(0)|D = 1, T = 1] &= \frac{\mathbb{E}[Y(0)|T = 0] - \mathbb{E}[Y(0)|D = 0, T = 1]Pr(D = 0|T = 1) + \delta}{Pr(D = 1|T = 1)} \\
&= \frac{\mathbb{E}[Y|T = 0] - \mathbb{E}[Y|D = 0, T = 1]Pr(D = 0|T = 1) + \delta}{Pr(D = 1|T = 1)}
\end{aligned} \tag{2}$$

Finally, the Average Treatment Effect on the Treated (ATT) is the difference between the treatment and non-treatment potential outcomes, taken solely among the treated, i. e. $\mathbb{E}[Y(1)|D = 1, T = 1] - \mathbb{E}[Y(0)|D = 1, T = 1]$. While we directly observe an estimate of the first quantity, the second term – the average outcome among the treated had they not taken the treatment – has now been given by the strategy above (Equation 2). We have thus identified the ATT,

$$\begin{aligned}
ATT &= \mathbb{E}[Y(1)|D = 1, T = 1] - \mathbb{E}[Y(0)|D = 1, T = 1] \\
&= \mathbb{E}[Y|D = 1, T = 1] - \left(\frac{\mathbb{E}[Y|T = 0] - \mathbb{E}[Y|D = 0, T = 1]Pr(D = 0|T = 1) + \delta}{Pr(D = 1|T = 1)} \right)
\end{aligned} \tag{3}$$

When the stability assumption is maintained, $\delta = 0$ and can thus be simply removed. In that case, despite appearing quite different, this estimator is equivalent to an instrumental variables approach in which "time" is the instrument and non-compliance is one-sided (see *Discussion*).

2.1 Extensions

A number of extensions are possible. First, investigators could use this tool to examine "what-if" scenarios: if clinicians have beliefs about the four quantities required here or evidence from past cases, we can compute ATT estimates to understand the underlying effect implied by those beliefs. Such an analysis is informal and only as good as the guesses that are used as data. However, it correctly produces an ATT estimate subject to those guesses, avoiding the errors that result from naive comparisons that may otherwise be made. An illustration of such errors is given below.

Second, the assumption that $\delta = 0$ can be replaced with one that allows a hypothetical or modeled change in the average non-treatment outcome over time such that $\delta \neq 0$. One application for this is sensitivity analysis: we can hypothesize shifts in the average non-treatment outcome and compute the corresponding effect, repeating this for different hypothesized shifts. This allows us to ask "how large a shift in the average non-treatment outcome must be permitted in order for our conclusions about the benefit (or harmfulness) of a treatment to change?" If a seemingly implausible shift is required to change our substantive conclusion – e. g. that non-treatment outcomes improved by 20 % despite no known changes in treatment protocols or compositional shifts in those who get the disease – then we would be able to rule out such concerns. These

analyses may or may not prove informative, but are an improvement over the naive comparisons that may otherwise be attempted. Alternatively, the stability assumption can be replaced with a known or estimated shift in non-treatment outcomes. If, for example, there are known compositional shifts or changes in available care besides the treatment of interest, and if we can model these or make reasonable estimates of how they may change average non-treatment outcomes, we can use this information to adjust the resulting treatment effect estimate for the treatment of interest.

3 Illustration: A harmful drug that appears to be beneficial

A simple illustration can demonstrate the method and how it can avoid the most dangerous errors that may arise due to direct comparisons that practitioners may be tempted to make. Suppose an aggressive form of cancer, once diagnosed at a certain stage, has only a 50 % one-year survival rate as measured over a period from 2005 to 2010 ($\widehat{\mathbb{E}}[Y(0)|T = 0] = 0.5$, where $\widehat{\mathbb{E}}[\cdot]$ designates a sample average). Between 2012 and 2016 suppose a new treatment becomes available, and that 30 % of the group diagnosed with this cancer after 2010 attempt this new treatment ($\widehat{Pr}(D = 1|T = 1) = 0.3$). Among this group, suppose that the one-year survival rate is 60 % ($\widehat{\mathbb{E}}[Y(1)|D = 1, T = 1] = 0.6$), while the one-year survival rate among those who chose not to take the treatment is only 40 % ($\widehat{\mathbb{E}}[Y(0)|D = 0, T = 1] = 0.4$). Under the stability assumption, we set $\delta = 0$ and can estimate the expected non-treatment outcome among those who took the treatment:

$$\mathbb{E}[Y(0)|T = 1, D = 1] = \frac{\mathbb{E}[Y|T = 0] - \mathbb{E}[Y|D = 0, T = 1]Pr(D = 0|T = 1)}{Pr(D = 1|T = 1)}$$
$$\approx \frac{0.5 - 0.4(0.7)}{0.3} = 0.73$$

To verify this result while reinforcing the intuition: first, under the stability assumption, if we could see how everybody in the later group fairs in the absence of the treatment, we know that (up to finite sample error) 50 % would survive at one year. Among the 70 % who did not take the treatment, only 40 % survive at one year. This "drop" in survival rate among the non-treated signals that it must have been those who were worse-off who chose not to take the treatment. Consequently, those who do take the treatment must have had higher (non-treatment) survival in order to bring this 40 % up to the required 50 % for the whole group. Specifically, the 30 % of the group who took the treatment must have had an average non-treatment survival of x in the equation $(0.4)(0.7) + (.3)x = .5$, which solves to $x = 0.73$.

The observed survival rate of 60 % under treatment thus no longer appears favorable compared to the 73 % who would have survived without treatment in this group, yielding an estimated ATT of 60 % – 73 % = −13 %. This finding of a harmful effect of treatment emphasizes both that this technique can return counter-intuitive results, and the ethical imperative to understand the impacts of experimental treatments. We emphasize that naive comparisons tell the opposite story: The treatment at first appears beneficial, with higher survival among those who took the drug compared to those before the drug existed (60 % versus 50 %), and higher survival among those who took it versus those who did not in the second period (60 % versus 40 %). This may persuade practitioners to recommend it, and patients to take it. Yet, it actually reduces survival by 13 % among those who take it. It would seem unethical not to make this information available to practitioners, patients, and regulators. Furthermore, if the assumption that $\delta = 0$ cannot confidently be defended, sensitivity analysis using a range of values for δ will characterize what we would conclude for any given assumption of δ.

4 Discussion

While motivated here by the problem of experimental medical treatments, this approach is quite general in its applicability to situations in which new treatments become available and individuals self-select to receive

them. The main concern investigators must keep in mind when choosing to apply this method is whether they can justify the stability assumption, or employ some value of δ other than zero. Fortunately, the stability of non-treatment outcomes is a straightforward assumption to understand, and can be well evaluated by experts in many cases. It is most plausible, first, if little has changed in the way the disease is treated over the time in question. This includes any *other* treatments that might be administered or withdrawn due to taking the new treatment in question. Second, a more subtle concern lies in compositional changes in the population who acquires the disease, such as changes in population health or competing risks, as these could also drive changes in the average non-treatment outcome. That said, if such compositional shifts do occur, they are likely to be slow-moving and so may be possible to rule out as problematic in the short-run. While no test can prove that the stability assumption (or any other assumption on δ) holds, investigators can check the stability of average outcomes over the course of many years prior to the introduction of the new treatment, which can boost the credibility that it remains stable thereafter. Altogether, if the outcome has been stable for many years prior to the introduction of the treatment in question, and there are no known changes in the use of other treatments or sudden changes in the composition of the group with the disease, then a strong case can be made for the stability of the average non-treatment outcome.

Pragmatically, only four numbers are required to form a point estimate: the estimated average outcome prior to treatment ($\mathbb{E}[Y|T = 0]$), the estimated average outcome among the treated and untreated after introduction of the treatment ($\mathbb{E}[Y|D = 1, T = 1]$ and $\mathbb{E}[Y|D = 0, T = 1]$ respectively), and the proportion who took the treatment after introduction, $Pr(D = 1|T = 1)$. In some cases, the required data may thus be available from existing sources, such as electronic medical records. In other cases, investigators may choose to run a trial of this type by design.

4.1 Comparisons to related methods

One alternative approach worth mentioning but less often applicable would be possible when we have a disease such that the prognosis in the absence of any new treatment is virtually certain, and thus the non-treatment outcome one would normally learn from a control group is already known. Suppose that nearly all individuals with a certain cancer at a certain stage die within one year (and those whose cancers do remit, if any, show no signs of their potential for remittance until it happens and thus would not have any basis for self-selecting into treatment). If a group – selected by any means – takes a new medication and has a 50 % survival rate at one year, then the improvement can reasonably be attributed to the new treatment, as we believe we know how those individuals would have faired under non-treatment, despite the absence of a control group. While possibly workable in some scenarios, such an approach is limited to cases where the outcome is nearly certain. By contrast, the approach here is more general, and recognizes that when outcomes are uncertain (such as a 50 % one year survival rate), there is non-trivial scope for self-selection.[3]

Second, this method may bear a resemblance to the Difference-in-Difference (DID) approach, but can operate in circumstances where DID is not possible, and provides a relaxation of DID in cases where it is possible. To conduct DID, we need either to measure each unit before and after (some are exposed to) treatment, or we must be able to place individuals into larger groupings that persist over time (such as states), with treatment being assigned at the level of those larger units at time $T = 1$. By contrast, the present method works even when there is no way to know if an individual observed at time $T = 0$ would have chosen treatment had they appeared at time $T = 1$. This is useful in cases such as new medical treatments: among those diagnosed with a given disease during $T = 0$, there is no way to say which would have taken the treatment at time $T = 1$, as would be required by DID.

The method is thus particularly useful where DID is not possible, however in arrangements where DID is possible (such as panel data), it provides an "adjustable" version of DID that allows prescribed deviations

---

3 Another approach that may be feasible in some circumstances occurs when a new treatment becomes available and all the individuals under study take it, leaving no control group at time $T = 1$. In this case, the method proposed here specializes to a simple cross-sectional "post-minus-pre" estimator, where identification is still achieved by assuming δ.

from the parallel trends assumption. Specifically, DID requires the parallel trends assumption, $\mathbb{E}[Y(0)|T = 1, D = 1] - \mathbb{E}[Y(0)|T = 0, D = 1] = \mathbb{E}[Y(0)|T = 1, D = 0] - \mathbb{E}[Y(0)|T = 0, D = 0]$, whereas the present method assumes $\mathbb{E}[Y(0)|T = 1] - \mathbb{E}[Y(0)|T = 0] = \delta$ for some choice δ. To understand the connection, consider that there are two ways to support a particular assumption of δ. The trend in average non-treatment outcomes over time could be different for the would-be-treated group and the would-be-control group, with the average of these trends (weighted by their population proportions) amounting to δ. Alternatively, we may propose a given δ by assuming that both the would-be-treated and would-be-control group changed by δ, in turn ensuring that the average $\mathbb{E}[Y(0)]$ across the two also changes by δ. This more restrictive claim is precisely the parallel trends assumption. Further, if we do assume parallel trends, this means we can learn the appropriate δ from the change over time in the control group alone. Setting δ to the estimated change in the control group, $\widehat{\mathbb{E}}[Y(0)|T = 1, D = 0] - \widehat{\mathbb{E}}[Y(0)|T = 0, D = 0]$, returns a value exactly equal to the DID estimate. However, if we wish to make any assumption other than parallel trends, this method allows it: any trends in the average non-treatment outcomes hypothesized for the treated and control group implies a choice of δ through their weighted average.

Third and perhaps most illuminating, in the case when $\delta = 0$ this procedure is identical to an instrumental variables approach in which "time" is the instrument. In the framework of [3], those at $T = 1$ are "encouraged" to take the treatment by its availability. When we assume $\delta = 0$, we assume $\mathbb{E}[Y(0)]$ does not change over time – thus the only way that the instrument (time) can influence outcomes is by switching some individuals ("compliers") into taking the treatment, satisfying the "exclusion restriction". The reader can verify that when $\delta = 0$ the estimator in (3) is numerically equal to the Wald estimator for instrumental variables. The proportion who take the treatment at $T = 1$ is the "compliance rate" or first stage effect. Defiers are assumed not to exist, and because of the unavailability of the treatment at time $T = 0$, non-compliance is known to be one-sided. As a consequence, the effect among the compliers is simply the average treatment effect on the treated. Going beyond the usual instrumental variable arrangement, when $\delta \neq 0$ is employed, this corresponds to allowing a prescribed violation of the exclusion restriction.[4] Accordingly, the required assumptions for this method (under the $\delta = 0$ assumption) can be partially represented by a Directed Acylic Graph (DAG) encoding an instrumental variable relationships, as in Figure 1 [4]. As with instrumental variables in general, the additional assumption of monotonicity or "no-defiers" on the $T \rightarrow D$ relationship is not represented on the DAG but must be stated. The absence of an edge between T and Y in this graph encodes the exclusion restriction corresponding to the "$\delta = 0$" case, though the method can allow for $\delta \neq 0$, not encoded on the non-parametric DAG.
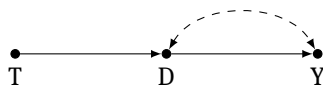

Figure 1: Graphical representation of time as an instrument. *Note:* Instrumental variables representation of the identification requirements. $T \in \{0, 1\}$ is the time period, $D \in \{0, 1\}$ is treatment status, and Y is the outcome. The required absence of defiers, and the possibility that $\delta \neq 0$, are not shown.

While I am not aware of any empirical or theoretical work describing the identification logic used here, the equivalence to using time as an instrument connects to an emergent set of medical studies in which the uptake of new treatments increases dramatically over time [5, 6, 7, 8, 9].[5] The approach described here helps to

---

4 The instrumental variables interpretation also suggests that this procedure could be used for treatments that were available in time $T = 0$ but whose uptake changed dramatically in $T = 1$, in which case the Wald estimator can be used to rescale the change in outcomes between $T = 0$ and $T = 1$ (i. e. the reduced form) by the change in treatment uptake, $Pr(D = 1|T = 1) - Pr(D = 1|T = 0)$ (i. e. the compliance rate). This may also be a reasonable strategy in some cases. In this case, however, the local average treatment effect identified is no longer the ATT.

5 Brookhart et al. [10] also provides a guide to the use of instrumental variables generally in comparative safety and effectiveness research, with a brief section discussing the use of calendar time as an instrument.

give clarity to the identification assumptions required of such work and how they can be judged. In addition, when identification depends upon an assumption on δ as described here, covariate adjustment procedures become unnecessary, or require explicit justification in terms of identification. Rather, simpler analyses using Equation 3, together with discussions of plausible values of δ are called for. Further, sensitivity analysis based on a range of these δ values can be a valuable addition to any such work.

4.2 Representational limitations of RCTs

Even when RCTs are possible, investigators may worry about two important representational limitations. For example if only a small fraction of people with a given disease are eligible or willing to consent, then how might this group be different from the ultimate target group who will use the treatment once approved? Clinical designs that allow partial self-selection in an effort to address these concern include "comprehensive cohort studies" [11] and "patient preference trials" [12]. In comprehensive cohort studies, it is proposed that those patients who refuse randomization be allowed to instead join the study but with the treatment of their preference. The randomized arms are compared as in any experiment. The outcomes (as well as the pre-treatment characteristics) of those in the preference groups are hoped to improve our understanding of generalizability, but it is unclear how to make reliable use of the information provided by these preference groups given confounding biases. In later work, "patient preference trials" encapsulate a variety of research designs in which patients' preferences are elicited, some individuals are randomized, and some receive a treatment of their choosing. In the recent proposal of [13], treatment preferences are elicited from all individuals, who are randomized into one group that will have their treatment assigned at random, and one that can choose its own treatment. This design allows for sharp-bound identification and sensitivity analysis for the average causal effects among those who would choose a given treatment.

The present method has the primary benefit of sidestepping the need for randomization, still required by the above designs. However, the allowance for self-selection alters the estimand in ways that may be preferable or complementary to an RCT, depending on the goals of the study. First, in retrospective work, the ATT identified here may be the ideal quantity of interest if we would like to know what effect a past treatment actually had. Second, if our goals are more prospective, the ATT from this method may say more about the potential effects of a new treatment in the clinical population likely to take it, if the RCT was highly restrictive in its eligibility criteria, or suffered low consent rates. On the other hand, the ATT may not be ideal to inform policy making decisions – such as promoting a new treatment to become a first line therapy – if the group likely to take the treatment under such a policy varies widely from those who have elected to take in the study period.

4.3 Conclusions

In summary, this note describes a simple identification procedure allowing estimation of the ATT regardless of self-selection into the sample. The simplest assumption – stability of average non-treatment outcomes ($\delta = 0$) – may be reasonable when we know that (a) the composition of those who acquire a certain condition has not changed, and (b) the availability and use of treatments have not changed, except for the new treatment of interest. Where investigators are uncertain that the average non-treatment outcomes have remained stable over time, one can model or propose a non-zero change ($\delta \neq 0$), or show the sensitivity of results to a range of δ values. In contrast to DID, the method works when different individuals are present in the two time periods, without any indication of who in the earlier time period would have been exposed to treatment had they been observed in the latter time period. In the $\delta = 0$ case it corresponds to using time as an instrument, clarifying the assumptions and required analyses of such an approach. In the $\delta \neq 0$ case, it allows a prescribed deviation from the exclusion restriction, as well as a sensitivity analysis. While the most obvious use of this method is when randomized trials are not possible, another potential benefit regards representation or external validity: The ATT estimated here may be more informative than the average effect from an RCT, depending upon our scientific goals, and who was able and willing to participate in the RCT.

This method is broadly applicable where treatment become newly available or popular, and an assumption on the stability of average non-treatment outcomes can be credibly made. Returning to the motivating case of "right to try" and other access to experimental medical treatments, the availability of this method does not change the deep and difficult set of ethical questions that must be answered about when and whether an experimental treatment should be made available. Rather, given the current laws, we must consider our ethical responsibility to learn what we can from such treatment regimes – not only to determine which therapies are promising for further trials, but to more quickly protect against harmful ones.

References

1. Hambine J. The disingenuousness of right to try. The Atlantic. 2018.
2. Neyman J, Dabrowska D, Speed T. On the application of probability theory to agricultural experiments. Essay on principles. Section 9. Stat Sci. 1923 [1990];5(4):465–72.
3. Angrist JD, Imbens GW, Rubin DB. Identification of causal effects using instrumental variables. J Am Stat Assoc. 1996;91(434):444–55.
4. Pearl J. Causality. Cambridge university press; 2009.
5. Johnston K, Gustafson P, Levy A, Grootendorst P. Use of instrumental variables in the analysis of generalized linear models in the presence of unmeasured confounding with applications to epidemiological research. Stat Med. 2008;27(9):1539–56.
6. Cain LE, Cole SR, Greenland S, Brown TT, Chmiel JS, Kingsley L, Detels R. Effect of highly active antiretroviral therapy on incident aids using calendar period as an instrumental variable. Am J Epidemiol. 2009;169(9):1124–32.
7. Shetty KD, Vogt WB, Bhattacharya J. Hormone replacement therapy and cardiovascular health in the united states. Med Care. 2009;600–5.
8. Mack CD, Brookhart MA, Glynn RJ, Meyer AM, Carpenter WR, Sandler RS, Stürmer T. Comparative effectiveness of oxaliplatin vs. 5-flourouricil in older adults: an instrumental variable analysis. Epidemiology (Camb, Mass). 2015;26(5):690.
9. Gokhale M, Buse JB, DeFilippo Mack C, Jonsson Funk M, Lund J, Simpson RJ, Stürmer T. Calendar time as an instrumental variable in assessing the risk of heart failure with antihyperglycemic drugs. Pharmacoepidemiol Drug Saf. 2018;27(8):857–66.
10. Brookhart MA, Rassen JA, Schneeweiss S. Instrumental variable methods in comparative safety and effectiveness research. Pharmacoepidemiol Drug Saf. 2010;19(6):537–54.
11. Olschewski M, Scheurlen H. Comprehensive cohort study: an alternative to randomized consent design in a breast preservation trial. Methods Inf Med. 1985;24(03):131–4.
12. Brewin CR, Bradley C. Patient preferences and randomised clinical trials. BMJ, Br Med J. 1989;299(6694):313.
13. Knox D, Yamamoto T, Baum MA, Berinsky A. Design, identification, and sensitivity analysis for patient preference trials. Technical report, Working Paper. 2014.