



# HHS Public Access

Author manuscript

*Stat Med.* Author manuscript; available in PMC 2022 June 15.

Published in final edited form as:

*Stat Med.* 2021 June 15; 40(13): 3124–3137. doi:10.1002/sim.8963.

## Fusion Designs and Estimators for Treatment Effects

Alexander Breskin<sup>1,2</sup>, Stephen R. Cole<sup>2</sup>, Jessie K. Edwards<sup>2</sup>, Ron Brookmeyer<sup>3</sup>, Joseph J. Eron<sup>4</sup>, Adimora A. Adimora<sup>4</sup>

<sup>1</sup>NoviSci, Durham, NC

<sup>2</sup>Department of Epidemiology, University of North Carolina at Chapel Hill, Chapel Hill, NC

<sup>3</sup>Department of Biostatistics, University of California – Los Angeles, Los Angeles, CA

<sup>4</sup>Institute of Global Health and Infectious Diseases, University of North Carolina at Chapel Hill, Chapel Hill, NC

### Abstract

While randomized trials remain the best evidence for treatment effectiveness, lack of generalizability often remains an important concern. Additionally, when new treatments are compared against existing standards of care, the potentially small benefit of the new treatment may be difficult to detect in a trial without extremely large sample sizes and long follow-up times. Recent advances in ‘data fusion’ provide a framework to combine results across studies that are applicable to a given population of interest and allow treatment comparisons that may not be feasible with traditional study designs. We propose a data fusion-based estimator that can be used to combine information from two studies: 1) a study comparing a new treatment to the standard of care in the local population of interest, and 2) a study comparing the standard of care to placebo in a separate, distal population. We provide conditions under which the parameter of interest can be identified from the two studies described and explore properties of the estimator through simulation. Finally, we apply the estimator to estimate the effect of triple- versus monotherapy for the treatment of HIV using data from two randomized trials. The proposed estimator can account for underlying population structures that induce differences in case mix, adherence, and outcome prevalence between the local and distal populations, and the estimator can also account for potentially informative loss to follow-up. Approaches like those detailed here are increasingly important to speed the approval and adoption of effective new therapies by leveraging multiple sources of information.

### Keywords

Data fusion; Causal inference; Randomized controlled trials; Generalizability; Transportability

---

Experiment remains the best judge of scientific truths. Two leading shortcomings of experimental science are that 1) results are *local* to the setting of the experiment (i.e., possible lack of generalizability), and 2) certainty of results is limited by the sample size, as

---

**Correspondance to:** Alexander Breskin, NoviSci, PMB 218, 201 W Main St, Ste 200, Durham, NC 27701, 917 – 593 – 9004, abreskin@novisci.com.

indicated through the resultant random error. Below we concentrate on an increasingly common experimental setting where both of these shortcomings play major roles.

The approval of new drugs by regulatory agencies (such as the FDA or EMA) often requires, due to appropriate ethical concerns, comparing new treatments to existing standard treatments, rather than placebo. While a new treatment may be more effective than a standard treatment, and offer an important alternative, standard treatments are typically partially effective. It is therefore often the case that new treatments provide only incremental benefit over existing, partially effective standard treatments, and therefore the demonstration of superiority (or non-inferiority) for a new treatment compared to the standard can require unfeasibly large or long studies. Unfortunately, this may lead to delays in development and approval of new treatments that provide incremental benefit over an existing partially effective standard treatment. Consider the following example.

We have a new treatment for HIV infection (e.g., three-drug antiretroviral therapy, or triple-therapy). We would like to demonstrate the effectiveness of this new treatment in a local setting. There is an existing standard treatment to prevent progression of HIV infection (e.g., two-drug antiretroviral therapy, or dual-therapy). The existing standard was tested, by experiment, against a placebo in a setting that is distal in geographical space. In our example, the distal setting differs from the local setting in at least three respects. First, the case mix of patients indicated for HIV treatment differs between the local and distal settings. Second, the adherence of HIV treatments is expected to differ between local and distal settings. Third, the risk of progressing to AIDS or death is lower in the local setting compared to the distal setting.

Ethical constraints preclude comparing the new treatment to placebo given the established effectiveness of a standard treatment. Additionally, the combination of a lower risk of HIV progression in the local setting with a (typically expected) small incremental improvement of the new treatment makes a simple direct randomized comparison of the new versus standard treatment infeasible in the local setting, as the study would require a prohibitively large sample size<sup>1</sup>. Considering these difficulties, a possible solution is to combine, or “fuse,” information from a randomized comparison of the new versus standard treatment in the local setting with existing information from a randomized comparison of the standard versus placebo treatment from the distal setting<sup>2</sup>. Here, fusion means combining information from multiple settings while accounting for important structural differences between the settings. Such a fusion of information must account for underlying causal structures that can induce differences in case mix, adherence, and HIV progression between the local and distal settings<sup>3</sup>.

In prior work, Rudolph et al. described an estimation approach that adapts a treatment comparison from a distal to the local setting while accounting for differences between settings<sup>4</sup>. This procedure, in which the study population is not a subset of the target population, is often referred to as “transporting” the effect estimate<sup>5</sup>. That prior approach required having information on both arms of the comparison of interest available from the distal setting, and only used information on case mix and adherence from the local setting. Often, however, the treatment comparison available from the distal setting is not of direct

interest, but it might be combined with additional sources of information from the local setting to estimate a different contrast involving a newly available treatment. Such an approach can allow for estimation of the effect of the new treatment compared to placebo among patients in the local setting. Here we describe an approach to obtaining evidence regarding a new treatment by fusing a chain of information from local and distal randomized studies. Through simulation, we explore several properties and operating characteristics of the proposed approach. Finally, we apply the approach to estimate the effect of triple- versus monotherapy for HIV using data from two randomized experiments conducted by the AIDS Clinical Trials Group (ACTG).

## METHODS

### Notation and the Parameter of Interest

Assume the study consists of  $n$  subjects. Let  $T$  be the time from randomization to HIV progression,  $D$  be the time from randomization to drop out,  $A$  be the time from randomization to first protocol deviation,  $\tau$  be the length of the study period, and  $T^* = \min(T, D, A, \tau)$ . If no protocol deviation occurs prior to the outcome or  $\tau$  it is assumed that treatment ends and, using notation similar to Tsiatis<sup>6</sup>,  $A = \infty$ . Let  $W$  denote setting, with  $W = d$  denoting distal and  $W = l$  denoting local. Let  $R = 0$  denote randomization to placebo,  $R = 1$  standard treatment, and  $R = 2$  new treatment. Collect measured covariates, other than  $W$ , into a vector denoted by  $X(t)$ . Drop out, protocol deviations, HIV progression, and data collection occur at discrete points in time,  $t \in \{0, 1, \dots, \tau\}$ . Let  $\delta_T(t) = 1$  if an individual has the outcome at time  $t$  (e.g.  $T = t$ ),  $\delta_A(t) = 1$  if an individual has a first protocol deviation at time  $t$  (e.g.  $A = t$ ), and  $\delta_D(t) = 1$  if an individual drops out at time  $t$  (e.g.  $D = t$ ). Finally, let  $T^{r,a}$  denote the time at which HIV progression would occur had, possibly counter to fact, the patient received treatment  $R = r$  and deviated from the protocol at time  $A = a$ . We denote the history of a random variable through  $t$  with an overbar, e.g.  $\bar{X}(t) = \{X(0), X(1), \dots, X(t)\}$ , and we denote  $\bar{X}(\tau) = \bar{X}$ . The observed data consist of independent and identically distributed samples of  $\{R, W, T^*, \bar{X}, \bar{\delta}_T, \bar{\delta}_A, \bar{\delta}_D\}$ . Note that if  $W = l$  then  $P(R = 0) = 0$ , and if  $W = d$  then  $P(R = 2) = 0$ .

The parameter of interest is a difference, at the end of the study period  $\tau$ , between the probability of HIV progression in the local setting under the protocol for new treatment  $P(T^{2,\infty} \leq \tau | W = l)$ , compared to the protocol for placebo  $P(T^{0,\infty} \leq \tau | W = l)$ .

Given the information from pending and existing randomized studies described above, respectively, along with identification conditions detailed below, first, the identified intent-to-treat parameters are

$$P(T^{2,A} \leq \tau | W = l) - P(T^{1,A} \leq \tau | W = l) \quad \text{eq. 1}$$

and

$$P(T^{1,A} \leq \tau | W = d) - P(T^{0,A} \leq \tau | W = d) \quad \text{eq. 2}$$

Note that here,  $T^{r,A}$  is the potential outcome under treatment level  $r$  and the observed time of protocol deviation. The first of these compares new treatment to standard treatment under observed adherence in the local setting, and the second compares standard treatment to placebo in the distal setting.

Second, we identify the per-protocol parameters

$$P(T^{2,\infty} \leq \tau | W = l) - P(T^{1,\infty} \leq \tau | W = l) \tag{eq. 3}$$

and

$$P(T^{1,\infty} \leq \tau | W = d) - P(T^{0,\infty} \leq \tau | W = d) \tag{eq. 4}$$

using measured information on protocol deviation times  $A$  to censor subjects when they first deviate from the protocol, and information on covariate history  $\bar{X}$  to account for possible induced informative censoring due to protocol deviation<sup>7</sup>. Here, the per-protocol effect means the effect had all individuals followed their assigned treatment protocol<sup>8</sup>.

Third, because of possible differences in the populations in the local and distal environments that can impact risk, it is necessary to compare all treatments in the same population. We therefore map eq. 4, which is defined among the population in the distal environment ( $W = d$ ), to the local environment ( $W = l$ ) using methods for generalizability (or transportability). Essentially, this amounts to a change in the population over which the parameter is defined. This mapping is achieved through the use of inverse odds of sampling weighting<sup>5</sup>. We thus obtain the parameter

$$P(T^{1,\infty} \leq \tau | W = l) - P(T^{0,\infty} \leq \tau | W = l) \tag{eq. 5}$$

Fourth, we note that eq. 3 and eq. 5, both contain the parameter  $P(T^{1,\infty} \leq \tau | W = l)$ , through which, by transitivity, we have identified a data “fusion” parameter

$$\psi = P(T^{2,\infty} < \tau | W = l) - P(T^{0,\infty} < \tau | W = l) \tag{eq. 6}$$

which is exactly the parameter of interest originally stated – a per-protocol comparison between the new treatment and the placebo in the local setting.

**Identification Conditions**

Identification of eq. 1 and eq. 2, the intent-to-treat parameters, is given by randomization. Identification of eq. 3 and eq. 4, the per-protocol parameters, can be achieved in two ways. First, one may conduct two separate trials, each involving subjects randomly sampled from the same target population, in which individuals are randomized to the new treatment vs. standard of care (in one trial) or standard of care vs. placebo (in the other trial). In each trial, individuals would be closely followed and all drop-out and protocol deviation would be prevented. Second, one may identify the per-protocol parameters by assuming a pair of exchangeability and positivity assumptions<sup>7,9</sup> and using data from two trials conducted in

different populations with possible drop-out and protocol deviation. Specifically, for eq. 3 and eq. 4, we require sequential ignorability of protocol deviations and dropout conditional on measured covariates, i.e.,

$$\delta_Z(t) \perp\!\!\!\perp T^{r, \infty} \mid \bar{X}(t), R, W, \bar{\delta}_A(t-1) = \bar{\delta}_D(t-1) = \bar{\delta}_T(t-1) = 0, \quad \forall t \in \{0, 1, \dots, \tau\}, Z \in \{A, D\} \tag{eq. 7}$$

and associated positivity conditions. In words, at each time, dropout and protocol deviations must be independent of the potential time to event conditional on covariate history, treatment, and setting, among those who have not yet experienced the outcome, dropped out, or had a protocol deviation by that time. Identification of eq. 5 is given by a pair of exchangeability and positivity assumptions for generalizability<sup>10-12</sup>. Specifically, we require ignorability of “selection” to the local setting conditional on measured baseline covariates, i.e.,

$$W \perp\!\!\!\perp T^{r, \infty} \mid X(0) \tag{eq. 8}$$

and an associated positivity condition<sup>10</sup>. In words, setting is independent of potential time to event conditional on baseline covariates. Finally, we assume independence between protocol deviations and drop out.

### Fusion Estimator

Given the identification conditions above, the risk of the outcome in the local setting under treatment  $r$  and full adherence can be rewritten as

$$P(T^{r, \infty} < \tau \mid W = 1) = E \left[ \frac{I(W = d)\pi_W(t)}{\pi_W(d)} \right]^{-1} E \left[ \frac{I(W = d)I(R = r)I(T^* \leq \tau)\delta_T(T^*)}{\pi_A(T^*)\pi_D(T^*)\pi_R} \times \frac{\pi_W(t)}{\pi_W(d)} \right],$$

where the nuisance functions, indexed by parameters  $\alpha$ ,  $\beta$ , and  $\gamma$ , are  $\pi_W(w) = P(W = w \mid X(0), R; \alpha)$ , protocol deviations  $\pi_A(t) = P[A > t \mid W, \bar{X}, R; \beta]$ , and drop out  $\pi_D(t) = P[D > t \mid W, \bar{X}, R; \gamma]$ . The parameters of the nuisance functions are summarized as  $\theta = (\alpha, \beta, \gamma)$ . We assume known randomization probability  $\pi_R = P(R \mid W)$ . This equivalence is demonstrated in discrete time as follows:

Beginning with the second expectation, we note that the indicator function  $I(T^* < \tau)$  can be rewritten as a sum of indicator functions, with at most one of those indicator functions taking value 1,

$$\begin{aligned} & E \left[ \frac{I(W = d)I(R = r)I(T^* \leq \tau)\delta_T(T^*)}{\pi_A(T^*)\pi_D(T^*)\pi_R} \times \frac{\pi_W(t)}{\pi_W(d)} \right] = \\ & E \left[ \frac{I(W = d)I(R = r) \left[ \sum_{t=1}^{\tau} I(T^* = t) \right] \delta_T(T^*)}{\pi_A(T^*)\pi_D(T^*)\pi_R} \times \frac{\pi_W(t)}{\pi_W(d)} \right] = \\ & \sum_{t=1}^{\tau} E \left[ \frac{I(W = d)I(R = r)I(T^* = t)\delta_T(T^*)}{\pi_A(t)\pi_D(t)\pi_R} \times \frac{\pi_W(t)}{\pi_W(d)} \right] \end{aligned}$$

Focusing on each element of the sum, we note that  $I(T^* = t)\delta_T(T^*) = I(T = t)I(D > t)I(A > t)$ , so

$$E \left[ \frac{I(W = d)I(R = r)I(T^* = t)\delta_T(T^*)}{\pi_A(t)\pi_D(t)\pi_R} \times \frac{\pi_W(l)}{\pi_W(W)} \right] =$$

$$E \left[ \frac{I(W = d)I(R = r)I(T = t)I(D > t)I(A > t)}{\pi_A(t)\pi_D(t)\pi_R} \times \frac{\pi_W(l)}{\pi_W(W)} \right].$$

Then by causal consistency we can replace  $T$  with the potential outcome  $T^{r,\infty}$ , so the previous quantity equals

$$E \left[ \frac{I(W = d)I(R = r)I(T^{r,\infty} = t)I(D > t)I(A > t)}{\pi_A(t)\pi_D(t)\pi_R} \times \frac{\pi_W(l)}{\pi_W(W)} \right],$$

which by randomization equals

$$E \left[ \frac{I(W = d)I(T^{r,\infty} = t)I(D > t)I(A > t)}{\pi_A(t)\pi_D(t)} \times \frac{\pi_W(l)}{\pi_W(W)} \right].$$

Using the law of total expectation

$$E \left[ \frac{I(W = d)I(T^{r,\infty} = t)I(D > t)I(A > t)}{\pi_A(t)\pi_D(t)} \times \frac{\pi_W(l)}{\pi_W(W)} \right] =$$

$$E \left[ E \left[ \frac{I(W = d)I(T^{r,\infty} = t)I(D > t)I(A > t)}{\pi_A(t)\pi_D(t)} \times \frac{\pi_W(l)}{\pi_W(W)} \middle| R, \bar{X}(t-1), \bar{\delta}_D(t-1), \bar{\delta}_A(t-1), \bar{\delta}_T(t-1), W \right] \right].$$

Next, we note that both  $I(D > t)$  and  $I(A > t)$  can be replaced by products of indicator terms as follows,

$$E \left[ E \left[ \frac{I(W = d)I(T^{r,\infty} = t) \prod_{k=0}^t I(\delta_D(k) = 0) \prod_{k=0}^t I(\delta_A(k) = 0)}{\pi_A(t)\pi_D(t)} \times \frac{\pi_W(l)}{\pi_W(W)} \right. \right.$$

$$\left. \left. \middle| R, \bar{X}(t-1), \bar{\delta}_D(t-1), \bar{\delta}_A(t-1), \bar{\delta}_T(t-1), W \right] \right].$$

With sequential ignorability, this latter quantity equals

$$E \left[ E \left[ \frac{I(W = d)I(T^{r,\infty} = t) \prod_{k=0}^{t-1} I(\delta_D(k) = 0) \prod_{k=0}^{t-1} I(\delta_A(k) = 0)}{\pi_A(t-1)\pi_D(t-1)} \times \frac{\pi_W(l)}{\pi_W(W)} \right. \right.$$

$$\left. \left. \middle| R, \bar{X}(t-1), \bar{\delta}_D(t-1), \bar{\delta}_A(t-1), \bar{\delta}_T(t-1), W \right] \right] =$$

$$E \left[ \frac{I(W = d)I(T^{r, \infty} = t)I(D > t - 1)I(A > t - 1)}{\pi_A(t-1)\pi_D(t-1)} \times \frac{\pi_W(l)}{\pi_W(W)} \right].$$

Iterating these steps  $t-1$  more times gives

$$E \left[ \frac{I(W = d)I(R = r)I(T^* = t)\delta_T(T^*)}{\pi_A(t)\pi_D(t)\pi_R} \times \frac{\pi_W(l)}{\pi_W(W)} \right] =$$

$$E \left[ I(W = d)I(T^{r, \infty} = t) \times \frac{\pi_W(l)}{\pi_W(W)} \right].$$

Next, again using the law of total expectation, we note that

$$E \left[ I(W = d)I(T^{r, \infty} = t) \times \frac{\pi_W(l)}{\pi_W(d)} \right] = E \left[ E \left[ I(W = d)I(T^{r, \infty} = t) \times \frac{\pi_W(l)}{\pi_W(d)} \middle| X(O) \right] \right].$$

From ignorability of region, the above equals

$$E[(\pi_W(l)I(T^{r, \infty} = t))],$$

Which by an application of Bayes' theorem equals

$$E \left[ \frac{P((X(0)|W = l)P(W = l))}{P(X(0))} I(T^{r, \infty} = t) \right],$$

which by the law of total expectation and region ignorability equals

$$E \left[ \frac{P((X(0)|W = l)P(W = l))}{P(X(0))} E[I(T^{r, \infty} = t) | X(0), W = l] \right].$$

By the definition of the expectation of an indicator, the three terms in the numerator of this latter quantity can be combined as

$$E \left[ \frac{P(X(0), T^{r, \infty} = t, W = 1)}{P(X(0))} \right].$$

By definition, this latter quantity is

$$E[P(T^{r, \infty} = t, W = l | X(0))],$$

Which is equal to

$$P(W = l)P(T^{r, \infty} = t | W = l).$$

Therefore,

$$\begin{aligned} \sum_{t=1}^{\tau} E \left[ \frac{I(W = d)I(R = r)I(T^* = t)\delta}{\pi_A(T^*)\pi_D(T^*)\pi_R} \times \frac{\pi_W(l)}{\pi_W(d)} \right] &= \\ \sum_{t=1}^{\tau} P(W = l)P(T^R = r, A = \infty = t | W = l) &= \\ P(W = l)P(T^R = r, A = \infty \leq \tau | W = l). & \end{aligned}$$

Finally, note that

$$E \left[ \frac{I(W = d)\pi_W(l)}{\pi_W(d)} \right]^{-1} = E[\pi_W(l)]^{-1} = P(W = l)^{-1},$$

So

$$E \left[ \frac{I(W = d)\pi_W(l)}{\pi_W(d)} \right]^{-1} P(W = l)P(T^R = r, A = \infty \leq \tau | W = l) = P(T^R = r, A = \infty \leq \tau | W = l).$$

A fusion estimator combines, or fuses, data sources. Specifically, a per-protocol fusion estimator for the risk of the outcome under treatment  $r$  and full adherence in the local setting is obtained by replacing expectations with sample averages as

$$\begin{aligned} \hat{P}_{fusion}(T^{r, \infty} < \tau | W = l; \hat{\theta}) &= \left( \sum_{i=1}^n \frac{I(W_i = d)\hat{\pi}_{W, i}(l)}{\hat{\pi}_{W, i}(d)} \right)^{-1} \sum_{i=1}^n \frac{I(W_i = d)I(R_i = r)I(T_i^* \leq \tau)\delta_{T^*}(T^*)}{\hat{\pi}_{A, i}(T_i^*)\hat{\pi}_{D, i}(T_i^*)\pi_{R, i}} \\ &\times \frac{\hat{\pi}_{W, i}(l)}{\hat{\pi}_{W, i}(d)}, \end{aligned}$$

with  $\hat{\theta} = (\hat{\alpha}, \hat{\beta}, \hat{\gamma})$  fit by maximum likelihood. For instance,  $\alpha$  can be estimated with logistic regression and  $\beta$  and  $\gamma$  can be estimated using proportional hazards models with the Breslow estimator of the baseline hazard function<sup>13</sup>. Note that the fusion occurs in the estimation of  $\alpha$ , as this parameter is used to standardize information from the distal setting to the local setting. Similarly, the local estimator for the risk of the outcome under treatment  $r$  and full adherence in the local setting is

$$\hat{P}_{local}(T^{r, \infty} < \tau | W = l; \hat{\theta}) = \left( \sum_{i=1}^n I(W_i = l) \right)^{-1} \sum_{i=1}^n \frac{I(W_i = l)I(R_i = r)I(T_i^* \leq \tau)\delta_{T^*}(T^*)}{\hat{\pi}_{A, i}(T_i^*)\hat{\pi}_{D, i}(T_i^*)\pi_{R, i}}.$$

Then the estimator for the parameter of interest is estimated with

$$\begin{aligned} \hat{\psi} &= \left( \hat{P}_{local}(T^{2, \infty} < \tau | W = l; \hat{\theta}) - \hat{P}_{local}(T^{1, \infty} < \tau | W = l; \hat{\theta}) \right) + \\ &\left( \hat{P}_{fusion}(T^{1, \infty} < \tau | W = l; \hat{\theta}) - \hat{P}_{fusion}(T^{0, \infty} < \tau | W = l; \hat{\theta}) \right). \end{aligned}$$





## SIMULATION RESULTS

As seen in Table 1 and the Figure, the proposed fusion estimator appears unbiased across both scenarios A and B, and across all sample sizes. With sample sizes of 1000 in both the local and distal settings, the bias of the fusion estimator was 0.009 in scenario A and 0.001 in scenario B, both substantially smaller than the standard error of the estimates. The naïve estimator is unbiased only in scenario A, where sampling and drop out are completely at random and there are no protocol deviations. With the same sample sizes as above, the bias of the naïve estimator was 0.008 in scenario A and  $-0.072$  in scenario B. The bias of the naïve estimator in scenario B was more than twice the magnitude of the standard error of the estimate. Root MSE is also always better for the proposed estimator in scenario B where sampling is differential and protocol deviations occur, but root MSE is lower for the naïve estimator in scenario A, where the simpler naïve estimator is consistent. With 1000 subjects in each setting in scenario A, the root MSEs were 0.05 and 0.04 for the fusion and naïve estimators, respectively. However, in scenario B, the root MSEs were 0.02 and 0.08 for the fusion and naïve estimators, respectively.

As shown in Table 2, the estimated SE for the proposed estimator appears to well approximates the Monte Carlo simulation standard error in scenario B, but is slightly greater than the Monte Carlo simulation standard error in scenario A. The maximum ratio of the estimated SE to the Monte Carlo SE for the fusion estimator was 1.40, which occurred with 100 individuals in the local setting and 1000 in the distal setting for scenario A. In all other scenarios this ratio was less than 1.2. The SE shrinks as a function of both the local and distal sample sizes, but it is more strongly related to the distal sample size in both scenarios. For the fusion estimator in scenario B, sample sizes of 100 in the low-incidence local and 1000 in the high-incidence distal setting yielded a SE of 0.043, while sample sizes of 1000 in the low-incidence local and 100 in the high-incidence distal yielded a SE of 0.039. In scenario A, in which both estimators are unbiased, the SE for the naïve estimator is always smaller than the fusion estimator.

## APPLICATION

To demonstrate the use of the proposed approach, we applied it to estimate the effect of triple- vs. monotherapy for HIV treatment. The data came from two randomized experiments conducted by the AIDS Clinical Trials Group (ACTG): ACTG 175<sup>15</sup>, which compared mono- with dual-therapy for HIV, and ACTG 320<sup>16</sup>, which compared dual- with triple-therapy for HIV. Specifically, we included subjects who were randomized to either Zidovudine (monotherapy), Zidovudine + Didanosine (dual-therapy), or Zidovudine + Zalcitabine (dual-therapy) from ACTG 175, and subjects who were randomized to Zidovudine + Lamivudine (dual-therapy) or Indinavir + Zidovudine + Lamivudine (triple therapy) from ACTG 320. We note that these trials had non-overlapping CD4+ cell count inclusion criteria, and therefore the results presented here should be not be considered meaningful clinical effects, but rather a demonstration of the proposed analytic approach. The outcome of interest was the 1-year risk of AIDS, death, or a 50% drop in CD4+ cell count.

A total of 821 subjects (33% monotherapy, 67% dual-therapy) from ACTG 175 and 1156 subjects (50% dual-therapy, 50% triple-therapy) from ACTG 320 were included. A sample of the baseline characteristics of the subjects, stratified by study, are presented in Table 3. Of note, subjects in ACTG 320 were slightly older, more likely to be of black race, more likely to inject drugs at baseline, and had a lower Karnofsky score at baseline.

Because data on adherence were not available, we did not consider non-adherence in our estimates, and instead used the proposed analytic approach to estimate the intention-to-treat effect. There remained two nuisance models to be fit – a model for the probability of being in ACTG 320 v. ACTG 175 and a model for the probability of remaining under follow-up. The former model was specified as a logistic regression adjusted for male sex, black race, injection drug use, age (modeled with B-splines with 3 knots), and Karnofsky score (modeled with B-splines with 3 knots). The latter model was specified as a Cox proportional hazards model with a Breslow estimator<sup>13</sup> of the baseline hazard, adjusted for the same variables as the previous model, but additionally stratified by treatment group. The model for dropout was fit separately for each study. For comparison, we also fit each model without adjusting for covariates in order to estimate the ‘crude’ effect of triple- versus monotherapy.

Using the proposed estimator, the crude effect was triple- versus monotherapy on the 1-year risk of AIDS, death, or a 50% drop in CD4+ cell count was a risk difference of  $-0.17$  (95% CI  $-0.23, -0.11$ ). Adjusting for male sex, black race, injection drug use, age, and Karnofsky score in each model yielded a risk difference of  $-0.20$  ( $-0.29, -0.10$ ). Of note, the standard error of the crude risk difference, 0.033, was smaller than standard error of the adjusted risk difference, 0.048, thus demonstrating the “cost” associated with using the proposed approach.

## DISCUSSION

It is possible to fuse evidence from multiple sources to estimate unbiased treatment contrasts that would not be possible using information from either source in isolation. The estimator we propose here uses information from randomized trials in local and distal settings to compare a new treatment to a placebo, even though such a comparison would be unethical in any single trial. Our proposed estimator is quite general, and can accommodate time-varying treatments and covariates, time to event outcomes, informative dropout and nonadherence, and differences in patient mix and disease risk between the local and distal setting.

Moreover, the general approach of fusion estimators and designs (combining sources of information) is widely applicable beyond the motivating example described here. Through simulation, we have found that, given the identification conditions we described, the fusion estimator appears unbiased in the settings we explored, and the proposed estimator for the standard error appears mildly conservative. Notably, for the set of scenarios explored here, the standard error of the fusion estimator was more sensitive to the sample size and number of events in the distal setting compared with the local setting, suggesting that, if data collection costs were equal, committing resources to the higher incidence distal setting may be preferred.

There is a ‘cost’, however, to using the fusion estimator. In particular, in scenarios where the fusion estimator is not necessary for unbiased effect estimation, such as our simulation scenario A, the standard error of the fusion estimator is larger than the standard error of the naïve estimator. The confidence afforded by the smaller standard error may be misplaced, however, as it is generally not possible to determine from the data alone when the naïve estimator will be unbiased. The loss of precision of the fusion estimator is thus the cost of insurance to avoid bias when fusing data sources.

As demonstrated by our applied example, implementation of the fusion estimator is straightforward and can be achieved by using many standard procedures from readily available software. In our example, the estimator required fitting logistic regressions and Cox proportional hazard models, making predictions from the models, and combining the predictions to form the estimator. In this example, the estimate using the proposed approach was slightly larger in magnitude than the crude estimate, albeit with a wider confidence interval. Additionally, the risk difference of  $-0.20$  is notably larger than the risk differences reported originally by each trial, where the effects of incremental improvements to treatment were estimated. By fusing results across trials, we were able to estimate a parameter that would not have been feasible using data from any available individual trial. Notably, a trial comparing triple- to monotherapy would not have been ethical to conduct due to knowledge that monotherapy is inferior to dual-therapy. The fusion approach therefore allowed for the estimation of a parameter that could not have been estimated easily otherwise.

We note that the conditions under which we identified the parameter of interest are sufficient but not necessary. Specifically, our derivation relied on conditional exchangeability with respect to the outcome across study locations. Essentially, this means that the risk of the outcome is the same in each location within levels of covariates. This is a stronger condition than is needed, as it may be possible to identify the parameter of interest by only requiring the effect measure of interest, rather than the risk itself, be the same in each location within levels of covariates. Identification under this weaker condition may allow the proposed approach to be used in a wider set of scenarios, but it should be noted that it also requires parametric model form knowledge that may not be available in all settings.

This work builds extensively on existing literature related to generalizing, transporting, and fusing effect estimates. We often refer to generalizing causal effects when the study population is a subset of the target population. Prior work by Lesko et al. described sufficient conditions estimators, expressed in terms of potential outcomes, for generalizing study results to a target population<sup>10</sup>. When the study population is not a subset of the target population, we often refer to “transporting” effects from one setting to another. In prior work by Rudolph and van der Laan<sup>17</sup>, conditions, identification results, and estimators (including semiparametric efficient, doubly robust estimators) for transporting effects are described, however these results are limited to single time point treatment and outcomes and do not account for dropout or non-adherence. Rudolph et al. present additional results that also allow for differences in non-adherence between the study and target populations<sup>4</sup>. When both generalizing and transporting results, no experimental evidence from the local setting is used. An alternative to the potential outcomes framework used here is the graphical framework for transporting and fusing information established by Pearl and Bareinboim<sup>182</sup>,

however, the literature on these graphical approaches has not yet included estimators for data fusion parameters or demonstrated their operating characteristics through simulation.

Our proposed approach is subject to limitations. First and foremost, the identification results rely on a set of untestable conditions described in equations 7 and 8. Though graphical approaches have been developed to aid in choosing sufficient adjustment sets and establishing identifiability<sup>18</sup>, the validity of the underlying causal structure is not testable given the data typically observed. Second, the proposed estimator relies on correctly specified models for the nuisance parameters  $\pi_A$ ,  $\pi_D$ ,  $\pi_W$ . If any of these models are misspecified, then the proposed estimator may be biased. A possible extension would be the development of multiply robust estimators that use additional nuisance parameters and are consistent if models for certain subsets of those parameters are correctly specified. Additionally, these multiply robust estimators achieve fast convergence rates and thus allow for extremely flexible, data-adaptive estimation of the nuisance parameters, enhancing the likelihood of correct model specification. Doubly robust methods have been developed for estimating effects of time-varying treatments on time-to-event outcomes in a study population<sup>19,20</sup> and for transporting the effects of single time point treatments on single time point outcomes<sup>17</sup>, but to our knowledge no doubly robust methods have been developed for the setting described in the present work. Third, our estimator for the standard error treats the nuisance parameters as known and therefore may deviate from the true standard error of the estimator. Correct standard errors that account for the estimation of the nuisance parameters could be estimated using stacked estimating equations<sup>21</sup> or using the bootstrap<sup>22</sup>. Fourth, our estimator can only be used to estimate per-protocol effects, not intention to treat effects. The approach described by Rudolph et al.<sup>4</sup> that uses an outcome-model based approach could be adapted to estimate such effects. Finally, further work is needed to determine what improvements in statistical power can be achieved by using the fusion approach to compare a new versus placebo treatment over a traditional comparison of a new versus standard treatment in the local setting. The relative power of each approach is likely strongly related to the relative effect sizes of each treatment and the characteristics of the local and distal populations, and thus extensive simulations will be needed to characterize the power of each approach. We anticipate that in many scenarios the fusion approach can be used to overcome issues related to incremental treatment effects that may lead to the need for untenably large studies by allowing new treatments to be compared with a placebo rather than existing standard treatments.

In addition to the scenario we introduced in the introduction, fusion estimators and designs may be widely applicable in settings of diseases that are changing rapidly with respect to case mix and geography, such as COVID-19. Trials for treatments or vaccines may take place in one location, but because of the movement of the disease, newly arising outbreaks, and multiple waves of infections, subsequent trials for other treatments will likely be conducted in new locations where incidence is highest. As new treatments are discovered and become standard, each subsequent trial will need to use the most recent standard identified from the previous location in the control arm. Fusion designs and estimators are well-equipped for such a scenario to allow information to be consistently fused across locations to facilitate the rapid development of new treatments for emerging, fast-moving diseases.

The proposed approach allows for the fusion of information across settings to estimate contrasts of interest that would not be feasible with standard study designs and analyses. In the era of ‘real-world evidence’ and ‘big data,’ such approaches will become increasingly important to speed the approval and adoption of effective new therapies by leveraging multiple sources of information. Additionally, given the high cost of conducting clinical trials, the ability to reuse the information generated and augment it with results from smaller local studies is an attractive approach to assessing treatment effects in populations that may not be included in large studies. Fusion approaches, such as the one presented here, provide an important path towards maximizing the value, accuracy, and impact of clinical research.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Appendix

Let  $\text{expit}(x) = \frac{e^x}{1 + e^x}$  and  $I(x)$  be the indicator function. The variable distributions for the simulation are:

### Scenario A:

$$P(R = 0 | W = l) = 0$$

$$P(R = 1 | W = l) = 0.5$$

$$P(R = 2 | W = l) = 0.5$$

$$P(R = 0 | W = d) = 0.5$$

$$P(R = 1 | W = d) = 0.5$$

$$P(R = 2 | W = d) = 0$$

$$P(X_1 = 1 | W) = \text{expit}(-1 + 2 * I(W = d))$$

$$P(X_{2,t=1} = 1) = 0.5$$

$$P(X_{2,t=2} = 1 | X_{2,t=1}) = \text{expit}(0.3 + 0.3 X_{2,t=1})$$

$$P(C_{t=1} = 1) = 0.1$$

$$P(C_{t=2} = 1 | A_{t=1} = C_{t=1} = Y_{t=1} = 0) = 0.1$$

$$P(A_t = 1 | W, R, X_{2,t}) = 0$$

$$P(Y_{t=1}^{r,\infty} | W, X_1, X_{2,t=1}) = \text{expit}(-3 + 0.5 * I(R = 1) + 1.5 * I(R = 0))$$

$$P(Y_{t=2}^{r,\infty} | W, X_1, X_{2,t=2}, Y_{t=1}^{r,\infty} = 0) = \text{expit}(-3 + 0.5 * I(R = 1) + 1.5 * I(R = 0))$$

### Scenario B:

$$P(R = 0 | W = l) = 0$$

$$P(R = 1 | W = l) = 0.5$$

$$P(R = 2 | W = l) = 0.5$$

$$P(R = 0 | W = d) = 0.5$$

$$P(R = 1 | W = d) = 0.5$$

$$P(R = 2 | W = d) = 0$$

$$P(X_1 = 1 | W) = \text{expit}(-1 + 2 * I(W = d))$$

$$P(X_{2,t=1} = 1) = 0.5$$

$$P(X_{2,t=2} = 1 | X_{2,t=1}) = \text{expit}(0.3 + 0.3 X_{2,t=1})$$

$$P(C_t = 1) = 0.1$$

$$P(C_t = 2 = 1 | A_t = 1 = C_t = 1 = Y_t = 1 = 0) = 0.1$$

$$P(A_t = 1 | W, R, X_{2,t}) = 1 - \text{expit}(I(W = 1) + 0.5 * I(R = 2) + 2 * I(R = 0) + X_{2,t})$$

$$P(Y_{t=1}^{r,\infty} | W, X_1, X_{2,t=1}) = \text{expit}(-8 + 0.5 * I(R = 1) + 1.5 * I(R = 0) + 5 * X_1 - 0.5 * X_{2,t=1})$$

$$P(Y_{t=2}^{r,\infty} | W, X_1, X_{2,t=2}, Y_{t=1}^{r,\infty} = 0) = \text{expit}(-8 + 0.5 * I(R = 1) + 1.5 * I(R = 0) + 5 * X_1 - 0.5 * X_{2,t=2})$$

### Data availability:

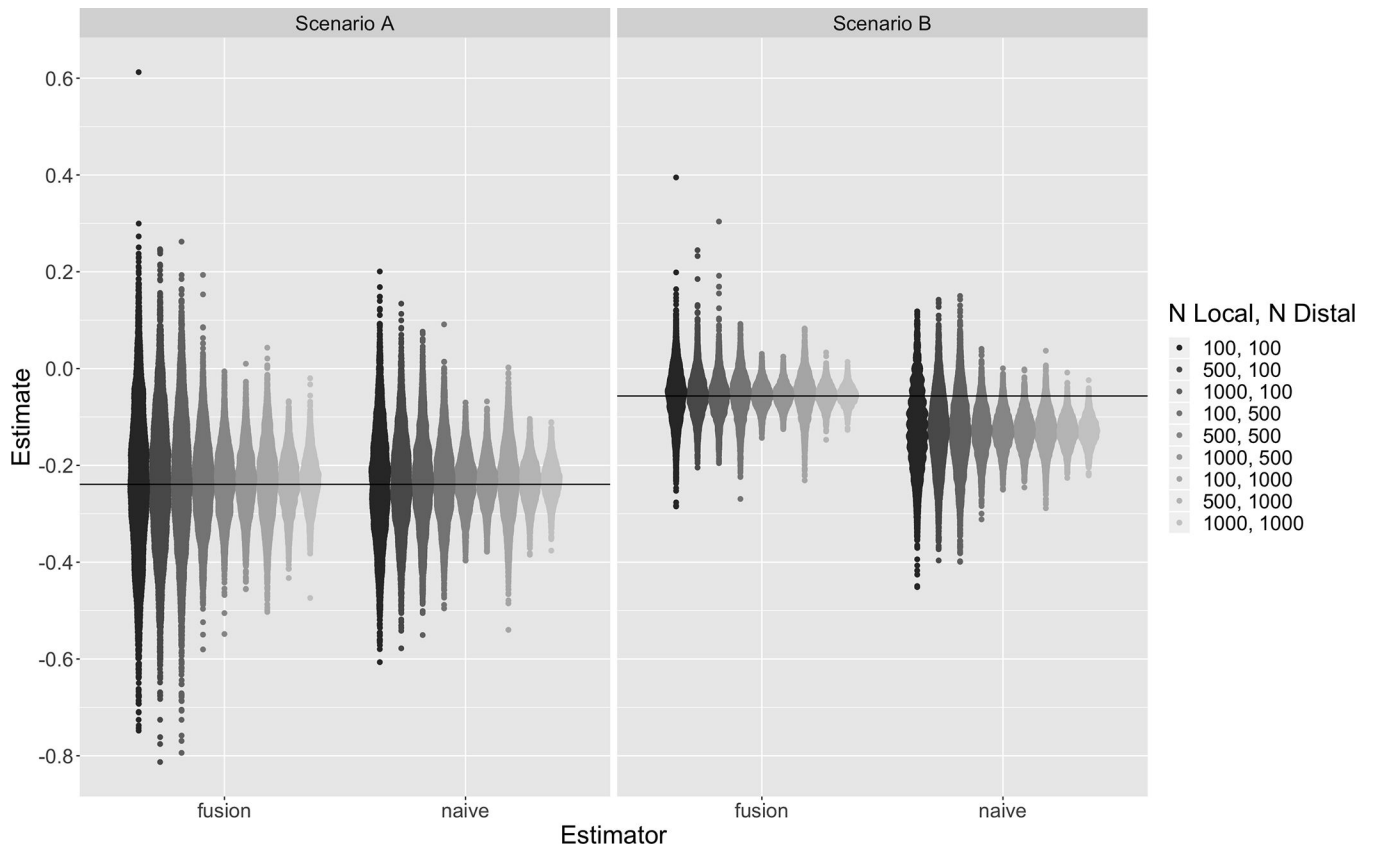
The data that support the findings of this study are available from the corresponding author upon reasonable request.

### REFERENCES

1. Adimora AA, Cole SR, Eron JJ. US Black Women and Human Immunodeficiency Virus Prevention: Time for New Approaches to Clinical Trials. *Clin Infect Dis* 2017;65(2):324–327. doi:10.1093/cid/cix313 [PubMed: 28383649]
2. Bareinboim E, Pearl J. Causal inference and the data-fusion problem. *Proc Natl Acad Sci* 2016;113(27):7345–7352. doi:10.1073/pnas.1510507113 [PubMed: 27382148]
3. Keiding N, Louis TA. Perils and potentials of self-selected entry to epidemiological studies and surveys. *J R Stat Soc Ser A Stat Soc* 2016;179(2):319–376. doi:10.1111/rssa.12136
4. Rudolph JE, Cole SR, Eron JJ, Kashuba AD, Adimora AA. Estimating Human Immunodeficiency Virus (HIV) Prevention Effects in Low-incidence Settings. *Epidemiology*. 2019;30(3):358–364. doi:10.1097/EDE.0000000000000966 [PubMed: 30640216]
5. Westreich D, Edwards JK, Stuart EA, Lesko CR, Cole SR. Transportability of trial results using inverse odds of sampling weights. *Am J Epidemiol* 5 2017;In press. doi:10.1002/elan.
6. Tsiatis AA (Anastasios A. Semiparametric Theory and Missing Data. Springer; 2006.
7. Robins JM, Finkelstein DM. Correcting for noncompliance and dependent censoring in an AIDS clinical trial with inverse probability of censoring weighted (IPCW) log-rank tests. *Biometrics*. 2000;56(3):779–788. doi:10.1111/j.0006-341X.2000.00779.x [PubMed: 10985216]
8. Hernán MA, Robins JM. Per-Protocol Analyses of Pragmatic Trials. *N Engl J Med* 2017;377(14):1391–1398. doi:10.1056/nejmsm1605385 [PubMed: 28976864]



9. Cain LE, Cole SR. Inverse probability-of-censoring weights for the correction of time-varying noncompliance in the effect of randomized highly active antiretroviral therapy on incident AIDS or death. *Stat Med* 2009;28(12):1725–1738. doi:10.1002/sim.3585 [PubMed: 19347843]
10. Lesko CR, Buchanan AL, Westreich D, Edwards JK, Hudgens MG, Cole SR. Generalizing Study Results: A Potential Outcomes Perspective. *Epidemiology*. 2017;28(4):553–561. doi:10.1097/EDE.0000000000000664 [PubMed: 28346267]
11. Buchanan AL, Hudgens MG, Cole SR, et al. Generalizing evidence from randomized trials using inverse probability of sampling weights. *J R Stat Soc Ser A Stat Soc* 2018;181(4):1193–1209. doi:10.1111/rssa.12357
12. Cole SR, Stuart EA. Generalizing evidence from randomized clinical trials to target populations the actg 320 trial. *Am J Epidemiol* 2010;172(1):107–115. [PubMed: 20547574]
13. Lin DY. On the Breslow estimator. *Lifetime Data Anal* 2007;13(4):471–480. doi:10.1007/s10985-007-9048-y [PubMed: 17768681]
14. Satten GA, Datta S. The Kaplan-Meier estimator as an inverse-probability-of-censoring weighted average. *Am Stat* 2001;55(3):207–210. doi:10.1198/000313001317098185 [PubMed: 28845048]
15. Hammer SM, Katzenstein DA, Hughes MD, et al. A Trial Comparing Nucleoside Monotherapy with Combination Therapy in HIV-Infected Adults with CD4 Cell Counts from 200 to 500 per Cubic Millimeter. *N Engl J Med* 1996;335(15):1081–1090. doi:10.1056/NEJM199610103351501 [PubMed: 8813038]
16. Hammer SM, Squires KE, Hughes MD, et al. A Controlled Trial of Two Nucleoside Analogues plus Indinavir in Persons with Human Immunodeficiency Virus Infection and CD4 Cell Counts of 200 per Cubic Millimeter or Less. *N Engl J Med* 1997;337(11):725–733. doi:10.1056/nejm199709113371101 [PubMed: 9287227]
17. Rudolph KE, van der Laan MJ. Robust estimation of encouragement design intervention effects transported across sites. *J R Stat Soc Ser B Stat Methodol* 2017;79(5):1509–1525. doi:10.1111/rssb.12213
18. Pearl J, Bareinboim E. External validity: From do-calculus to transportability across populations. *Stat Sci* 2014;29(4):579–595. doi:10.1214/14-STS486
19. Bang H, Robins JM. Doubly robust estimation in missing data and causal inference models. *Biometrics*. 2005;61(4):962–972. doi:10.1111/j.1541-0420.2005.00377.x [PubMed: 16401269]
20. Van Der Laan MJ, Gruber S. Targeted minimum loss based estimation of causal effects of multiple time point interventions. *Int J Biostat* 2014;8(1). doi:10.1515/1557-4679.1370
21. Stefanski LA, Boos DD. The Calculus of M-Estimation. *Am Stat* 2012;56(May):37–41. doi:10.2307/3087324
22. Efron B, Tibshirani R. Bootstrap Methods for Standard Errors, Confidence Intervals, and Other Measures of Statistical Accuracy. *Stat Sci* 1986;1(1):54–75. doi:10.1214/ss/1177013817



**Figure 1: Results of 4000 simulations of scenarios A and B.**  
The solid lines represent the true treatment effect.

**TABLE 1:**

Bias and root MSE for the naïve and fusion estimators, 4000 simulation trials

Scenario	N		Bias		Root Mean Squared Error	
	Local	Distal	Fusion	Naïve	Fusion	Naïve
A	100	100	0.003	0.004	0.156	0.113
A	100	500	0.005	0.006	0.091	0.078
A	100	1000	0.007	0.008	0.079	0.073
A	500	100	0.008	0.008	0.142	0.097
A	500	500	0.007	0.007	0.069	0.051
A	500	1000	0.005	0.005	0.053	0.041
A	1000	100	0.009	0.008	0.136	0.090
A	1000	500	0.007	0.007	0.065	0.045
A	1000	1000	0.009	0.008	0.049	0.036
B	100	100	0.002	-0.069	0.056	0.108
B	100	500	0.001	-0.071	0.044	0.085
B	100	1000	0.003	-0.070	0.041	0.081
B	500	100	0.003	-0.069	0.042	0.103
B	500	500	0.001	-0.072	0.024	0.080
B	500	1000	0.001	-0.072	0.021	0.077
B	1000	100	0.001	-0.073	0.040	0.106
B	1000	500	0.001	-0.072	0.021	0.079
B	1000	1000	0.001	-0.072	0.018	0.077

**TABLE 2:**

Average and Monte Carlo standard errors for the naive and fusion estimators, 4000 simulation trials

Scenario	N		Average Standard Error <sup>a</sup>		Monte Carlo Standard Error <sup>b</sup>	
	Local	Distal	Fusion	Naïve	Fusion	Naïve
A	100	100	0.163	0.123	0.156	0.113
A	100	500	0.108	0.096	0.091	0.078
A	100	1000	0.110	0.105	0.079	0.073
A	500	100	0.149	0.105	0.141	0.096
A	500	500	0.073	0.055	0.068	0.050
A	500	1000	0.057	0.046	0.053	0.041
A	1000	100	0.147	0.103	0.136	0.090
A	1000	500	0.069	0.050	0.065	0.044
A	1000	1000	0.051	0.039	0.048	0.035
B	100	100	0.054	0.087	0.056	0.082
B	100	500	0.043	0.056	0.044	0.047
B	100	1000	0.043	0.057	0.041	0.040
B	500	100	0.041	0.081	0.042	0.076
B	500	500	0.025	0.039	0.024	0.037
B	500	1000	0.022	0.030	0.021	0.028
B	1000	100	0.039	0.080	0.040	0.077
B	1000	500	0.021	0.037	0.021	0.034
B	1000	1000	0.018	0.028	0.018	0.026

<sup>a</sup> Average standard error is defined as the average of the standard errors across Monte Carlo simulations, estimated with the standard error estimator described in the text.

<sup>b</sup> Monte Carlo standard error is defined as the standard deviation of the estimates across Monte Carlo simulations.

Select characteristics of subjects from AIDS Clinical Trials Group studies 175 and 320<sup>†</sup>

**TABLE 3:**

	ACTG 175		ACTG 320	
	Monotherapy N = 271	Dual-therapy N = 550	Dual-therapy N = 579	Triple-therapy N = 577
<b>Male Sex</b>	221 (81.5)	448 (81.5)	485 (83.8)	471 (81.6)
<b>Age (Median, IQR)</b>	35 (30, 41)	35 (30, 41)	38 (33, 44)	38 (33, 44)
<b>Black race</b>	70 (25.8)	138 (25.1)	165 (28.5)	163 (28.2)
<b>Injection drug use</b>	33 (12.2)	79 (14.4)	93 (16.1)	91 (15.8)
<b>Karnofsky Score (Mean, SD)</b>	95.0 (6.3)	95.4 (6.0)	91.4 (7.7)	91.2 (7.7)

<sup>†</sup>All numbers are N (%) unless otherwise noted