The **T**riangulation **WI**thin A **ST**udy (**TWIST**) framework for causal inference within Pharmacogenetic research

Jack Bowden[1], Luke C. Pilling[2], Deniz Türkmen[2],
Chia-Ling Kuo[3] & David Melzer[2]

[1]*Exeter Diabetes Group (ExCEED), College of Medicine and Health, University of Exeter, Exeter, U.K.*
[2]*Epidemiology and Public Health Group, College of Medicine and Health, University of Exeter, Exeter, U.K.*
[3] *Connecticut Convergence Institute for Translation in Regenerative Engineering, University of Conneticut, CT, U.S.A*

*Address for correspondence:
Jack Bowden, University of Exeter College of Medicine and Health
RILD Building, RD&E Hospital,
Barrack Road, Exeter, Devon, EX2 5DW
j.bowden2@exeter.ac.uk

Abstract

In this paper we review the methodological underpinnings of the general pharmacogenetic approach for uncovering genetically-driven treatment effect heterogeneity. This typically utilises only individuals who are treated and relies on fairly strong baseline assumptions to estimate what we term the 'genetically mediated treatment effect' (GMTE). When these assumptions are seriously violated, we show that a robust but less efficient estimate of the GMTE that incorporates information on the population of untreated individuals can instead be used. In cases of partial violation, we clarify when Mendelian randomization and a modified confounder adjustment method can also yield consistent estimates for the GMTE. A decision framework is then described to decide when a particular estimation strategy is most appropriate and how specific estimators can be combined to further improve efficiency. Triangulation of evidence from different data sources, each with their inherent biases and limitations, is becoming a well established principle for strengthening causal analysis. We call our framework '**T**riangulation **WI**thin a **ST**udy' (TWIST)' in order to emphasise that an analysis in this spirit is also possible within a single data set, using causal estimators that are approximately statistically uncorrelated, but reliant on different sets of assumptions. We illustrate these approaches by re-analysing primary-care-linked UK Biobank data relating to *CYP2C19* genetic variants, Clopidogrel use and stroke risk, and data relating to *APOE* genetic variants, statin use and Coronary Artery Disease.

Key Words: Pharmacogenetics, Difference-in-differences, Mendelian randomization, Confounder adjustment, Triangulation.

1 Background

Over the last 20 years the field of Epidemiology has embraced the exploitation of random genetic inheritance to help uncover causal mechanisms of disease using the technique of Mendelian randomization (MR). [1]. The basic premise of MR is illustrated by the causal diagrams in Figure 1: Genetic variants, usually Single Nucleotide Polymorphisms or SNPs, G, are found which robustly explain variation in a modifiable risk factor, X, where X is typically continuous (for example body mass index). The association between the exposure and an outcome, Y, hypothesised to be a downstream consequence of X, may be contributed to in observational data by unobserved confounding U. If present, such confounding would mean that the naive association between X and Y would not reflect the causal effect of X on Y. If the important confounders could be appropriately measured and adjusted for, then this bias could be removed, but an MR analysis aims to circumvent any potential confounding by instead measuring the association between the outcome and the portion of the exposure that can be genetically predicted by the SNPs. Provided that the SNPs are independent of the confounders U, do not affect the outcome through any other pathway except the exposure, and the causal effect of a unit increase in the exposure is the same across individuals, then MR can consistently estimate the average causal effect of intervening on X (e.g. to reduce or increase) on the outcome Y. The Instrumental Variable (IV) assumptions of the MR approach are denoted IV1-IV4 in Figure 1. Genetic variants can also play an important role in help-

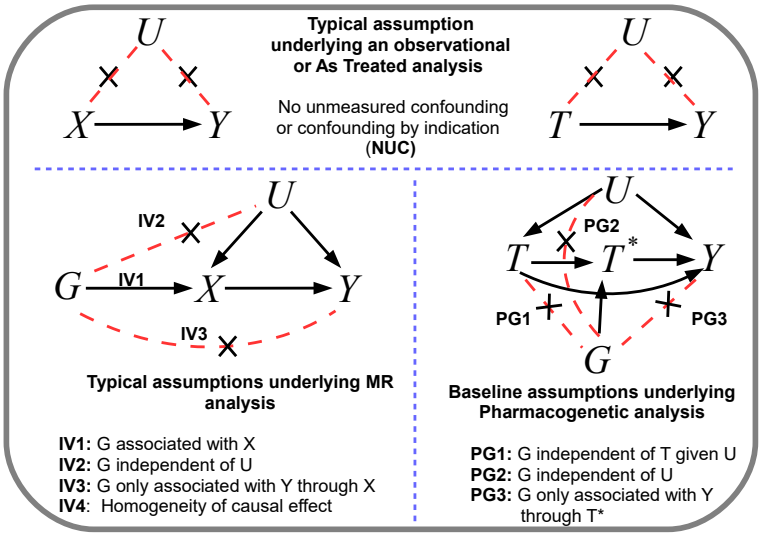

Figure 1: *Top: typical assumptions underlying an the causal interpretation of an exposure X or treatment T on an outcome Y following standard regression analysis: Bottom-left: Assumptions underlying a standard MR analysis using genetic variant G as an IV to estimate the causal effect of X on Y. Bottom-right: Assumptions underlying a standard pharmaco-genetic analysis.*

ing to explain treatment effect heterogeneity, by the science of Pharmacogenetics. A canonical example is Clopidogrel: the primary drug for ischemic stroke prevention in the UK and many other countries (see `https://cks.nice.org.uk/antiplatelet-treatment`). It requires CYP2C19 enzyme activation in order to be properly metabolised into the active form of the drug and thus work to its fullest extent. However, it has long been known that both common loss-of-function and gain-of-function variants within the *CYP2C19* gene region can massively impact each patient's ability to metabolise the drug [2]. Consequently, when prescribed in a primary care setting its effectiveness is heterogeneous, working for some and not for others. Estimating the true effectiveness of a treatment from observational data is challenging due to 'confounding by indication'[3] (see Figure 1 top-right). For example, Clopidogrel use will, quite rightly, strongly depend on an individual's underlying risk of stroke, but socio-economic factors may also influence an individual's ability to access appropriate healthcare [4]. Use for a sustained period may additionally depend on whether they tolerate the drug without side-effects. Whilst these confounding factors could in principle be directly accounted for in a statistical analysis if complete information on a patient's clinical state were available, this is seldom the case. A well known example, albeit in a non-pharmacogenetic context, is hormone replacement therapy, which was linked to increased cancer risk

in observational data but not in subsequent randomized trials [5]. Thankfully, the need for adjustment can be circumvented if the purpose of the analysis is is instead to compare the relative effectiveness of a treatment across across genetic groups (see Figure 1 bottom-right). In the case of Clopidogrel and *CYP2C19*, typically one might assume that *CYPC219* variants G: do not predict whether an individual receives Clopidogrel (T); are not associated with any confounders (U) predicting Clopidogrel use and stroke (Y); and only affect stroke risk through their interaction with Clopidogrel (denoted by the variable T^*). We will refer to these as the 'Pharmacogenetic' (PG) assumptions as a counterpart to the IV assumptions utilised by MR. A key difference exists between the role of the gene in MR and the role of the gene in pharmacogenetics: In MR, genes are assumed to directly influence the modifiable exposure. In Pharmacogenetics we can think of treatment as the exposure, and the genes are hypothesised to alter treatment once it is taken.

In recent work, Pilling et al [6] use data from GP-linked UK Biobank participants on Clopidogrel treatment to estimate the its effect in different *CYP2C19* genetic subgroups and from this the number of strokes that could potentially be avoided if all individuals could experience the same benefit as the group with the most favourable genotype (through either dose modification or through switching to an alternative drug). In this paper we review the methodological underpinnings of the general PG approach, which utilises only individuals who are treated and relies on fairly strong baseline PG assumptions to estimate what we refer to as the 'genetically mediated treatment effect' (GMTE). When the PG assumptions are seriously violated, we show that a robust but less efficient estimate of the GMTE that incorporates information on the population of untreated individuals can instead be used. In cases of partial violation, we clarify when Mendelian randomization and traditional confounder adjustment can also yield consistent estimates for the GMTE. A decision framework is then described to decide when a particular estimation strategy is most appropriate and how specific estimators can be combined to further improve efficiency.

Triangulation of evidence from different data sources, each with their inherent biases and limitations, is becoming a well established principle for strengthening causal analysis [7]. We call this framework '**T**riangulation **WI**thin a **ST**udy' (TWIST)' in order to emphasise that an analysis in this spirit is also possible within a single data set, using causal estimators that are approximately statistically uncorrelated, but reliant on different sets of assumptions. We illustrate these approaches by re-analysing primary-care-linked UK Biobank data relating to *CYP2C19* genetic variants, Clopidogrel use and stroke risk, and data relating to *APOE* genetic variants, statin use and Coronary Artery Disease (CAD).

2 Methods

Suppose that we are interested in evaluating the maximal effectiveness of a treatment T on an outcome Y using observational data. For simplicity we will assume initially that T is a binary treatment indicator so that, if prescribed, it is taken in full, that Y is a continuous or binary outcome variable and we are interested in estimating the treatment effect as a mean or risk difference contrast. A naive way of estimating this effect would be to compare outcomes across those who are treated and those who are untreated. Borrowing terminology from the clinical trials literature, we refer to this as the 'As treated' (AT) estimate

$$\hat{\beta}_{AT} = \hat{E}[Y|T=1] - \hat{E}[Y|T=0] \tag{1}$$

The AT estimate may not directly address our research needs for two reasons. Firstly, although we may understand many of the factors which influence whether an eligible individual is prescribed treatment by their doctor, there may be unmeasured variables, U, which influence both the decision to prescribe treatment and the outcome. Indeed, even if the treatment is truly effective in reducing the severity or risk of Y, it is highly likely that the population of treated individuals may still experience worse outcomes than those who are untreated. This would mean that the sign of the AT estimate could be positive, and thus qualitatively different than the true causal effect. This is classic confounding by indication. Secondly, a pharmacological investigation may suggest that the treatment does not in fact work for a certain proportion of the population at all, has a markedly reduced effectiveness, or increases the risk of side-effects. We would therefore like to estimate the difference in patient outcomes if all patients who take treatment experienced the 'full' effect, as experienced by those with a genotype G=1, versus the reduced (or zero) effect experienced by those with genotype G=0. This effect could be realised in practice by switching to an alternative medication whose effect is not affected by patient genotype.

To make the target of our analysis more explicit we use potential outcomes notation. Let $Y_i(g,t)$ be the potential outcome for individual i if assigned to genotype level $G = g$ and treatment level $T = t$ and define the potential treatment variable $T_i(j)$ as the treatment experienced by patient i under allocation to genotype level j. Our quantity of interest is the genetically mediated treatment effect (GMTE) on the outcome:

$$\beta_{GMTE}(Y) = E[Y_i(g, T_i(1))] - E[Y_i(g, T_i(0)] \tag{2}$$

This estimand reflects the extent to which a gene affects the outcome through the pathway of treatment only, and not the effect of the genotype on the outcome through any other mechanism. In order to make ideas more concrete, consider the following model for the expectation of Y given treatment, T, genetic variant G and unmeasured confounder U:

$$\begin{aligned} E[Y|T,G,U] &= \gamma_{Y0} + \beta_1 TG + \beta_0 T(1-G) + \gamma_{GY} G + \gamma_{YU} U & (3) \\ &= \gamma_{Y0} + \beta_0 T + (\beta_1 - \beta_0) TG + \gamma_{GY} G + \gamma_{YU} U & (4) \end{aligned}$$

Under model (3), β_1 and β_0 reflect the treatment effect experienced by those with genotype G=1 and G=0 respectively. The parameter γ_{GY} represents the direct effect of G on Y and γ_{YU} represents the direct effect of U on Y. The GMTE is equal to the difference in treatment effects experienced by the two genetic groups, $\beta_1 - \beta_0$. If the two genetic groups experience distinct and non-zero treatment effects because both β_1 and β_0 are non-zero, model (3) violates the Homogeneity assumption (IV4) typically invoked in Mendelian randomization (see Figure 1). When model (3) is re-written as in (4), we see that Homogeneity violation can be understood as T affecting Y in two ways: directly through a main effect (β_0) as well as indirectly through its interaction with G ($\beta_1 - \beta_0$). If those who take treatment with genotype level G=0 experience no treatment effect all ($\beta_0 = 0$), then the Homogeneity assumption would be satisfied and model (3) reduces to:

$$E[Y|T,G,U] = \gamma_{Y0} + \beta_1 TG + \gamma_{GY} G + \gamma_{YU} U$$

In the general case of treatment effect heterogeneity (model (3)), when assumptions PG1-PG3 hold, so that $G \perp\!\!\!\perp T|U$, $G \perp\!\!\!\perp U$ and $G \perp\!\!\!\perp Y|T,U$ (this last condition would imply that γ_{GY} in equation (5) is zero), then we can consistently estimate the GMTE using only treated individuals via the 'GMTE(1)' estimate:

$$\hat{\beta}_{GMTE(1)}(Y) = \hat{E}[Y|T=1, G=1] - \hat{E}[Y|T=1, G=0], \tag{5}$$

where $\hat{E}[.]$ refers to the sample mean. We use the additional subscript '(1)' to denote that the estimate conditions on T=1 and the bracket '(Y)' to denote the outcome, Y. Although the latter notation may seem redundant at first sight, it is needed for subsequent development. Further intuition for this estimate is provided using DAGs in Figure 2 (top-left) and Section 2.1. Note that in the DAG we replace the product TG with the variable T^* to denote the mechanistic interaction. In Appendix 1 we discuss the precise conditions under which the GMTE(1) estimate is consistent for the GMTE under our assumed model. For example, it can still be consistent even when assumption PG1 is violated if $T \perp\!\!\!\perp U$. We call this the No Unmeasured Confounder (NUC) assumption from now on.

2.1 A robust GMTE estimator

The GMTE(1) estimate is simple, intuitive and can be calculated entirely in the sub-population of those who are treated. However, in this context it generally requires assumptions PG1-PG3. Violation of PG1 implies that an individual's genotype directly influences the likelihood that they receive treatment. For example, it could be that those with a G=0 genotype have an increased risk of side effects on treatment and choose to immediately come off the drug. An alternative explanation could be genetic population stratification [8]: e.g the allele frequency of the genetic variant G and the rate of treatment could simultaneously vary across individuals from different ethnic groups. An example of PG2 violation would be if the genetic variant increases the likelihood of an unmeasured risk factor for the outcome, and this risk factor also increases their likelihood of being treated. An example of PG3 violation would be if an individual's genotype directly affects the outcome through a pathway completely independent of either treatment or any confounding factor, which could be viewed as horizontal pleiotropy [9]. When any of these assumptions are violated $\hat{\beta}_{GMTE(1)}(Y)$ will reflect the genetically mediated effect of treatment plus the bias due to PG1-3 violation. The first and second examples above are illustrated in Figure 2 (top-left). In this case the bias is the sum of

- b_{PG1} via the $G \to T \leftarrow U \to Y$ pathway (due to violation of PG1);
- b_{PG2} via the $G \to U \to Y$ pathway (due to violation of PG2);
- b_{PG3} via the $G \to Y$ pathway (due to violation of PG3)

.

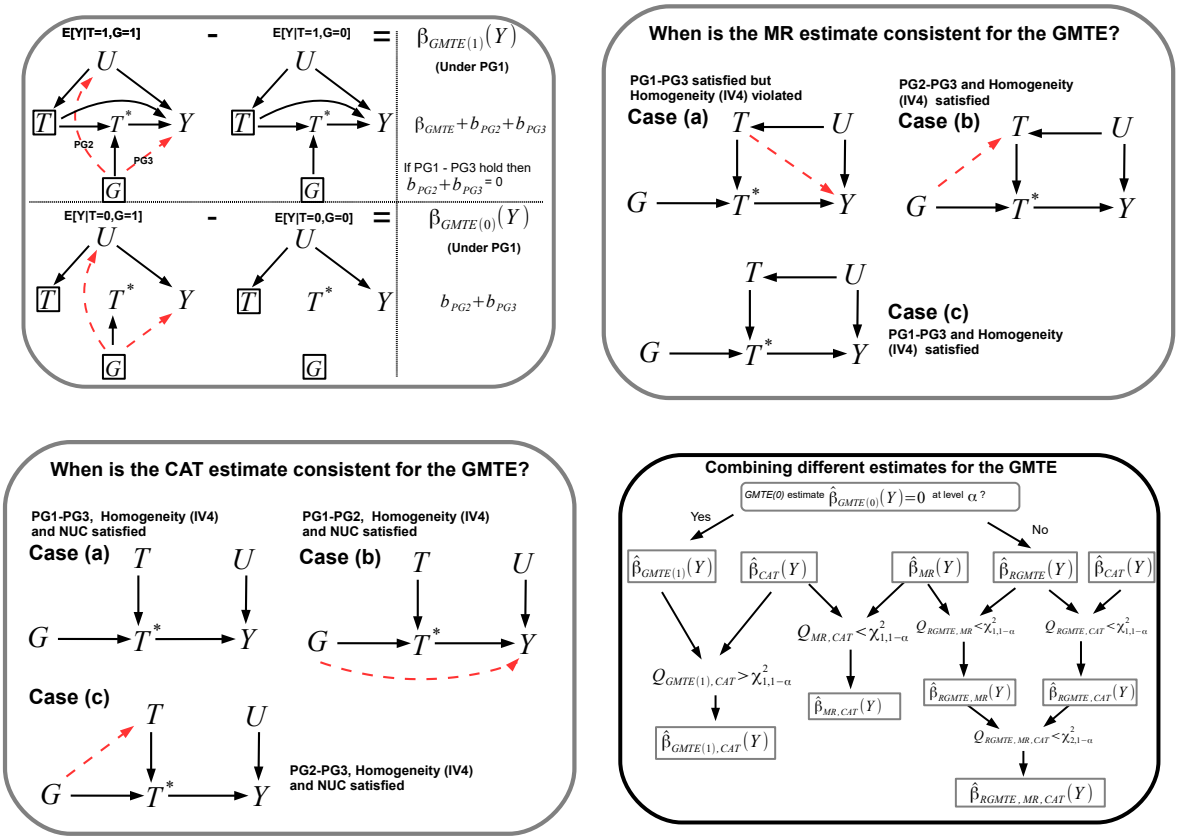

Figure 2: *Top-left: The $\hat{\beta}_{GMTE(1)}(Y)$ in the treated group is equal to the true GMTE plus bias due to violation of PG1-PG3. Square boxes around T and G denotes that we are conditioning on these variables. The bias of $\hat{\beta}_{GMTE(1)}(Y)$ can be quantified in the untreated group via $\hat{\beta}_{GMTE(0)}(Y)$ and subtracted from $\hat{\beta}_{GMTE(1)}(Y)$ to give a more robust estimate of the GMTE. Top-right: Three cases (a)-(c) for which the MR estimate is consistent for the GMTE. Bottom-left: Three cases (a)-(c) for which the CAT estimate is consistent for the GMTE. Bottom-right: A decision framework for choosing when to combine statistically uncorrelated estimates of the GMTE.*

Whilst the bias contributions b_{PG2} and b_{PG3} are clear, bias contribution b_{PG1} is perhaps less so; it occurs because the GMTE estimate explicitly conditions on treatment T, the presence of an association between G and T makes T a 'collider' [10]. Thankfully, when assumption PG1 is satisfied but PG2-3 are violated, the additional bias terms b_{PG2} and b_{PG3} can be removed by incorporating information on the untreated population. This is achieved by calculating the equivalent GMTE(1) estimate for the untreated group,

$$\beta_{GMTE(0)}(Y) = \hat{E}[Y|T=0, G=1] - \hat{E}[Y|G=0, T=0] \tag{6}$$

and subtracting it from that in the treated group. The 'Robust' genetically mediated treatment effect, RMTE(Y), is therefore defined as

$$\hat{\beta}_{RGMTE}(Y) = \hat{\beta}_{GMTE(1)}(Y) - \hat{\beta}_{GMTE(0)}(Y) \tag{7}$$

This bares a strong resemblance to the 'difference-in-differences' approach used in econometrics [11]. For further clarification see Figure 2 (top-left). In Appendix 2 we discuss the precise conditions under which the RGMTE estimate is consistent for the GMTE under our assumed model. Although assumption PG1 is key for the RGMTE, we show that it can still consistently estimate the GMTE if PG1 is violated but the NUC assumption is satisfied.

2.2 The Mendelian randomization estimate

Given data on both treated and untreated individuals, it is possible to obtain an estimate for the GMTE by using the genetic variant G as an instrumental variable for the mechanistic interaction variable $T^* = TG$, as in Mendelian randomization. The MR estimate with a single instrument is known to have the form

$$\beta_{MR}(Y) = \frac{E[Y|G=1] - E[Y|G=0]}{E[T^*|G=1] - E[T^*|G=0]} \tag{8}$$

where T^* is the *de-facto* exposure. As discussed previously, the MR estimate is usually justified by assuming that G satisfies assumptions IV1-IV3 with respect to the exposure (in our case, T^*) and that the Homogeneity assumption (IV4) is satisfied. However, unless those who have genotype level G=0 experience no treatment effect, the Homogeneity assumption will be violated in this setting. In Appendix 3, we show that the MR estimate in (8) is consistent for the GMTE if PG2-PG3 hold and either PG1 holds, or the Homogeneity assumption holds (β_0=0). Further intuition for this approach is given in Figure 2 (top-right).

2.3 A corrected 'As Treated' estimate

As previously discussed, the As Treated estimate suffers in general from confounding by indication, and a lack of specificity to the genetic variant driving the mechanistic interaction with the treatment. However, if the Homogeneity and NUC assumptions are satisfied and either PG1 or PG3 are satisfied, then the 'Corrected' As Treated estimate (CAT):

$$\beta_{CAT}(Y) = \frac{E[Y|T=1] - E[Y|T=0]}{E[G|T=1]}, \tag{9}$$

can consistently estimate the GMTE. Further details are provided in Appendix 4 and for further intuition see Figure 2 (bottom-left).

2.4 Implementation of the methods

In Table 1 we provide summary information for the GMTE(1), RGMTE, MR and CAT estimates, including the assumptions that they rely on and how they can be obtained using generic regression models, where the variable Z is included as a measured covariate. We also include a column that shows, how to test for important potential confounders of treatment and outcome using an estimate-specific approach (see Appendix 1-4 for more details). A nice property is that, for each estimate of the GMTE, one calculates exactly the same estimate, but using a potential confounder U as the outcome. If this estimate is non-zero then this indicates a meaningful bias, unless said confounder is adjusted for.

When the outcome is continuous, the approaches can be implemented using linear regression to estimate the GMTE as a mean difference. With a binary outcome, we recommend estimating risk differences directly using either a linear probability model, or using a logistic regression model to furnish estimates on the risk difference scale. In the latter case, a risk difference can be constructed by calculating the mean difference between predicted probabilities under manipulation of the appropriate variable between levels 1 and 0, as suggested by Gelman and Hill [12]. For example, in the case of the GMTE(1) estimator

$$\hat{\beta}_{GMTE(1)}(Y) = \frac{1}{n}\sum_{i=1}^{n} \hat{\pi}(Y)(T^*=1, T=1, Z=z_i) - \hat{\pi}(Y)(T^*=0, T=1, Z=z_i)$$

where $\hat{\pi}(Y)(T^* = t_i^*, T = j, Z = z_i)$ is the estimated fitted value from a logistic regression of Y on T,T^* and Z at treatment level j. For time-to-event data, we recommend analysing the data using an Aalen additive hazard model, as suggested by Tchetgen et. al. [13]. Specifically, one would assume that model (3) holds on the additive hazard scale, so that at time t_y the hazard:

$$h(t_y|T, U, G, Z) = \gamma_{Y0} + \beta_1 TG + \beta_0 T(1-G) + \gamma_{GY} G + \gamma_{YU} U + \gamma_{YZ} Z \tag{10}$$

The proposed modelling framework for estimating mean, risk differences or additive hazard differences is intended to guarantee that we can obtain estimates for the GMTE from different estimators on the same scale, because these measures are collapsible. That is, they should remain constant when marginalised over unobserved confounders [14]. This is especially important for being able to effectively combine methods, as described in the next section.

Estimate	Key Assumptions	Confounder test	Fit
CAT $\hat{\beta}_{CAT}(Y)$	$\{PG1 \cup PG3\} \cap NUC \cap Hom$	$\hat{\beta}_{CAT}(G) = 0$, $\hat{\beta}_{CAT}(U) = 0$	Coef. of T_{CAT} in $Y \sim T_{CAT} + Z$
GMTE(0) $\hat{\beta}_{GMTE(0)}(Y)$	$PG3 \cap G \perp\!\!\!\perp U \mid T = 0$	$\hat{\beta}_{GMTE(0)}(U) = 0$	Coef. of T^{*-} in $Y \sim T^{-} + T^{*-} + Z$
GMTE(1) $\hat{\beta}_{GMTE(1)}(Y)$	$PG3 \cap G \perp\!\!\!\perp U \mid T = 1$	$\hat{\beta}_{GMTE(1)}(U) = 0$	Coef. of T^{*} in $Y \sim T + T^{*} + Z$
RGMTE $\hat{\beta}_{RGMTE}(Y)$	$Cov(U, G) \perp\!\!\!\perp T$	$\hat{\beta}_{RGMTE}(U) = 0$	Coef. of T^{*} in $Y \sim T + T^{*} + \hat{T}^{*} + Z$
MR $\hat{\beta}_{MR}(Y)$	$\{PG1 \cup Hom\} \cap PG2 \cap PG3$	$\hat{\beta}_{MR}(U) = 0$	Coef. of $\hat{T}^{*}$ in $Y \sim \hat{T}^{*} + Z$

Table 1: A summary of the identifying assumptions, confounder test formulae and model implementation for the four estimators of the GMTE. For the GMTE(0) estimate, $T_{CAT} = \hat{E}[G|T = 1]T$, for the GMTE(0) estimate $T^{-} = 1 - T$, $T^{*-} = T^{-}G$, for the RGMTE estimate $T^{*} = TG$ and $\hat{T}^{*} = \hat{E}[T^{*}|G]$.

2.5 Which estimates can be combined?

When two estimates are highly correlated, we gain little knowledge when they are observed to be similar. However, when two uncorrelated or weakly correlated estimates are similar, it gives credence to the hypothesis that they are estimating the same underlying quantity, and there is the potential to combine them into a single, more precise estimate. In Appendix 5 we show that the RGMTE and MR estimates are asymptotically uncorrelated. We also show that the CAT estimate is mutually uncorrelated with the GMTE(1) and RGMTE estimates, and uncorrelated with the MR estimate when $G \perp\!\!\!\perp T$. In cases where G and T are not perfectly independent, but G is a modest predictor of T (a highly plausible scenario in most pharmcogenetic contexts), the correlation between the MR and CAT estimate will be non-zero but practically negligible. The fact that most estimate-pairs are uncorrelated makes them easy to combine via a simple inverse variance weighted average or meta-analysis. In order to decide whether two uncorrelated estimates can be combined, we propose the use of a simple heterogeneity statistic. Taking the GMTE(1) and CAT estimates as an example, we would first calculate

$$Q_{GMTE(1),CAT} = \sum_{e} w_e(\hat{\beta}_e(Y) - \hat{\beta}_C(Y))^2, \text{ where } \hat{\beta}_c(Y) = \frac{\sum_e w_e \hat{\beta}_e}{\sum_e w_e} \tag{11}$$

where $e = \{GMTE(1), CAT\}$ and where w_e is the inverse variance of estimate $\hat{\beta}_e$. If $Q_{GMTE(1),CAT} < \chi^2_{1,1-\alpha}$ (where α is pre-specified) then we judge the GMTE(1) and CAT estimates to be sufficiently similar to combine into a single estimate $\hat{\beta}_{GMTE(1),CAT}(Y)$. If $Q_{GMTE(1),CAT}$ is greater than the pre-specified threshold then the two estimates should be left separate.

The RMGTE and MR estimates are in general highly correlated with the GMTE(1) estimate, and should therefore not be combined. In Appendix 6 we show that the MR and RMGTE estimates can be viewed as complementary functions of the GMTE(1) and GMTE(0) estimates, and that the combined RGMTE/MR estimate is exactly equivalent to the GMTE(1) estimate when $G \perp\!\!\!\perp T$ and the proportion of treated and untreated participants in the data is the same. Figure 2 (Bottom-right) shows a pictorial diagram of all single and combined estimates that can be derived using the above heterogeneity statistic criteria. This comprises four original estimates (CAT,GMTE1,RMGTE,MR), four 'paired estimates (CAT/GMTE1, CAT/RGMTE, CAT/MR, RGMTE/MR) and one 'triplet' estimate (CAT/RGMTE/MR), making nine in total. One possible analysis option would be to report all single and valid combined estimates which are sufficiently homogeneous according to a particular significance threshold. Another option would be

to allow the GMTE(0) estimate to initially guide the analysis towards either the GMTE(1) estimate (and its possible combination with the CAT estimate) or the RMGTE estimate (and its possible combination with either MR estimate, the CAT estimates or both).

3 Simulation illustration

Trial data comprising a binary genotype G, treatment indicator T, observed covariate Z and a continuous outcome Y are simulated for n=10,000 patients using the following data generating model which is consistent with the DAG in Figure 3:

$$\begin{aligned}
G_i &\sim Bern(0.5) \\
U_i &\sim N(0,1) + \gamma_{GU} G_i \\
\eta_{Ti} &\sim \gamma_{T0} + \gamma_{UT} U_i + \gamma_{GT} G_i + \gamma_{ZT} Z_i + \epsilon_{Ti} \\
Pr(T_i = 1 | U_i, G_i) &= \text{expit}(\eta_{Ti}) \\
Y_i | T_i, G_i, Z_i, U_i &= \gamma_{Y0} + \beta_1 T_i G_i + \beta_0 T_i (1 - G_i) + \gamma_{GY} G_i + \gamma_{YZ} Z_i + \gamma_{YU} U_i + \epsilon_{Yi}
\end{aligned}$$

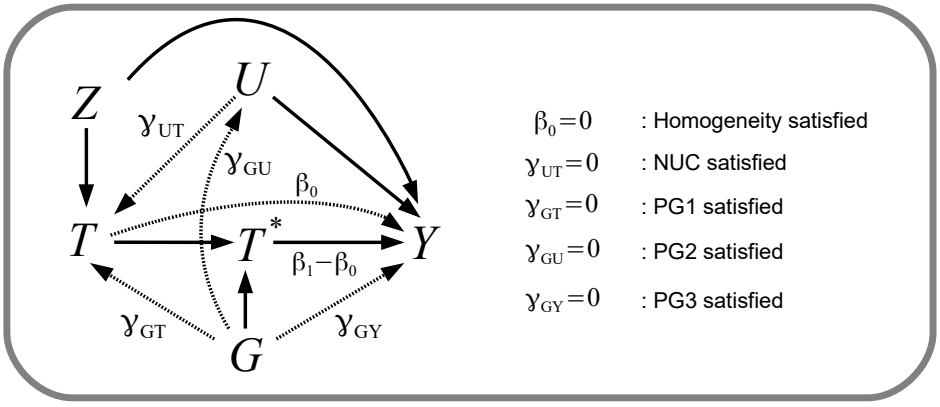

Figure 3: *Causal DAG consistent with the data generating model in the simulation study. Solid lines represent relationships between variables that remain constant across all simulations. Dashed lines represent possible relationships between variables that are either included or excluded in a given simulation scenario.*

Under this model, assumptions PG1-PG3 are violated if γ_{GT}, γ_{GU} and γ_{GY} are non-zero respectively. Finally, the Homogeneity (Hom) and NUC assumption ($U \perp\!\!\!\perp T$) are violated when β_0 and γ_{UT} are non-zero respectively. Note that if the NUC assumption holds, then PG2 is in a sense automatically satisfied because U is no longer a confounder. However, in this case there may still be a path from G to Y via U. This would then form all or part of any PG3 violation.

Figure 4 shows the distribution of estimates for the GMTE obtained across 500 independent data sets and six simulation scenarios, using the CAT, GMTE(1), RGMTE and MR estimators. We also show the distribution of the GMTE(0) estimate in each case, as a helpful guide to understand the extent of bias that can be estimated from the data. In all scenarios the true GMTE is fixed at -0.5. Table 2 shows the mean point estimates, standard errors and 95% confidence interval coverage corresponding to the same six scenarios. Table 3 shows mean point estimates, standard errors and 95% confidence interval coverage for the 5 combined estimators shown within the decision framework in Figure 2 (bottom-right), using a significance threshold of α=0.05.

In Scenario 1 assumption PG3 is violated but all others (PG1, PG2, Hom, NUC) are satisfied. In this case both the CAT and RMGTE estimators are unbiased, with the RGMTE having the smallest standard error. In Table 3 we see that the combined RGMTE/CAT estimate is consequently unbiased with a standard error of 0.085, which is smaller than either the RMGTE or CAT estimates.

In Scenario 2 the NUC assumption is violated but all others (PG1, PG2, Hom) are satisfied. In this case the GMTE(1), RGMTE and MR estimators are unbiased, with the GMTE(1) estimate being the most precise. In Table 3 we see that the combined RGMTE/MR estimate is consequently unbiased with

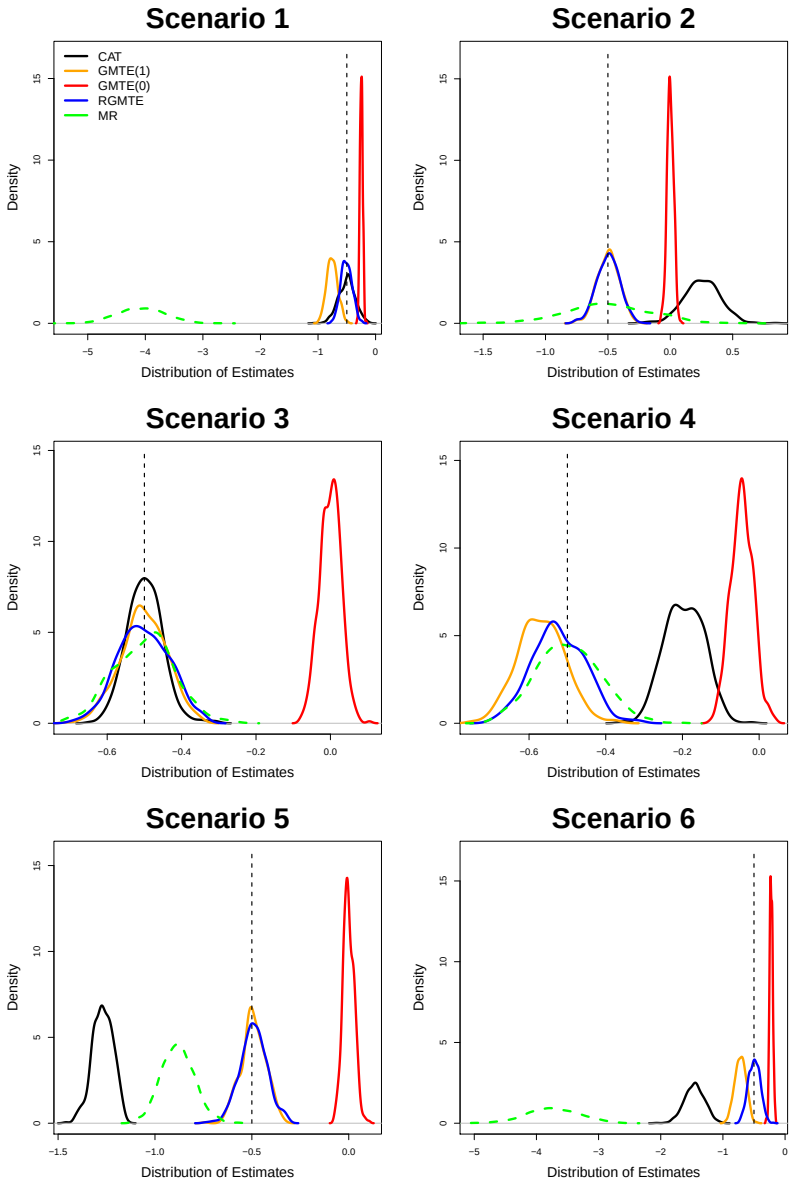

Figure 4: *Distribution of estimates for the CAT, GMTE(1), GMTE(0), RGMTE and MR estimators across six simulation scenarios. In each case, the true GMTE is fixed at -0.5.*

Scenario & Assumption(s) violated		Estimator CAT	GMTE(1)	GMTE(0)	RGMTE	MR
1. PG3 violated	Est.	**-0.502**	-0.755	-0.248	**-0.507**	-4.068
	S.E	**0.150**	0.093	0.025	**0.096**	0.350
	Cov^{ge}	**0.95**	0.22	0	**0.95**	0.00
2. NUC violated	Est.	0.255	**-0.498**	0.000	**-0.498**	**-0.497**
	S.E	0.140	**0.089**	0.025	**0.092**	**0.320**
	Cov^{ge}	0.00	**0.96**	0	**0.95**	**0.94**
3. PG1 violated	Est.	**-0.499**	**-0.502**	0.000	**-0.502**	**-0.501**
	S.E	**0.050**	**0.063**	0.028	**0.069**	**0.081**
	Cov^{ge}	**0.95**	**0.95**	0	**0.95**	**0.96**
4. PG1 and NUC violated	Est.	-0.195	-0.568	-0.045	-0.523	**-0.497**
	S.E	0.050	0.061	0.028	0.067	**0.079**
	Cov^{ge}	0.00	0.79	0	0.93	**0.94**
5. PG2 & Hom violated	Est.	-1.270	**-0.493**	-0.001	**-0.492**	-0.883
	S.E	0.050	**0.063**	0.028	**0.069**	0.083
	Cov^{ge}	0.00	**0.95**	0	**0.94**	0.01
6. All except PG1	Est.	-1.458	-0.719	-0.223	**-0.496**	-3.728
	S.E	0.150	0.092	0.025	**0.095**	0.360
	Cov^{ge}	0.00	0.34	0	**0.96**	0.00

Table 2: *Point estimates, standard errors and coverage (of 95% confidence interval) for the CAT, GMTE(1), GMTE(0) RGMTE and MR estimates across six simulation scenarios. In each case, the true GMTE is fixed at -0.5. Unbiased estimates and associated standard errors/coverage are highlighted in bold.*

Scenario & Assumption(s) violated		Estimator RGMTE/MR	RGMTE/CAT	MR/CAT	GMTE1/CAT	RGMTE/MR/CAT
1. PG3 violated	Est.	-0.754	**-0.506**	-1.041	-0.684	-0.684
	S.E	0.092	**0.080**	0.140	0.079	0.078
	Cov^{ge}	0.22	**0.96**	0.02	0.35	0.35
2. NUC violated	Est.	**-0.498**	-0.274	0.134	-0.286	-0.286
	S.E	**0.089**	0.077	0.130	0.075	0.075
	Cov^{ge}	**0.96**	0.15	0.01	0.16	0.16
3. PG1 violated	Est.	**-0.501**	**-0.500**	**-0.500**	**-0.500**	**-0.500**
	S.E	**0.052**	**0.040**	**0.042**	**0.039**	**0.036**
	Cov^{ge}	**0.93**	**0.95**	**0.91**	**0.94**	**0.92**
4. PG1 and NUC violated	Est.	**-0.512**	-0.313	-0.282	-0.346	-0.350
	S.E	**0.051**	0.040	0.042	0.039	0.036
	Cov^{ge}	**0.94**	0.01	0.00	0.02	0.02
5. PG2 & Hom violated	Est.	-0.652	-1.001	-1.167	-0.969	-0.979
	S.E	0.053	0.040	0.043	0.039	0.036
	Cov^{ge}	0.19	0.00	0.00	0.00	0.00
6. All except PG1	Est.	-0.709	-0.761	-1.813	-0.913	-0.905
	S.E	0.092	0.081	0.140	0.079	0.079
	Cov^{ge}	0.37	0.11	0.00	0.00	0.00

Table 3: *Point estimates, standard errors and coverage (of 95% confidence interval) for the five combined estimates across six simulation scenarios. In each case, the true GMTE is fixed at -0.5. Unbiased estimates and associated standard errors/coverage are highlighted in bold.*

a standard error near-identical to the GMTE(1) estimate, in line with the theoretical prediction outlined in Appendix 6. In Scenario 3 assumption PG1 is violated and (PG2, PG3, Hom, NUC) are satisfied. In this case all estimators are unbiased. In Table 3 we show in this case that the most efficient unbiased estimate of all comes from combining the RGMTE, CAT and MR estimates. In Scenario 4 PG1 and NUC are violated but the remaining assumptions (PG2, PG3, Hom) are satisfied. In this case only the MR estimate is unbiased. Consequently, no combined estimator is unbiased. In Scenario 5, PG2 and Hom are satisfied but (PG1, PG3, NUC) are satisfied. In this case the GMTE(1) and RGMTE estimators are unbiased, with the GMTE(1) estimator being the most efficient. No combined estimate is unbiased. In Scenario 6 all assumptions except PG1 are violated. In this case only the RGMTE estimate is unbiased and, again, no combined estimate is unbiased.

In conclusion, our simulation study provides an empirical verification of the strengths and limitations of each approach, and when any two uncorrelated estimates can be effectively combined via a simple inverse variance weighted meta-analysis. Although the standard error of any estimate that combines the CAT and MR estimates requires G to be independent of T to be strictly valid (since this implies a zero correlation between their estimates) when this assumption is violated in Scenario 3 it only induces a modest loss of coverage (e.g. 91% for the CAT/MR estimate and 92% for the CAT/RGMTE/MR estimate). Across the simulations the RGMTE is emerges as the most robust estimate.

4 Applied analyses

4.1 Clopidogrel, CYPC219 & Stroke risk

Clopidogrel is a widely used anti-platelet therapy that impairs platelet aggregation with consequent reductions in risk of atherothrombotic events such as myocardial infarctions and ischemic strokes [15]. Clopidogrel is a pro-drug that requires activation by liver enzymes, primarily CYP2C19. Genetic variants in *CYP2C19* impair function with subsequently reduced clopidogrel active plasma levels [16], and we have previously shown using primary care linked data on UK Biobank participants that carriers of these variants have increased risks of ischemic stroke and myocardial infarction (MI) whilst prescribed clopidogrel in the primary care setting [6]. In this work we calculated the population attributable fraction using established methods by analysing data on only those who were treated with Clopidogrel, but we revisit the analysis and apply the full decision framework proposed in this paper.

The UK Biobank study recruited 503,325 volunteers from the community who attended one of 22 assessment centres in England, Wales or Scotland between 2006 and 2010 [17]. Participants were aged 40 to 70 years at the time of assessment, and baseline assessment included extensive questionnaires on demographic, health, and lifestyle information. Blood samples were taken, allowing analysis of participant genetics. Ethical approval for the UK Biobank study was obtained from the North West Multi-Centre Research Ethics Committee. This research was conducted under UK Biobank application 14631 (PI: DM).

Linked electronic medical records from primary care are available for 230,096 (45.7%) of participants, which includes >57 million prescribing events between 1998 and 2017. Detailed description of the data extracted and limitations are available from UK Biobank. For this analysis we excluded 5,353 participants missing any genetics data, then 14,856 of non-European genetic ancestry, then 555 missing any *CYP2C19* loss of function genotype data, leaving 209,333 participants with sufficient primary care and genetic data. N=198,868 never received a Clopidogrel prescription. N=938 only ever received one prescription, so did not have sufficient exposure time for study. Of the 9,527 participants remaining, in 2,044 the prescribing frequency was less than once every 2 months, and these were also excluded. This left 7,483 participants with at least two Clopidogrel prescriptions for analysis. Baseline information on the included participants is shown in Table 4.

CYP2C19 loss-of-function (LoF) carriers (any $*2$-$*8$ alleles) had significantly increased ischemic stroke risk (Hazard Ratio (HR) 1.53: 95% CIs 1.04 to 2.26, p=0.031) and separately MI (HR 1.14: 1.04 to 1.26, p=0.008) whilst on Clopidogrel, compared to non-carriers in Cox's proportional hazards regression models adjusted for age at first Clopidogrel prescription, sex, and the first 10 genetic principal components of ancestry.

Variable	Prescribed Clopidogrel (n=7,483)	Never Prescribed Clopidogrel n = 198,868
Age at recruitment	61.4± 6.2	56.5±8.0
Age at first prescription	64.1±7.3	-
Sex(Female%)	2,559(34.2%)	110,569(55.6%)
CYPC219 LoF carrier	2,145(28.7%)	56,043(28.2%)
Incident Ischemic Stroke diagnosis*	110 (1.5%)	2,078(1.0%)
Incident MI diagnosis*	1,822 (24.8%)	13,796(6.9%)

Table 4: *Baseline data on UK biobank participants in the Clopidogrel analysis set. *Based on hospital episode statistics data.*

4.1.1 Estimating the GMTE

To estimate the GMTE in this case we modelled the time to stroke using an Aalen additive hazards model, as described in Section 2.4. All models were adjusted for age at recruitment or first Clopidogrel prescription, sex, and the first 10 genetic principal components of ancestry. Figure 5 and Table 5 show the results for this analysis, which reflect the genetically mediated effect of Clopidogrel treatment on the hazard of stroke per year, expressed as a percentage. The GMTE(1) estimate suggests that being a CYP2C19 LoF carrier (G=1) increases the risk of stroke by 0.28% (p=0.048) compared to those without the LoF variant (G=0). To put this figure in context, if we could reduce the LoF carrier's risk by this amount then, when multiplied by the 5264 LoF carrier patient years in the data, it would lead to an expected 13.2% reduction in the total number of strokes (or a reduction of 15 strokes from 110 to 95).

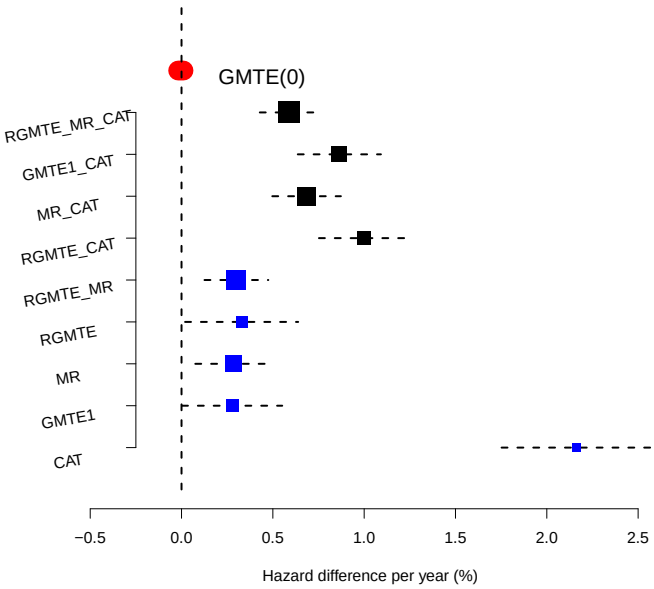

Figure 5: *Forest plot of Results for the Clopidogrel data. Blue squares show individual causal estimates as well as combined estimators that pass the heterogeneity test at the 5% level. Black squares show combined estimates that fail the heterogeneity test at the 5% level.*

To test for potential bias in the GMTE(1) estimate, we calculate the GMTE(0) estimate in the untreated population. Thankfully, it is close to zero (Hazard diff = -0.0039%, p=0.61), although slightly negative. Taken at face value, this suggest LoF carriers have a slightly reduced risk of stroke through pathways other than Clopidogrel use. Next we calculate the Corrected As Treated (CAT) estimate. As discussed, the validity of this method rests strongly on being able to identify all confounders of Clopidogrel use and stroke. With the data available, it was only possible to adjust for age, sex and genetic principal components and perhaps unsurprisingly, the CAT estimate is an order of magnitude larger (Hazard diff = 2.2%, p$\leq 2\times10^{-16}$). Consequently, the $Q_{CAT,GMTE(1)}$ statistic detects large heterogeneity and suggests that the CAT and GMTE(1) estimates should not be combined.

Estimator	Estimate (% scale)	S.E	p-value	Combined Estimates		
				Q statistic	Q p-value	Combine?
CAT	2.2	0.210	0.0e+00			
GMTE(0)	-3.9×10^{-3}	7.5×10^{-3}	6.1e-01			
GMTE(1)	0.28	0.140	4.8e-02			
MR	0.29	0.110	7.9e-03			
RGMTE	0.33	0.160	3.7e-02			
RGMTE/MR	0.3	0.089	7.5e-04	0.05	8.3e-01	Yes
RGMTE/CAT	1.0	0.130	2.0e-15	49	2.4e-12	No
MR/CAT	0.68	0.096	9.1e-13	64	1.3e-15	No
GMTE(1)/CAT	0.86	0.120	1.0e-13	56	5.9e-14	No
RGMTE/MR/CAT	0.59	0.082	6.6e-13	68	2.1e-15	No

Table 5: Hazard difference estimates (LoF carriers versus non-carriers) on percentage scale for all single and combined estimates

For completeness, we next calculate the difference between the RGMTE estimate for the GMTE hazard difference. Since this is itself the difference between the GMTE(1) and GMTE(0) estimates, and given they are of opposite sign, the RMGTE estimate is slightly larger at 0.33% (p=0.037), suggesting 17 strokes could have been avoided. The MR estimate for the GMTE hazard difference is similar at 0.29% (p=0.008). Heterogeneity analysis reveals that the MR and RGMTE estimates are sufficiently similar to combine into a more precise single estimate of the GMTE ($Q_{MR,RGMTE}$=0.8). The combined estimate is 0.3 (p=7.5×10^{-04}), or that 16 strokes could have been avoided. No other combination of estimates are sufficiently similar to combine.

4.2 Statins, *APOE* & CAD

We now apply our framework to estimate the extent to which genetic variation at the *APOE* locus modulates the risk of coronary artery disease (CAD) due to statin treatment using UK Biobank data. Our full data comprises 155,409 unrelated participants of European descent, with primary care data available (updated to March 2017) and up-to-date hospital admission data as of December 2020. Of this sample, we excluded: n=11 participants with missing *APOE* genotypes; n=6,456 non-regular statin users with less than four prescriptions per year or residuals from the linear regression for total statin prescriptions on years of statin treatment greater than 3 or less than -3; n=1,273 non-statin users diagnosed with CAD at baseline (or prior to baseline); n=4,566 participants starting statin after a doctor's diagnosis of coronary artery disease (CAD) based on the hospital admission records.

Among the included samples (n=143,103), 57,682 (59.5%) were female. Of these, 46,179 (32.3%) were statin users, with a median of 9.4 (inter-quartile range: 6.6 to 13.5) statin prescriptions per year and a median of 5.6 (inter-quartile range: 1.2 to 9.9) years of statin treatment. Several SNPs were associated with LDL cholesterol response to statins based on a genome-wide association study, where the *APOE* e2 defining SNP rs7412 showed a larger LDL cholesterol lowering response to statins compared to e3e3s [18]. APOE genotypes (diplotypes essentially) were determined based on genotypes at rs7412 and rs429358. Inspecting the *APOE* genotype distribution, the majority of participants were classed as e_3e_3 (n=83,813, 58.6%), followed by e_3e_4 (n=33,597, 23.5%), e_2e_3s (n=17,811, 12.4%), e_2e_4s (n=3,616, 2.5%), e_4e_4s (n=3,366, 2.4%), and e_2e_2s (n=900, 0.6%). These groups are mutually exclusive. Summary statistics for statin users and non-statin users are presented in Table 6.

4.3 Results

Using the $e3e3$ group as a reference, we fitted Aalen additive hazard models within each mutually exclusive genetic group additionally adjusting for sex, age on statin or age at recruitment, and the top 10 genetic principal components. For brevity, we focus on the results of the e2e2 versus e3e3 and e4e4 versus e3e3 analyses, which are shown in Table 7 and account for approximately 72% of the patient data. Estimates reflect the hazard or risk difference of a CAD event per year, expressed as a percentage. Only results of combined estimates that pass a heterogeneity test at the 5% level are shown. Equivalent

	Statin Users (n=46,179)	Non-Statin Users (n=96,924)
Age starting statin (left)		
age at recruitment (right)	60.5± 7.9	54.7± 8.0
Sex (=Female)	20,921 (45.3%)	57,682 (59.5%)
APOE genotype		
e_3e_3	26,938 (58.3%)	56,875 (58.7%)
e_2e_2	258 (0.6%)	642 (0.7%)
e_2e_3	4,772 (10.3%)	13,039 (13.5%)
e_2e_4	1,065 (2.3%)	2,551 (2.6%)
e_3e_4	11,849 (25.7%)	21,748 (22.4%)
e_4e_4	1,297 (2.8%)	2,069 (2.1%)
Incident CAD (MI or angina) cases	7,259 (15.7%)	2,758 (2.8%)

Table 6: Baseline covariates, genetic data and incident CAD cases on statin users and non-users in UK Biobank.

estimates for the remaining genetic groups showed no evidence of a non-zero genetically mediated effect. Results for all genetic groups are given in Appendix 7.

Estimator	Estimate	S.E	p-value	Expected Avoided CAD events in e3e3 group (95% CI)
		e4e4 versus e3e3		
CAT	3.9000	0.082	0.000	-
GMTE1	-0.0310	0.015	0.043	-85 (-168, -3)
MR	0.0069	0.017	0.690	19 (-76,115)
RGMTE	-0.0370	0.018	0.046	-103 (-204,-2)
GMTE0	0.0110	0.0063	0.073	-
RGMTE/MR	-0.0140	0.013	0.280	-39 (-108,31)
		e2e3 versus e3e3		
CAT	1.2000	0.0240	0.0e+00	-
GMTE1	-0.0100	0.0087	2.4e-01	-28 (-75,19)
MR	-0.0460	0.0110	3.4e-05	-128 (-189,-67)
RGMTE	-0.0055	0.0098	5.7e-01	-15 (-69,38)
GMTE0	-0.00014	0.0025	9.5e-01	-

Table 7: Hazard difference estimates on the % scale for all single and valid combined estimates for the e2e3 and e4e4 genotype groups

4.3.1 e4e4 versus e3e3

Inspecting the $e4e4$ genetic subgroup first, the GMTE(1) estimate suggests that the risk of CAD could be reduced by 0.031% per year if e3e3 patients experienced the same treatment effect as e4e4 patients (p=0.043). This estimate is valid if the e4e4 genotype only affects the risk of CAD through modulating the effectiveness of statins (i.e. assumptions PG1-PG3 hold). In order to probe this we calculate the equivalent GMTE(0) estimate in non-statin users. The e_4e_4 group is now seen to have a 0.011% larger risk of CAD than $e3e3$ (p=0.07), which suggests that PG1-PG3 violation is possible. Since the RMGTE(0) and RGMTE(1) estimates are of opposite signs, the RGMTE estimate, which is robust to PG2-PG3 violation, infers the risk difference between e4e4 and e3e3's is larger at -0.037% per year (p=0.046). The MR estimate of the GMTE is also positive (0.0069%), but very close to zero (p=0.69). This is, however, sufficiently similar to the RGMTE estimate for it to be combined with the MR estimate, and the combined value suggests a hazard difference of -0.014% per year (p=0.28). The CAT estimate for the hazard difference in these data is a 3.9% increase per year. Its magnitude is so large compared to the other estimates that we can infer adjustment for age, sex and genetic PCs is again not sufficient to remove confounding by indication. Consequently, no other estimate is sufficiently similar in order to combine

with the CAT estimate, as shown in Figure 6 (left). In the final column of Table 7 we translate the hazard difference estimate per year implied by the GMTE1, MR, RGMTE and combined RMGTE/MR estimate to give an expected number of CAD events that could be avoided if all 26,938 e3e3 statin user patients could receive the same benefit as the e4e4 patients, by multiplying the per-year risk reduction over the relevant 278,409 patients-years in the data. Using the RGMTE estimate for this risk reduction gives a figure of 103. The GMTE1 and combined RGMTE/MR estimates imply more modest reductions of 85 and 39 CAD events respectively.

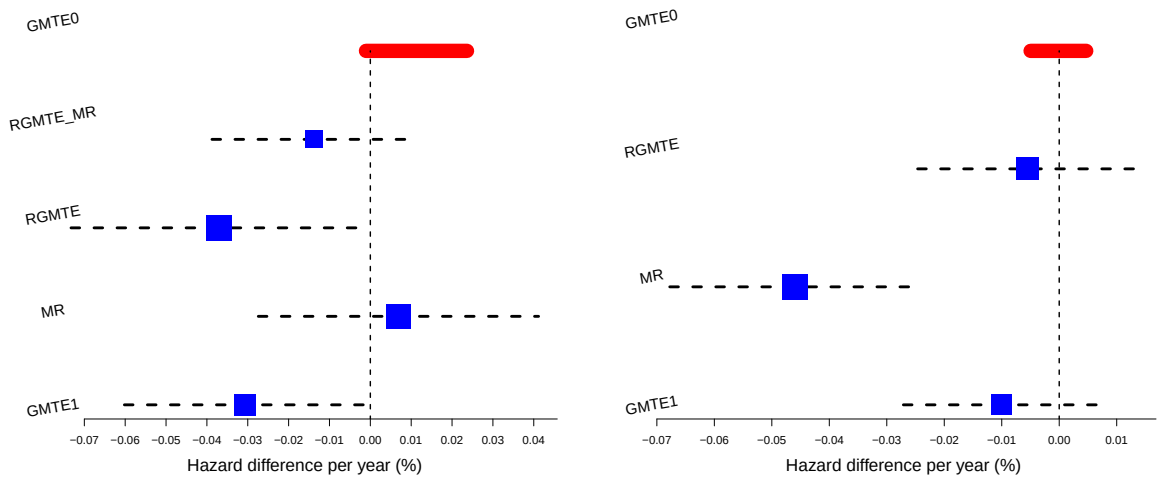

Figure 6: *Hazard difference estimates for the e4e4 versus e3e3 (left) and e2e3 versus e3e3 (right) analyses.*

4.3.2 e2e3 versus e3e3

Turning our attention to the e2e3 subgroup in Table 7 and Figure 6 (right), we again see a large, non-credible CAT hazard difference estimate for the GMTE of a 1.2% per year between e2e3 and e3e3 groups. The GMTE(1), GMTE(0) and RGMTE estimates for the GMTE are all close to zero and non-significant at the 5% level. In contrast, the MR estimate for the GMTE suggests that e2e3's have a 0.046% reduced risk of CAD per year ($p=3\times10^{-5}$). Using this estimate, the expected number of CAD events that could be avoided if all 26,938 e3e3 patients could receive the same benefit as the e2e3 patients is 128. This is valid if we believe that assumptions PG2-PG3 hold, but either PG1 or the Homogeneity assumption are violated. The GMTE1 and RGMTE estimates imply more modest reductions of 28 and 15 CAD events respectively. In this example, no two single uncorrelated estimates are sufficiently similar in order to combine.

5 Discussion

In this paper we propose a general framework for estimating the genetically mediated treatment effect that combines several distinct but complementary causal inference techniques. We propose a rudimentary decision framework for choosing when to combine approaches based on heterogeneity statistics. In practice, expert knowledge and prior evidence should also be leveraged to decide whether the particular assumptions of the causal estimation strategy are likely to be met, in order to put more or less weight on their findings. Our inverse variance meta-analysis procedure for combining estimates is very simple, and simulations showed that exhibited good statistical properties even when small correlations between constituent estimates were present. As future work, we plan to develop a more sophisticated procedure to explicitly account for this correlation. We also plan to develop a hierarchical testing procedure that can be applied across the TWIST framework to control family-wise or false-discovery error rates.

In our analysis of the statin data, we estimated the GMTE in several mutually exclusive genetic groups, which resulted in an inevitable loss of precision. Efficiency could potentially be regained by collapsing genetic subsets together if they give similar estimates, or by making a linearity assumption about the magnitude of effect across genotypic groups (e.g. between e3e3, e3e4 and e4e4). This would not be

defensible if the genetic groups were ordered with respect to the magnitude of their causal estimate, but would be defensible if genetic groups could be ordered by their effect on increasing drug metabolism. In the case of 3 genetic groups, G_i and T_i^* could take a value in {0,1,2}. This would enable the data to be pooled in order to target a combined estimand

$$\beta_{GMTE}(Y) = E[Y_i(g, T(m))] - E[Y_i(g, T(m-1)], \tag{12}$$

for all m in $\{1, 2\}$. If such a model were correct, it opens up the possibility of making the analysis even more robust to violations of the PG assumptions, because an additional causal parameter could be jointly estimated alongside the GMTE to reflect, for example the direct effect of G on Y. This is an important avenue for further research.

Although the CAT estimate can, in principle consistently estimate the GMTE estimand, it relies heavily on the NUC assumption. In both applied analyses we were not able to sufficiently control for confounding by indication to deliver an estimate close to any other GMTE estimate, due to a lack of relevant covariate data. In future work we plan to revisit both analyses after collating a much larger set of relevant information. More-sophisticated approaches such as Propensity Scores, matching methods and inverse probability weighting may then offer some utility [19]. So too may methods for multi-variable Mendelian randomization, where instead of directly adjusting for confounders of treatment and outcome, we instead adjust for their genetically predicted value. This latter approach could be more robust to collider bias [10].

In supplementary material we provide R code for fitting the TWIST framework for continuous, binary and time-to-event data as well as code used in the simulation study. Work is ongoing to produce a single R package to apply TWIST and visualise its results.

References

[1] George Davey Smith and Shah Ebrahim. "'Mendelian randomization': can genetic epidemiology contribute to understanding environmental determinants of disease?" In: *International Journal of Epidemiology* 32.1 (Feb. 2003), pp. 1–22. ISSN: 0300-5771. DOI: 10.1093/ije/dyg070. eprint: http://oup.prod.sis.lan/ije/article-pdf/32/1/1/11213726/320001_dyg070.pdf. URL: https://doi.org/10.1093/ije/dyg070.

[2] Michael V. Holmes et al. "CYP2C19 Genotype, Clopidogrel Metabolism, Platelet Function, and Cardiovascular Events: A Systematic Review and Meta-analysis". In: *JAMA* 306.24 (Dec. 2011), pp. 2704–2714. ISSN: 0098-7484. DOI: 10.1001/jama.2011.1880. eprint: https://jamanetwork.com/journals/jama/articlepdf/1105555/jcr15007_2704_2714.pdf. URL: https://doi.org/10.1001/jama.2011.1880.

[3] Demetrios N. Kyriacou and Roger J. Lewis. "Confounding by Indication in Clinical Research". In: *JAMA* 316.17 (Nov. 2016), pp. 1818–1819. ISSN: 0098-7484. DOI: 10.1001/jama.2016.16435. eprint: https://jamanetwork.com/journals/jama/articlepdf/2576568/jgm160006.pdf. URL: https://doi.org/10.1001/jama.2016.16435.

[4] P J Veugelers and A M Yip. "Socioeconomic disparities in health care use: Does universal coverage reduce inequalities in health?" In: *Journal of Epidemiology & Community Health* 57.6 (2003), pp. 424–428. ISSN: 0143-005X. DOI: 10.1136/jech.57.6.424. eprint: https://jech.bmj.com/content/57/6/424.full.pdf. URL: https://jech.bmj.com/content/57/6/424.

[5] Nancy Krieger et al. "Hormone replacement therapy, cancer, controversies, and women's health: historical, epidemiological, biological, clinical, and advocacy perspectives". In: *Journal of Epidemiology & Community Health* 59.9 (2005), pp. 740–748. ISSN: 0143-005X. DOI: 10.1136/jech.2005.033316. eprint: https://jech.bmj.com/content/59/9/740.full.pdf. URL: https://jech.bmj.com/content/59/9/740.

[6] Luke C. Pilling et al. "Genetic variation in activating clopidogrel: longer-term outcomes in a large community cohort". In: *medRxiv* (2021). DOI: 10.1101/2021.04.19.21255559. eprint: https://www.medrxiv.org/content/early/2021/04/20/2021.04.19.21255559.full.pdf. URL: https://www.medrxiv.org/content/early/2021/04/20/2021.04.19.21255559.

[7] Debbie A Lawlor, Kate Tilling, and George Davey Smith. "Triangulation in aetiological epidemiology". In: *International Journal of Epidemiology* 45.6 (Jan. 2017), pp. 1866–1886. ISSN: 0300-5771. DOI: 10.1093/ije/dyw314. eprint: https://academic.oup.com/ije/article-pdf/45/6/1866/32531204/dyw314.pdf. URL: https://doi.org/10.1093/ije/dyw314.

[8] Jacklyn N. Hellwege et al. "Population Stratification in Genetic Association Studies". In: *Current Protocols in Human Genetics* 95.1 (2017), pp. 1.22.1–1.22.23. DOI: https://doi.org/10.1002/cphg.48. eprint: https://currentprotocols.onlinelibrary.wiley.com/doi/pdf/10.1002/cphg.48. URL: https://currentprotocols.onlinelibrary.wiley.com/doi/abs/10.1002/cphg.48.

[9] Gibran Hemani, Jack Bowden, and George Davey Smith. "Evaluating the potential role of pleiotropy in Mendelian randomization studies". In: *Human Molecular Genetics* 27.R2 (May 2018), R195–R208. ISSN: 0964-6906. DOI: 10.1093/hmg/ddy163. eprint: https://academic.oup.com/hmg/article-pdf/27/R2/R195/25229892/ddy163.pdf. URL: https://doi.org/10.1093/hmg/ddy163.

[10] Marcus R Munafò et al. "Collider scope: when selection bias can substantially influence observed associations". In: *International Journal of Epidemiology* 47.1 (Sept. 2017), pp. 226–235. ISSN: 0300-5771. DOI: 10.1093/ije/dyx206. eprint: http://oup.prod.sis.lan/ije/article-pdf/47/1/226/25088130/dyx206.pdf. URL: https://doi.org/10.1093/ije/dyx206.

[11] Alberto Abadie. "Semiparametric Difference-in-Differences Estimators". In: *The Review of Economic Studies* 72.1 (Jan. 2005), pp. 1–19. ISSN: 0034-6527. DOI: 10.1111/0034-6527.00321. eprint: https://academic.oup.com/restud/article-pdf/72/1/1/18327455/72-1-1.pdf. URL: https://doi.org/10.1111/0034-6527.00321.

[12] Andrew Gelman and Iain Pardoe. "Average predictive comparisons for models with nonlinearity, interactions, and variance components". In: *Sociological Methodology* 37.1 (2007), pp. 23–51.

[13] Eric Tchetgen Tchetgen et al. "Instrumental variable estimation in a survival context". In: *Epidemiology* 26.3 (2015), pp. 402–410.

[14] Anders Huitfeldt, Mats J. Stensrud, and Etsuji Suzuki. "On the collapsibility of measures of effect in the counterfactual causal framework". In: *Emerging Themes in Epidemiology* 16.1 (Jan. 2019), p. 1. ISSN: 1742-7622. DOI: 10.1186/s12982-018-0083-9. URL: https://doi.org/10.1186/s12982-018-0083-9.

[15] Jessica L. Mega et al. "Cytochrome P-450 Polymorphisms and Response to Clopidogrel". In: *New England Journal of Medicine* 360.4 (2009). PMID: 19106084, pp. 354–362. DOI: 10.1056/NEJMoa0809171. eprint: https://doi.org/10.1056/NEJMoa0809171. URL: https://doi.org/10.1056/NEJMoa0809171.

[16] Tabassome Simon and Nicolas Danchin. "Clinical Impact of Pharmacogenomics of Clopidogrel in Stroke". In: *Circulation* 135.1 (2017), pp. 34–37. DOI: 10.1161/CIRCULATIONAHA.116.025198. eprint: https://www.ahajournals.org/doi/pdf/10.1161/CIRCULATIONAHA.116.025198. URL: https://www.ahajournals.org/doi/abs/10.1161/CIRCULATIONAHA.116.025198.

[17] Cathie Sudlow et al. "UK Biobank: An Open Access Resource for Identifying the Causes of a Wide Range of Complex Diseases of Middle and Old Age". In: *PLOS Medicine* 12.3 (Mar. 2015), pp. 1–10. DOI: 10.1371/journal.pmed.1001779. URL: https://doi.org/10.1371/journal.pmed.1001779.

[18] Iris Postmus et al. "Pharmacogenetic meta-analysis of genome-wide association studies of LDL cholesterol response to statins". In: *Nature Communications* 5.1 (Oct. 2014), p. 5068. ISSN: 2041-1723. DOI: 10.1038/ncomms6068. URL: https://doi.org/10.1038/ncomms6068.

[19] Elizabeth A. Stuart. "Matching Methods for Causal Inference: A Review and a Look Forward". In: *Statistical Science* 25.1 (2010), pp. 1–21. DOI: 10.1214/09-STS313. URL: https://doi.org/10.1214/09-STS313.

Appendix 1: Consistency of the GMTE(1) estimate

Under model (5) we have that the estimand targeted by the $\hat{\beta}_{GMTE(1)}(Y)$ is equal to

$$
\begin{aligned}
E[Y|T=1,G=1]-E[Y|T=1,G=0] &= \gamma_{Y0}+\beta_0+(\beta_1-\beta_0)+\gamma_{GY}+\gamma_{YU}E[U|T=1,G=1] \\
&- (\gamma_{Y0}+\beta_0+\gamma_{YU}E[U|T=1,G=0]) \\
&= (\beta_1-\beta_0)+\gamma_{GY}+\gamma_{YU}\left\{E[U|T=1,G=1]-E[U|T=1,G=0]\right\} \\
&= \beta_{GMTE(1)}+\gamma_{GY}+\gamma_{YU}\beta_{GMTE(1)}(U)
\end{aligned}
$$

For this estimand to equal β_{GMTE} we require assumption PG3 (γ_{GY}=0) and that that $G \perp\!\!\!\perp U|T=1$.

Under model (5) we have that the estimand targeted by the $\hat{\beta}_{GMTE(0)}(Y)$ is equal to

$$
\begin{aligned}
E[Y|T=0,G=1]-E[Y|T=0,G=0] &= \gamma_{Y0}+\gamma_{GY}+\gamma_{YU}E[U|T=0,G=1] \\
&- (\gamma_{Y0}+\gamma_{YU}E[U|T=0,G=0]) \\
&= \gamma_{GY}+\gamma_{YU}\left\{E[U|T=0,G=1]-E[U|T=0,G=0]\right\} \\
&= \gamma_{GY}+\gamma_{YU}\beta_{GMTE(0)}(U)
\end{aligned}
$$

For this estimand to equal β_{GMTE} we require assumption PG3 (γ_{GY}=0) and that that $G \perp\!\!\!\perp U|T=0$.

Appendix 2: Consistency of the RGMTE estimate

Under model (5) we have that RGMTE(Y) estimand is equal to

$$
\begin{aligned}
& E[Y|T=1,G=1]-E[Y|T=1,G=0]-(E[Y|T=0,G=1]-E[Y|T=0,G=0]) \\
= \;& (\beta_1-\beta_0)+\gamma_{GY}+\gamma_{YU}\left\{E[U|T=1,G=1]-E[U|T=1,G=0]\right\} \\
- \;& \gamma_{GY}+\gamma_{YU}\left\{E[U|T=0,G=1]-E[U|T=0,G=0]\right\} \\
= \;& \beta_{GMTE}+\gamma_{YU}\left(\left\{E[U|T=1,G=1]-E[U|T=1,G=0]\right\}-\left\{E[U|T=0,G=1]-E[U|T=0,G=0]\right\}\right) \\
= \;& \beta_{GMTE}+\gamma_{YU}\beta_{RGMTE}(U)
\end{aligned}
$$

The RGMTE(Y) estimand is therefore equivalent to β_{GMTE} if the RGMTE(U) estimand is zero. This will be zero whenever the strength of association between U and G is independent of T or that

$$Cov(U,G) \perp\!\!\!\perp T$$

Appendix 3: Consistency of the MR estimate

Using equation (3) as the data generating model, the MR estimand is equal to:

$$\beta_{MR}(Y)=\frac{E[Y|G=1]-E[Y|G=0]}{E[T^*|G=1]-E[G|T=1]} \tag{13}$$

We first note that the denominator of the MR estimand simplifies to $E[T=1|G=1]$ since $E[T^*|G=1]=E[TG|G=1]=E[T|G=1]$ and that $E[T^*|G=0]=E[TG|G=0]=0$. We can then write the estimand as

$$
\begin{aligned}
\beta_{MR}(Y)=\frac{E[Y|G=1]-E[Y|G=0]}{E[T|G=1]} &= (\beta_1-\beta_0) \\
&+ \frac{\beta_0(E[T|G=1]-E[T|G=0])}{E[T|G=1]} \\
&+ \frac{\gamma_{YU}(E[U|G=1]-E[U|G=0])}{E[T|G=1]} \\
&+ \frac{\gamma_{GY}}{E[T|G=1]} \\
&= \beta_{GMTE} \\
&+ \beta_0\left(1-\frac{E[T|G=0]}{E[T|G=1]}\right)+\gamma_{YU}\beta_{MR}(U)+\frac{\gamma_{GY}}{E[T|G=1]}
\end{aligned}
$$

From this we can see that the consistency of the MR estimate relies unambiguously on assumptions PG2 and PG3 being satisfied in order for the second and third bias terms to disappear, but there are two possible ways for the first bias term to disappear: Either assumption PG1 is satisfied, so that $E[T|G=1]$-$E[T|G=0]=0$, or the Homogeneity assumption is satisfied, so that $\beta_0=0$. We can therefore write the identifying condition of the MR estimand as

$$\{PG1 \cup Hom\} \cap PG2 \cap PG3$$

Note that the magnitude of the bias terms in the MR estimand will be inversely proportional to the probability of being treated given genotype G=1.

Appendix 4: Consistency of the CAT estimate

Under model (5) we have that CAT(Y) estimand is equal to

$$\beta_{CAT}(Y) = \frac{E[Y|T=1]-E[Y|T=0]}{E[G|T=1]},$$

The numerator of the estimand equals

$$\begin{aligned} & \beta_0 + (\beta_1-\beta_0)E[G|T=1] + \gamma_{GY}E[G|T=1] + \gamma_{YU}E[U|T=1] - (\gamma_{GY}E[G|T=0] + \gamma_{YU}E[U|T=0]) \\ = \; & \beta_0 + \beta_{GMTE}E[T|G=1] + \gamma_{GY}\left\{E[G|T=1]-E[G|T=0]\right\} + \gamma_{YU}\left\{E[U|T=1]-E[U|T=0]\right\} \end{aligned} \tag{14}$$

Dividing through by $E[G|T=1]$ we see that the full estimand equals

$$\beta_{CAT}(Y) = \beta_0 + \beta_{GMTE} + \gamma_{GY}\beta_{CAT}(G) + \gamma_{YU}\beta_{CAT}(U)$$

The CAT estimand is therefore equivalent to β_{GMTE} if the Homogeneity and the NUC assumptions hold, and either $G \perp\!\!\!\perp T$ or PG3 hold. We can therefore express the identifying condition of the GMTE(1) estimand as

$$\{PG1 \cup PG3\} \cap U \perp\!\!\!\perp T \cap Hom$$

Appendix 5: Which estimates are uncorrelated?

Let $\bar{Y}_{ij} = \hat{E}[Y|T=i, G=j]$, $\text{Var}(\bar{Y}_{ij}) = \sigma^2$, $\pi_{ij} = \hat{E}[T=i|G=j]$ and $\tau_{ij} = \hat{E}[G=j|T=i]$. We now write the CAT, GMTE(1), RGMTE and MR estimates as

$$\begin{aligned} \hat{\beta}_{CAT(Y)} &= \frac{\bar{Y}_{10}\tau_{10} + \bar{Y}_{11}\tau_{11} - (\bar{Y}_{00}\tau_{00} + \bar{Y}_{01}\tau_{01})}{\tau_{11}} \\ \hat{\beta}_{GMTE(1)}(Y) &= \bar{Y}_{11} - \bar{Y}_{10} \\ \hat{\beta}_{MR(Y)} &= \frac{\bar{Y}_{11}\pi_{11} + \bar{Y}_{01}\pi_{01} - (\bar{Y}_{10}\pi_{10} + \bar{Y}_{00}\pi_{00})}{\pi_{11}} \\ \hat{\beta}_{RGMTE}(Y) &= \bar{Y}_{11} - \bar{Y}_{10} - (\bar{Y}_{01} - \bar{Y}_{00}) \end{aligned}$$

Ignoring the denominator of $\hat{\beta}_{MR}$, the covariance between its numerator and $\hat{\beta}_{RGMTE}$ can be shown to equal zero. Let n_{ij} equal the number of subjects with variables $T=i$ and $G=j$ in the sample and let n_{Gj} equal the number of subjects with variable $G=j$ in the sample. It then follows that:

$$\begin{aligned} \pi_{11}\text{Cov}(\hat{\beta}_{MR}(Y), \hat{\beta}_{RGMTE}(Y)) &= \pi_{11}\text{Var}(\bar{Y}_{11}) - \pi_{10}\text{Var}(\bar{Y}_{10}) - \pi_{01}\text{Var}(\bar{Y}_{01}) + \pi_{00}\text{Var}(\bar{Y}_{00}) \\ &= \pi_{11}\frac{\sigma^2}{n_{11}} + \pi_{10}\frac{\sigma^2}{n_{10}} - \pi_{01}\frac{\sigma^2}{n_{01}} - \pi_{00}\frac{\sigma^2}{n_{00}} \\ &= \frac{\sigma^2}{n_{G1}} + \frac{\sigma^2}{n_{G1}} - \frac{\sigma^2}{n_{G0}} - \frac{\sigma^2}{n_{G0}} \\ &= 0 \end{aligned}$$

Using similar arguments it is easy to show that

$$\begin{aligned}
\tau_{11}\text{Cov}(\hat{\beta}_{CAT}(Y), \hat{\beta}_{RGMTE}(Y)) &= 0 \\
\tau_{11}\pi_{11}\text{Cov}(\hat{\beta}_{CAT}(Y), \hat{\beta}_{MR}(Y)) &= 0 \text{ if } G \perp\!\!\!\perp T, \\
\text{Cov}(\hat{\beta}_{CAT}(Y), \hat{\beta}_{GMTE(1)}(Y)) &= 0
\end{aligned}$$

In cases where $G \perp\!\!\!\perp T$ does not hold, but G predicts a relatively small amount of variation in T, the correlation between the CAT and MR estimate will be non-zero but negligible. This is a reasonable assumption in almost all pharmacogenetic contexts. This means that there are effectively only two pairs of correlated estimates: (GMTE(1), RGMTE), and (GMTE(1), MR).

Appendix 6: When does $\hat{\beta}_{RGMTE,MR}(Y)$=$\hat{\beta}_{GMTE(1)}(Y)$?

Using the notation from the previous appendix, we can then write the MR estimand as

$$\begin{aligned}
\frac{E[Y|G=1] - E[Y|G=0]}{E[T|G=1]} &= \frac{\bar{Y}_{01}\pi_{01} + \bar{Y}_{11}\pi_{11} - (\bar{Y}_{00}\pi_{00} + \bar{Y}_{10}\pi_{10})}{\pi_{11}} \\
&= \frac{\bar{Y}_{11}\pi_{11} - \bar{Y}_{10}\pi_{10} + (\bar{Y}_{01}\pi_{01} + \bar{Y}_{00}\pi_{00})}{\pi_{11}}
\end{aligned}$$

Under the assumption that $G \perp\!\!\!\perp T$ we have that $\pi_{10} = \pi_{11} = Pr(T=1) = \pi_1$, $\pi_{00} = \pi_{01} = Pr(T=0)$ $= \pi_0$ and the MR estimand equals

$$\beta_{GMTE(1)}(Y) + \beta_{GMTE(0)}(Y)\frac{\pi_1}{\pi_0}$$

The RGMTE estimand is $\beta_{GMTE(1)}(Y) - \beta_{GMTE(0)}(Y)$. Therefore, when additionally $Pr(T=1) = Pr(T=0)$, both the RGMTE and MR estimates will have equal precision and the GMTE(1) estimate will be the average of the RGMTE estimand and the MR estimand. In general, an inverse variance weighted average of the RGMTE and MR estimates is likely to be a close approximation to the GMTE(1) estimate.

Appendix 7: full results for the APOE, Statin and CAD analysis

Estimator	Estimate	S.E	p-value
e2e2 versus e3e3			
CAT	18.800	0.407	0.000
GMTE1	0.004	0.036	0.906
MR	-0.006	0.047	0.892
RGMTE	-0.002	0.038	0.961
RGMTE/MR	-0.004	0.030	0.902
GMTE0	-0.003	0.010	0.725
e2e3 versus e3e3			
CAT	1.180	0.024	0.000
GMTE1	-0.010	0.009	0.245
MR	-0.046	0.011	0.000
RGMTE	-0.006	0.010	0.571
GMTE0	0.000	0.002	0.954
e2e4 versus e3e3			
CAT	4.700	0.100	0.000
GMTE1	-0.005	0.017	0.797
MR	-0.022	0.021	0.288
RGMTE	-0.001	0.020	0.977
RGMTE/MR	-0.011	0.015	0.452
GMTE0	0.002	0.005	0.666
e3e4 versus e3e3			
CAT	0.580	0.011	0.000
GMTE1	-0.003	0.006	0.586
MR	0.011	0.007	0.107
RGMTE	-0.002	0.007	0.786
RGMTE/MR	0.005	0.005	0.346
GMTE0	0.001	0.002	0.615
e4e4 versus e3e3			
CAT	3.860	0.082	0.000
GMTE1	-0.031	0.015	0.043
MR	0.007	0.017	0.694
RGMTE	-0.037	0.018	0.046
RGMTE/MR	-0.014	0.013	0.276
GMTE0	0.011	0.006	0.073

Table 8: Hazard difference estimates on the % scale for all single and valid combined estimates