Published in final edited form as:
Pharmacogenet Genomics. 2014 March ; 24(3): 146–155. doi:10.1097/FPC.0000000000000027.

Comparison of study designs used to detect and characterize pharmacogenomic interactions in non-experimental studies: a simulation study

Christy L. Avery, **Jane S. Der**, **Eric A Whitsel**, and **Til Stürmer**
Departments of Epidemiology (CLA, JSD, EAW, TS) and Medicine (EAW), The University of North Carolina at Chapel Hill, Chapel Hill, NC, 27514

Abstract

Objectives—Adverse drug reactions are common, serious, difficult to predict, and may be influenced by genetics, prompting the increasing popularity of pharmacogenomic studies. Many pharmacogenomic studies are conducted in non-experimental settings, yet little is known about the influence of confounding by contraindication. We therefore compared the two designs (the overall population (OPD) and the treated-only (TOD) design) by simulating a pharmacogenomic study of the electrocardiographic QT interval (QT).

Methods—Simulations were informed by data from the Atherosclerosis Risk in Communities Study and a literature review examining QT, QT-prolonging drug use, and modification by single nucleotide polymorphisms (SNP). Drug treatment was assigned based on age, gender, and QT_{long}, representing confounding by contraindication. QT was simulated as a function of drug treatment, one SNP, the drug-SNP interaction, and clinical covariates.

Results—Failure to adjust for confounding by contraindication produced a varying degree of bias in the OPD, while the TOD was biased by the SNP main effect. For example, in the OPD, the false positive proportion (FPP) for the drug-SNP interaction was 5% across the range of SNP main effects (0–10 ms), but increased to 19% without adjusting for confounding by contraindication. In the TOD, the FPP increased to 89% with SNP main effects >4 ms, although bias was reduced by 39% with adjustment for covariates affected by the SNP.

Conclusions—The potential for bias from confounding by contraindication (OPD) should be weighed against bias from SNP main effects (TOD) when selecting the study design that best suits the given context.

Keywords

confounding; study design; QT interval; genetic epidemiology

INTRODUCTION

Adverse drug reactions (ADR) cause approximately 100,000 deaths and 2.2 million serious events annually in the United States and occur even when drugs are administered according to clinical guidelines [1]. Although factors that increase susceptibility to ADRs are not always clear [2], genetic factors may improve their prediction. For example, the U.S. Food and Drug Administration had approved labeling that incorporates information on genetic variants affecting drug metabolism for upwards of 70 drugs [3], including warfarin, where

Corresponding author: Christy Avery, Department of Epidemiology, University of North Carolina Chapel Hill, Bank of America Center, 137 E. Franklin St., Suite 306, Chapel Hill, NC 27514, (919) 966-4312 (voice), (919) 966-9800 (fax), christy_avery@unc.edu.

genes encoding the VKORC1 and CYP2C9 proteins are now routinely evaluated when assigning dosage regimens [4–6]. Other examples of the use of pharmacogenomics to avoid ADRs include *ITPA* assessment to prevent anemia during hepatitis C treatment [7] and multigene assays to predict the recurrent of tamoxifen-treated, node-negative breast cancer [8]. Thus, an improved understanding of genetic susceptibility underlying ADRs is of interest to patients, clinicians, the pharmaceutical industry, and regulatory agencies.

Contemporary pharmacogenomic research typically leverages exposure and outcome data from randomized trials [9–11] or non-experimental studies [12,13]. Analytical approaches also vary, with some investigators favoring a treated-only design (TOD) that limits confounding by contraindication (i.e. when a factor is associated with the avoidance of a treatment and independently influences the outcome) because every participant is treated [14,15]. Others employ an overall population design (OPD) that includes treated and untreated participants and addresses confounding by contraindication by adjustment [16,17], which allows estimation of interaction on different scales as well as the separation of SNP main effects from gene-environment effects. However, no studies have contrasted these two commonly used pharmacogenomic study designs to determine whether scenarios exist where one study design might be preferred over the other.

We therefore conducted a series of simulations contrasting the OPD and TOD. The simulations were informed by empirical data from the Atherosclerosis Risk in Communities study (ARIC), an ongoing and population-based epidemiologic study in four United States communities [18]. QT interval duration (QT), a common measure of ventricular repolarization that is heritable [19,20], reliable [21], heterogeneous among populations exposed to arrhythmogenic drugs [22,23], and associated with life-threatening cardiac arrhythmias [24,25], served as the simulation model.

METHODS

Simulation overview

To design the simulations, we conducted a literature review to inform conceptualization of the relationship between QT, QT-prolonging drug use, and modification by single nucleotide polymorphisms (SNP, Figure 1) using published clinical and genome-wide association studies (GWAS) to establish plausible effect sizes for drug, SNP, and drug-SNP effects on QT [10,26–28] as well as risk factors for QT prolongation [29]. Based on the review, a directed acyclic graph (DAG) [30,31] was constructed, which is a causal diagram used to identify variables that must be controlled to obtain an unbiased effect estimate [29, 30]. Our DAG included two study visits (visit 1=baseline and visit 2=follow-up). QT (milliseconds, ms) was measured at each visit and QT_2 was the dependent variable. QT-prolonging drug use, initiated between visit 1 and visit 2 and measured at visit 2, was the exposure of interest. One SNP was then included as a modifier of the drug-QT_2 effect. By design, QT_1 and QT_2 were affected by the SNP because QT is heritable [19,20]. In addition to the SNP, we also included three covariates affecting QT_1 and QT_2 (sex, age, and U, representing unknown/unmeasured variation in QT). Drug treatment was affected by sex and age (confounders of the drug-QT association). Finally, QT_{long} was created from QT_1 to represent confounding by contraindication and was based on the assumption that participants with a longer QT were less likely to be prescribed QT-prolonging medications than those without a prolonged QT. QT_{long} was not included in the DAG because it was derived from QT_1 and therefore does not belong to the causal structure. QT_1 also did not affect QT_2 by design outside the drug because QT measured at one occasion does not have a causal effect on QT measured at another point in time. Instead, the correlation between the QT_1 and QT_2 was a function of unmeasured covariates (“U”, e.g. congenital heart disease, drugs, and SNPs other than the SNP included in the simulation, hypothyroidism, hypokalemia, and

myocardial infarction, among others). Consistent with DAG theory, the SNP and U did not confound the drug-QT association if we controlled for QT_{long} or QT_1. Further, using DAG methodology [30,31], the SNP-QT_2 association is unconfounded because there are no unblocked backdoor paths.

Simulation parameters and values

Age for each observation was simulated using a normal distribution with a mean (54 years) and standard deviation (6 years) equal to that observed at the ARIC baseline visit. Sex was simulated as a uniform random variable with a defined probability of being male (45%). A uniform distribution also was used to simulate SNP genotype according to a pre-specified minor allele frequency (MAF, ranging 0.05–0.45). The probabilities of a participant being heterozygous or homozygous for the major or minor allele were calculated under the assumption of Hardy-Weinberg equilibrium. Unknown/unmeasured confounding (U) was simulated using a normal distribution with a mean (0) and standard deviation (2) to produce a correlation between QT_1 and QT_2 (r = 0.62) similar to that observed in the ARIC study.

QT_1 was then simulated as a linear function of the SNP, age (centered), sex, U, and a standard deviation of 16 ms based on data from the ARIC baseline visit:

$$QT_1=\beta_0+\beta_1 SNP+\beta_2 Age+\beta_3 Sex+\beta_4 U+\varepsilon \quad \text{eq (1)}$$

where mean QT was simulated to be 418 ms when all other variables equaled 0. An additional copy of the minor allele, one-year increase in age, male sex, and 1 unit increments in U were associated with 5, 0.3, −8, and 2 ms changes in QT_1, respectively (Table 1). QT_{long} was then defined as QT_{long}=1 when $QT_1 \geq$ 450 ms and QT_1=0 otherwise, using a published threshold for QT abnormalities [32].

Initiation of drug treatment between the first and second visit was predicted conditional on age, sex and QT_{long}, using a logit function:

$$logit(pr(Drug=1))=\alpha_0+\alpha_1 Age+\alpha_2 Sex+\alpha_3 QT_{long} \quad \text{eq (2)}$$

where α_0 (−3.222) corresponded to a 5% prevalence of drug use (Table 1). Each one-year increase in age, male sex, and having a prolonged QT were each associated with changes of 0.041, 0.693, and −2.30 in the log odds of drug exposure, respectively. Drug treatment was then assigned using a binomial distribution with size =1 and this calculated probability.

Finally, we calculated QT_2 using a linear model as:

$$QT_2=\beta_0+\beta_1 Drug+\beta_2 SNP+\beta_3 Drug \times SNP+\beta_4 Age+\beta_5 Sex+\beta_6 U+\varepsilon \quad \text{eq (3)}$$

where β_0 was 418 ms when all other variables equaled 0. An additional copy of the minor allele, one-year increase in age, male sex, 1 unit increment in U, and treatment with the QT-prolonging drug were associated with 5, 0.3, −8, 2, 5 ms changes in QT_2, respectively (Table 1). The drug-SNP interaction was set to 3 ms. Changes in the basic scenario presented in equation 3 involved varying the SNP main effect from 0 to 10 ms and the drug-SNP interaction from −6 to 6 ms, and the MAF from 0.05–0.45 while the remaining parameters were fixed at their basic scenario values, unless otherwise specified. A total of 10,000 closed cohort studies of n=25,000 participants were then simulated for each baseline and alternative simulation scenario (Table 1). All analyses were performed using the statistical programming package SAS (Cary, N.C.).

Analyses of drug-SNP interactions: OPD versus TOD

We used our simulations to contrast two study designs, the OPD that included treated and untreated participants and the TOD, which was restricted to participants treated with a QT-prolonging drug at visit 2. For each design, we evaluated full (SNP, age, sex, and QT_{long} in the TOD; SNP, age, sex, QT_{long}, drug, and drug-SNP interaction in the OPD) and reduced (inclusive of all parameters in the OPD except QT_{long}) models, the later to evaluate the influence of failure to account for confounding by contraindication, e.g., a cross-sectional study without measures of the contraindication. This approach was then repeated substituting QT_1 for QT_{long}.

In the OPD, the drug-SNP interaction effect was estimated using linear regression with QT_2 as the dependent variable and the SNP, drug, drug-SNP interaction, QT_{long} (or QT_1), age centered at the mean, and sex as independent variables, as follows:

$$QT_2=\beta_0+\beta_1 Drug+\beta_2 SNP+\beta_3 Drug\times SNP+\beta_4 Age+\beta_5 Sex+\beta_6 QT_{long}. \quad \text{eq (4)}$$

For the TOD, the drug-SNP interaction effect (β_1) was estimated conditional on treatment using linear regression with QT_2 as the dependent variable, and the SNP, QT_{long}(or QT_1), age, and sex as independent variables:

$$QT_2=\beta_0+\beta_1 SNP+\beta_2 Age+\beta_3 Sex+\beta_4 QT_{long} \quad \text{eq (5)}$$

RESULTS

The simulated mean QT at visit 1 was 407 (26) ms in the OPD and 2 ms lower in the TOD (Table 2). The prevalence of prolonged QT was 4.6% in the OPD and by design only 0.76% in the TOD, reflecting confounding by contraindication. There also was a higher proportion of males in the TOD, consistent with findings that QT is generally longer in females than in males [33,34].

Simulations with drug-SNP interaction = 0 and varied SNP main effect

First, we evaluated the performance of the OPD and TOD in the absence of a simulated drug-SNP interaction (i.e. the drug-SNP interaction = 0), but in the presence of a SNP main effect (range: 0 to 10 ms), which represent the effect of the SNP independent of its interactive effect with the drug. Estimates of the drug-SNP interaction in the OPD were unconfounded across simulated SNP main effects when QT_{long} or QT_1 were included in the model (Figure 2, panels A and C). When QT_{long} or QT_1 were excluded from the OPD, moderate bias was observed, which increased with increasing SNP main effect (bias= −1.36 ms when SNP main effect =10 ms). Bias of the drug-SNP interaction was much more pronounced in the TOD and varied based on adjustment for QT_{long} or QT_1 (Figure 2, panels B and D). For example, substantial bias of the drug-SNP interaction (bias = 8.69 ms when SNP main effect = 10 ms) for the TOD was observed regardless of adjustment for QT_{long}. However, adjustment for QT_1 markedly decreased the potential for bias in the TOD (bias = 3.88 ms when SNP main effect = 10 ms). No variation in the magnitude of bias by MAF was seen for either the OPD or TOD.

Estimates of the FPP in the OPD remained 5% when adjusting for QT_{long} or QT_1, but increased in the absence of adjustment for either QT_{long} or QT_1 (FPP=0.19 when SNP main effect=10) (Figure 2, panels E and G). For the TOD, the estimated FPP was 5% in the absence of a SNP effect, but rose with increasing SNP main effect (FPP= 89% when SNP main effect = 4 ms and MAF = 0.25) (Figure 2, panel F), regardless of adjustment for

QT_{long}. Notably, the FPP estimated for SNPs with MAF = 0.45 approached 99% at SNP main effects of 4 ms. However, adjustment for QT_1 reduced the FPP markedly (FPP= 39% when SNP main effect = 4 ms and MAF = 0.25) (Figure 2, panel H).

Simulations with SNP main effect = 0 and varied drug-SNP interaction

Next, we evaluated the performance of the OPD and TOD in the presence of a simulated drug-SNP interaction (range: −6 to 6 ms) in the absence of a SNP main effect. For the OPD and TOD, results were unbiased across the range of drug-SNP effects regardless of adjustment for QT_{long} or QT_1 (Figure 3, panels A–D). The proportion of significant studies in the OPD and TOD models, estimated when the drug-SNP effect=0, was identical (5%) (Figure. 3, panel E–H)regardless of adjustment for QT_{long} or QT_1. Power to detect the interaction (estimated when the drug-SNP effect ≠ 0) also was comparable for the OPD and TOD, increased symmetrically as the simulated drug-SNP effect increased, and was higher for SNPs with higher simulated MAFs. Power also was slightly higher when adjusting for QT_1 (Figure 3, panels G and H).

Simulations with SNP main effect = 5 and varied drug-SNP interaction

Finally, a strong SNP main effect (SNP main effect = 5) and a varied drug-SNP effect (range: −6 to 6 ms) were simulated. For the OPD, the full model was slightly biased across the range of drug-SNP effects (bias = 0.14) (Figure 4, panels A and C) and became more pronounced in the absence of adjusting for QT_{long} or QT_1 (bias = −0.59). A bias of 4.43 ms across the range of drug-SNP effects was observed for the TOD (Figure 4, panels B and D), which did not vary with adjustment for QT_{long}. However, adjustment for QT_1 decreased the bias to 1.93 ms. The FPP was 5% in the OPD when the drug-SNP interaction was 0 and increased slightly to 6.7% in the absence of adjusting for QT_{long} or QT_1 (Figure 4, panels E and G).

Differences in statistical power also were observed for the OPD when including or excluding QT_{long} or QT_1. For example, there was a 25.8% reduction in power for a simulated drug-SNP effect = −2 ms and MAF = 0.25 in the absence of controlling for QT_{long} orQT_1, although differences in power were greatly attenuated when simulated drug-SNP interactions exceeded |4| ms. For the TOD, the FPP was 97.7% when the simulated drug-SNP interaction = 0 and MAF = 0.25 (Figure 4, panel F). The TOD FPP did not change with adjustment for QT_{long}, but was reduced to 55.3% when QT_1 was included in the model (Figure 4, panel H). Interestingly, the nominal 5% level was observed for the TOD when the simulated drug-SNP interaction was −4 ms in the absence of adjusting for QT_1 (Figure 4, panel F) and −2 ms otherwise (Figure 4, panel H). As expected, statistical power for both the OPD and TOD was greatest for SNPs with MAF = 0.45 and lowest for SNPs with MAF = 0.05.

DISCUSSION

In these simulations, we examined the influence of bias from confounding by contraindication in non-experimental pharmacogenomic studies. Our simulations identified several scenarios where confounding by contraindication affected estimated drug-SNP interactions, albeit to different degrees and conditional on study design. Adjustment for confounding by contraindication in the OPD and by variables affected by the SNP in the TOD, as well as the magnitude of the simulated SNP main effect, also influenced the degree of bias in these simulated pharmacogenomic studies.

Particularly striking was the observed bias in and proportion of false positive effects estimated from drug-SNP interactions in the TOD when a SNP main effect was simulated.

These observations reflect the inseparability of the SNP and drug-SNP interaction effects in models restricted to treated participants. Dependence between the SNP and drug-SNP effects was not observed in the OPD. While these results are a logic consequence of the corresponding designs, we believe that researchers might not fully appreciate the magnitude of bias in the TOD. Interestingly enough, the bias in the TOD resulting from a SNP main effect was reduced by conditioning on factors affected by the SNP, including QT_1. Specifically, in the TOD conditioning on QT_1 opens a backdoor path between the SNP and U, although in our simulation this appears to considerably reduce bias. However, this result may vary by the scenario under investigation. Also, controlling for QT_1 requires longitudinal data from two time-points and in such a setting researchers would tend to implement the OPD.

A scenario analogous to that seen in the TOD in the presence of a simulated SNP effect occurs when case-only approaches for the assessment of gene-environment interaction are used in case-control studies [35]. In these circumstances, even modest SNP-environment associations in the population can produce inflated type I errors. For both the TOD and the case-only design, the error is bounded by the degree of departure from independence. In genomic studies of QT, estimated SNP main effects generally range between 2–3 ms per copy of the minor allele [26,27], corresponding to an estimated bias and FPP of 1.79–2.68 ms and 36–67% when the true drug-SNP interaction is 0 and the TOD is selected. However, adjusting for variables affected by the SNP reduced the bias and FPP to 0.76–1.15 and 14–24%. Therefore, skepticism is warranted when interpreting results from pharmacogenomic studies of QT that use the TOD and report interactions for genes with established main effects, although adjustment for variables affected by the SNP may help reduce such skepticism.

Despite the potential for bias and type I error associated with the TOD, scenarios exist where its use may be preferred. Confounding by contraindication (or indication), which results from the assignment of different treatments to patients with different prognoses, is a major challenge in pharmacoepidemiology and potentially pharmacogenomics as well. In this simulation study, confounding by contraindication, and thus a biased drug-QT effect, generally had modest influence on estimates of the drug-SNP interaction. However, the magnitude of confounding by contraindication is a function of the strength of the confounder effect on treatment and outcome (independent of treatment) as well as its prevalence in the population [36]. In this scenario, we assumed a strong effect on drug prescribing, but the prevalence of QT_{long} was low, limiting its influence. Although the potential for substantial bias and an increase in the proportion of false positive associations also were observed with strong (i.e. > 5 ms) SNP effects in the presence of uncontrolled confounding by contraindication for the OPD, prior main effect and pharmacogenomic GWAS have not identified associations with QT of this magnitude [26,36].

However, bounds placed on the expected magnitude of SNP and interactive effects may not hold for future studies of rare variants, which may harbor much stronger effects. Also concerning are the potential for bias and elevated false positive proportions resulting from SNPs in the “gray zone”, i.e. those with relatively modest effects that are difficult to identify by conventional methods. Researchers therefore must balance the tradeoffs of the OPD and TOD on a study-by-study basis. Indeed, several methods have been developed that leverage small SNP main effects to boost statistical power for detecting gene-environment effects in large-scale genomics studies [37].

Although our simulations evaluated gene-environment interaction for a continuous outcome, the associated challenges apply to assessments of interaction between any pair of risk factors and a (e.g., binary) trait. Interaction exists if there is a different effect of a treatment on risk

of disease in individuals with different genotypes [38] and may be synergistic, antagonistic or null for additive and multiplicative scales [39]. In addition to the challenges described above, a further drawback of the TOD is that if the outcome is binary, only multiplicative interaction can be estimated, although a number of epidemiologists have advocated additive interaction as the relevant scale for the assessment of biological interaction [40–42].

The lack of an ideal study design in non-experimental pharmacogenomic research may prompt suggestions to preferentially choose clinical trials, where confounding by contraindication is avoided by design. However, non-experimental studies have advantages over clinical trials that exclude potential participants experiencing early manifestations of drug intolerance before they are randomized (e.g. during a run-in phase) and/or those with relevant co-morbidities. Such exclusions often yield selected populations that are less prone to ADRs than community-based populations [43]. Clinical trials also may not be large enough to detect pharmacogenomic effects given the requisite scale of even modestly well-powered genetic epidemiologic studies. For example, a recently published pharmacogenomic GWAS of QT conducted in non-experimental settings reported that longitudinal studies exceeding 30,000 participants were required to detect pharmacogenomic effects [28]. Clearly, very few clinical trials of the same drug or drug class have been performed on such large numbers of participants and have the necessary genetic data.

Several limitations of the present study warrant further consideration as future efforts to evaluate the influence of study design in pharmacogenomic studies. The simulations presented in this manuscript are of course limited by the scope of the scenarios and the selection of the parameter space. As such, conclusions may be made for the specified scenarios, but generalization to scenarios with different confounding patterns should be done with caution. We also limited our evaluation to two study designs: the TOD and OPD. Other designs that deserve consideration are those using active comparators, e.g. participants initiating other hypertensive agents could serve as the reference group in a pharmacogenomic study of thiazide diuretics. Outside of the current scope also was the consideration of selection bias and treatment duration, although both warrant future investigation.

CONCLUSION

Our simulations suggest that there is no ideal design for pharmacogenomic studies conducted in non-experimental settings. Although effects of confounding by contraindication were modest, thus favoring the OPD, the TOD may be preferred when strong confounding by contraindication is suspected or prior data support the absence of a SNP main effect. Ultimately, pharmacogenomic researchers using data collected in non-experimental settings must carefully weigh the counterbalancing influences of all error sources before selecting the design that best suits the given context.

Acknowledgments

This effort was partially supported by an award from the National Heart, Lung, and Blood Institute (CLA, EAW, TS, R01-HL-103612). CLA was supported in part by grant R00-HL-098458 from the National Heart, Lung, and Blood Institute. JSD acknowledges support by the Gary G. and Carolyn J. Koch Merit Scholarship. TS was supported in part by grant R01 AG023178 from the National Institute of Aging.

References

1. Lazarou J, Pomeranz BH, Corey PN. Incidence of adverse drug reactions in hospitalized patients: a meta-analysis of prospective studies. JAMA. 1998; 279:1200–5. [PubMed: 9555760]

2. Aronson JK, Ferner RE. Joining the DoTS: new approach to classifying adverse drug reactions. BMJ. 2003; 327:1222–5. [PubMed: 14630763]
3. Khoury MJ, Gwinn M, Clyne M, Yu W. Genetic epidemiology with a capital E, ten years after. Genet Epidemiol. 2011; 35(8):845–52. [PubMed: 22125223]
4. Daneshjou R, Tatonetti NP, Karczewski KJ, Sagreiya H, Bourgeois S, Drozda K, Burmester JK, Tsunoda T, Nakamura Y, Kubo M, Tector M, Limdi NA, Cavallari LH, Perera M, Johnson JA, Klein TE, Altman RB. Pathway analysis of genome-wide data improves warfarin dose prediction. BMC Genomics. 2013; 14 (Suppl 3):S11. [PubMed: 23819817]
5. Niinuma Y, Saito T, Takahashi M, Tsukada C, Ito M, Hirasawa N, Hiratsuka M. Functional characterization of 32 CYP2C9 allelic variants. Pharmacogenomics J. 2013 In press.
6. Perera MA, Cavallari LH, Limdi NA, Gamazon ER, Konkashbaev A, Daneshjou R, Pluzhnikov A, Crawford DC, Wang J, Liu N, Tatonetti N, Bourgeois S, Takahashi H, Bradford Y, Burkley BM, Desnick RJ, Halperin JL, Khalifa SI, Langaee TY, Lubitz SA, Nutescu EA, Oetjens M, Shahin MH, Patel SR, Sagreiya H, Tector M, Weck KE, Rieder MJ, Scott SA, Wu AH, Burmester JK, Wadelius M, Deloukas P, Wagner MJ, Mushiroda T, Kubo M, Roden DM, Cox NJ, Altman RB, Klein TE, Nakamura Y, Johnson JA. Genetic variants associated with warfarin dose in African-American individuals: a genome-wide association study. Lancet. 2013; 382:790–6. [PubMed: 23755828]
7. Fellay J, Thompson AJ, Ge D, Gumbs CE, Urban TJ, Shianna KV, Little LD, Qiu P, Bertelsen AH, Watson M, Warner A, Muir AJ, Brass C, Albrecht J, Sulkowski M, McHutchison JG, Goldstein DB. ITPA gene variants protect against anaemia in patients treated for chronic hepatitis C. Nature. 2010; 464:405–8. [PubMed: 20173735]
8. Paik S, Shak S, Tang G, Kim C, Baker J, Cronin M, Baehner FL, Walker MG, Watson D, Park T, Hiller W, Fisher ER, Wickerham DL, Bryant J, Wolmark N. A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. N Engl J Med. 2004; 351:2817–26. [PubMed: 15591335]
9. Link E, Parish S, Armitage J, Bowman L, Heath S, Matsuda F, Gut I, Lathrop M, Collins R. SLCO1B1 variants and statin-induced myopathy--a genomewide study. N Engl J Med. 2008; 359:789–99. [PubMed: 18650507]
10. Volpi S, Heaton C, Mack K, Hamilton JB, Lannan R, Wolfgang CD, Licamele L, Polymeropoulos MH, Lavedan C. Whole genome association study identifies polymorphisms associated with QT prolongation during iloperidone treatment of schizophrenia. Mol Psychiatry. 2009; 14:1024–31. [PubMed: 18521091]
11. Aberg K, Adkins DE, Liu Y, McClay JL, Bukszar J, Jia P, Zhao Z, Perkins D, Stroup TS, Lieberman JA, Sullivan PF, van den Oord EJ. Genome-wide association study of antipsychotic-induced QTc interval prolongation. Pharmacogenomics J. 2012; 12:165–72. [PubMed: 20921969]
12. van Winkel R, Moons T, Peerbooms O, Rutten B, Peuskens J, Claes S, van Os J, De Hert M. MTHFR genotype and differential evolution of metabolic parameters after initiation of a second generation antipsychotic: an observational study. Int Clin Psychopharmacol. 2010; 25:270–6. [PubMed: 20523222]
13. Kaab S, Crawford DC, Sinner MF, Behr ER, Kannankeril PJ, Wilde AA, Bezzina CR, Schulze-Bahr E, Guicheney P, Bishopric NH, Myerburg RJ, Schott JJ, Pfeufer A, Beckmann BM, Martens E, Zhang T, Stallmeyer B, Zumhagen S, Denjoy I, Bardai A, Van Gelder IC, Jamshidi Y, Dalageorgou C, Marshall V, Jeffery S, Shakir S, Camm AJ, Steinbeck G, Perz S, Lichtner P, Meitinger T, Peters A, Wichmann HE, Ingram C, Bradford Y, Carter S, Norris K, Ritchie MD, George AL Jr, Roden DM. A large candidate gene survey identifies the KCNE1 D85N polymorphism as a possible modulator of drug-induced torsades de pointes. Circ Cardiovasc Genet. 2012; 5:91–9. [PubMed: 22100668]
14. de Keyser CE, Eijgelsheim M, Hofman A, Sijbrands EJ, Maitland-van der Zee AH, van Duijn CM, Uitterlinden AG, Witteman JC, Ch Stricker BH. Single nucleotide polymorphisms in genes that are associated with a modified response to statin therapy: the Rotterdam Study. Pharmacogenomics J. 2011; 11:72–80. [PubMed: 20195290]
15. Becker ML, Visser LE, van Schaik RH, Hofman A, Uitterlinden AG, Stricker BH. Genetic variation in the organic cation transporter 1 is associated with metformin response in patients with diabetes mellitus. Pharmacogenomics J. 2009; 9:242–7. [PubMed: 19381165]

16. Hajjar I, Kritchevsky S, Newman AB, Li R, Yaffe K, Simonsick EM, Lipsitz LA. Renin angiotensin system gene polymorphisms modify angiotensin-converting enzyme inhibitors' effect on cognitive function: the health, aging and body composition study. J Am Geriatr Soc. 2010; 58:1035–42. [PubMed: 20722844]

17. Lee CR, North KE, Bray MS, Couper DJ, Heiss G, Zeldin DC. Cyclooxygenase polymorphisms and risk of cardiovascular events: the Atherosclerosis Risk in Communities (ARIC) study. Clin Pharmacol Ther. 2008; 83:52–60. [PubMed: 17495879]

18. ARIC Investigators. The Atherosclerosis Risk in Communities (ARIC) Study: design and objectives. The ARIC investigators. Am J Epidemiol. 1989; 129:687–702. [PubMed: 2646917]

19. Havlik RJ, Garrison RJ, Fabsitz R, Feinleib M. Variability of heart rate, P–R, QRS and Q–T durations in twins. J Electrocardiol. 1980; 13:45–8. [PubMed: 7188949]

20. Newton-Cheh C, Larson MG, Corey DC, Benjamin EJ, Herbert AG, Levy D, D'Agostino RB, O'Donnell CJ. QT interval is a heritable quantitative trait with evidence of linkage to chromosome 3 in a genome-wide linkage analysis: The Framingham Heart Study. Heart Rhythm. 2005; 2:277–84. [PubMed: 15851319]

21. Vaidean GD, Schroeder EB, Whitsel EA, Prineas RJ, Chambless LE, Perhac JS, Heiss G, Rautaharju PM. Short-term repeatability of electrocardiographic spatial T-wave axis and QT interval. J Electrocardiol. 2005; 38:139–47. [PubMed: 15892024]

22. Pratt CM, Ruberg S, Morganroth J, McNutt B, Woodward J, Harris S, Ruskin J, Moye L. Dose-response relation between terfenadine (Seldane) and the QTc interval on the scalar electrocardiogram: distinguishing a drug effect from spontaneous variability. Am Heart J. 1996; 131:472–80. [PubMed: 8604626]

23. McAllister RG Jr, Kirsten EB. The pharmacology of verapamil. IV. Kinetic and dynamic effects after single intravenous and oral doses. Clin Pharmacol Ther. 1982; 31:418–26. [PubMed: 7060323]

24. Yap YG, Camm AJ. Drug induced QT prolongation and torsades de pointes. Heart. 2003; 89:1363–72. [PubMed: 14594906]

25. Zhang Y, Post WS, Blasco-Colmenares E, Dalal D, Tomaselli GF, Guallar E. Electrocardiographic QT interval and mortality: a meta-analysis. Epidemiology. 2011; 22:660–70. [PubMed: 21709561]

26. Newton-Cheh C, Eijgelsheim M, Rice KM, de Bakker PI, Yin X, Estrada K, Bis JC, Marciante K, Rivadeneira F, Noseworthy PA, Sotoodehnia N, Smith NL, Rotter JI, Kors JA, Witteman JC, Hofman A, Heckbert SR, O'Donnell CJ, Uitterlinden AG, Psaty BM, Lumley T, Larson MG, Stricker BH. Common variants at ten loci influence QT interval duration in the QTGEN Study. Nat Genet. 2009; 41:399–406. [PubMed: 19305408]

27. Pfeufer A, Sanna S, Arking DE, Muller M, Gateva V, Fuchsberger C, Ehret GB, Orru M, Pattaro C, Kottgen A, Perz S, Usala G, Barbalic M, Li M, Putz B, Scuteri A, Prineas RJ, Sinner MF, Gieger C, Najjar SS, Kao WH, Muhleisen TW, Dei M, Happle C, Mohlenkamp S, Crisponi L, Erbel R, Jockel KH, Naitza S, Steinbeck G, Marroni F, Hicks AA, Lakatta E, Muller-Myhsok B, Pramstaller PP, Wichmann HE, Schlessinger D, Boerwinkle E, Meitinger T, Uda M, Coresh J, Kaab S, Abecasis GR, Chakravarti A. Common variants at ten loci modulate the QT interval duration in the QTSCD Study. Nat Genet. 2009; 41:407–14. [PubMed: 19305409]

28. Avery CL, Sitlani CM, Arking DE, Arnett DK, Bis JC, Boerwinkle E, Chen Y-DI, de Craen AJM, Eijgelsheim M, Enquohabarie D, Evans DS, Ford I, Garcia ME, Gudnason V, Harris TB, Heckbert SR, Hochner H, Hofman A, Hsueh W-C, Isaacs A, Jukema JW, Knekt P, Kors JA, Krijthe BP, Kristiansson K, Laaksonen M, Liu Y, Li X, MacFarlane PW, Newton-Cheh C, Nieminen MS, Oostra BA, Peloso GM, Porthan K, Rice K, Rivadeneira FF, Rotter JI, Salomaa V, Sattar N, Siscovick DS, Slagboom PE, Smith AV, Sotoodehnia N, Stott DJ, Stricker BH, Stürmer T, Trompet S, Uitterlinden AG, van Duijn CM, Westendorp RGJ, Witteman JC, Whitsel EA, Psaty BM. Drug-gene interactions and the search for missing heritability: a cross-sectional pharmacogenomics study of the QT interval. The Pharmacogenomics Journal. 2013 In press.

29. Benoit SR, Mendelsohn AB, Nourjah P, Staffa JA, Graham DJ. Risk factors for prolonged QTc among US adults: Third National Health and Nutrition Examination Survey. Eur J Cardiovasc Prev Rehabil. 2005; 12:363–8. [PubMed: 16079644]

30. Geneletti S, Gallo V, Porta M, Khoury MJ, Vineis P. Assessing causal relationships in genomics: From Bradford-Hill criteria to complex gene-environment interactions and directed acyclic graphs. Emerg Themes Epidemiol. 2011; 8:5. [PubMed: 21658235]

31. Greenland S, Pearl J, Robins JM. Causal diagrams for epidemiologic research. Epidemiology. 1999; 10:37–48. [PubMed: 9888278]

32. Robbins J, Nelson JC, Rautaharju PM, Gottdiener JS. The association between the length of the QT interval and mortality in the Cardiovascular Health Study. Am J Med. 2003; 115:689–94. [PubMed: 14693320]

33. Merri M, Benhorin J, Alberti M, Locati E, Moss AJ. Electrocardiographic quantitation of ventricular repolarization. Circulation. 1989; 80:1301–8. [PubMed: 2805266]

34. Lepeschkin E, Surawicz B. The duration of the Q–U interval and its components in electrocardiograms of normal persons. Am Heart J. 1953; 46:9–20. [PubMed: 13057814]

35. Albert PS, Ratnasinghe D, Tangrea J, Wacholder S. Limitations of the case-only design for identifying gene-environment interactions. Am J Epidemiol. 2001; 154:687–93. [PubMed: 11590080]

36. Bross ID. Spurious effects from an extraneous variable. J Chronic Dis. 1966; 19:637–47. [PubMed: 5966011]

37. Manning AK, LaValley M, Liu CT, Rice K, An P, Liu Y, Miljkovic I, Rasmussen-Torvik L, Harris TB, Province MA, Borecki IB, Florez JC, Meigs JB, Cupples LA, Dupuis J. Meta-analysis of gene-environment interaction: joint estimation of SNP and SNP x environment regression coefficients. Genet Epidemiol. 2011; 35:11–8. [PubMed: 21181894]

38. Ottman R. Gene-environment interaction: definitions and study designs. Prev Med. 1996; 25:764–70. [PubMed: 8936580]

39. Brennan P. Gene-environment interaction and aetiology of cancer: what does it mean and how can we measure it? Carcinogenesis. 2002; 23:381–7. [PubMed: 11895852]

40. Rothman, KJ. Epidemiology: an introduction. Oxford: University press; 2002. Measuring interactions.

41. Andersson T, Alfredsson L, Kallberg H, Zdravkovic S, Ahlbom A. Calculating measures of biological interaction. Eur J Epidemiol. 2005; 20:575–9. [PubMed: 16119429]

42. Rothman KJ. The estimation of synergy or antagonism. Am J Epidemiol. 1976; 103:506–11. [PubMed: 1274952]

43. Wilke RA, Xu H, Denny JC, Roden DM, Krauss RM, McCarty CA, Davis RL, Skaar T, Lamba J, Savova G. The emerging role of electronic medical records in pharmacogenomics. Clin Pharmacol Ther. 2011; 89:379–86. [PubMed: 21248726]

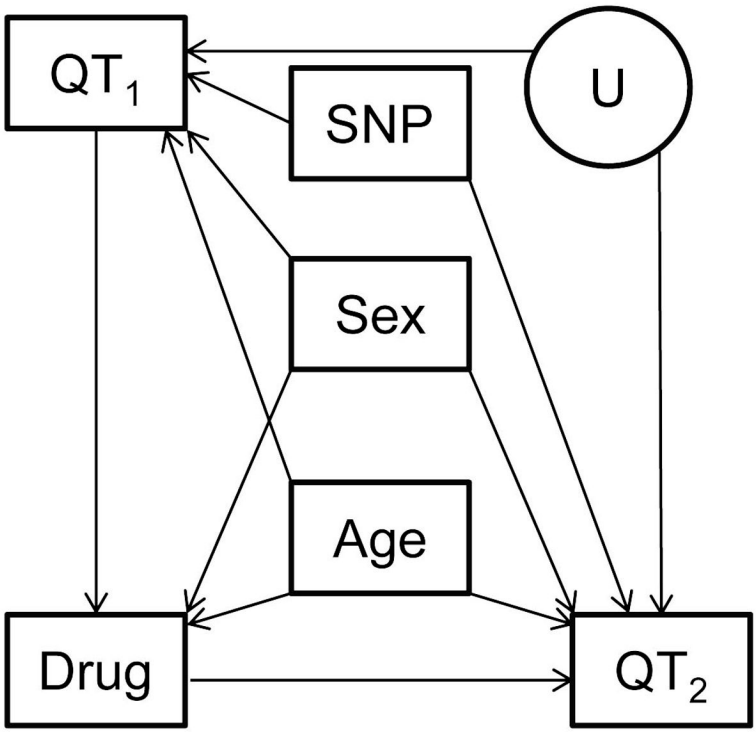

Figure 1.
Directed acyclic graph (DAG) of the relationship between QT-prolonging drug use and QT assessed at visit 2 (QT_2, ms). Single nucleotide polymorphism (SNP) dosage was added as an effect measure modifier. The confounders sex and age, independent of the SNP, were included to examine the degree to which confounders of the drug-QT_2 association bias the drug-SNP interaction. U represents unknown or unmeasured factors influencing the correlation between QT_1 and QT_2, including other SNPs, drugs, and clinical covariates. QT at visit 1 (QT_1), was added to model confounding by contraindication, knowing that participants with a prolonged QT are less likely to be prescribed QT-prolonging medications.

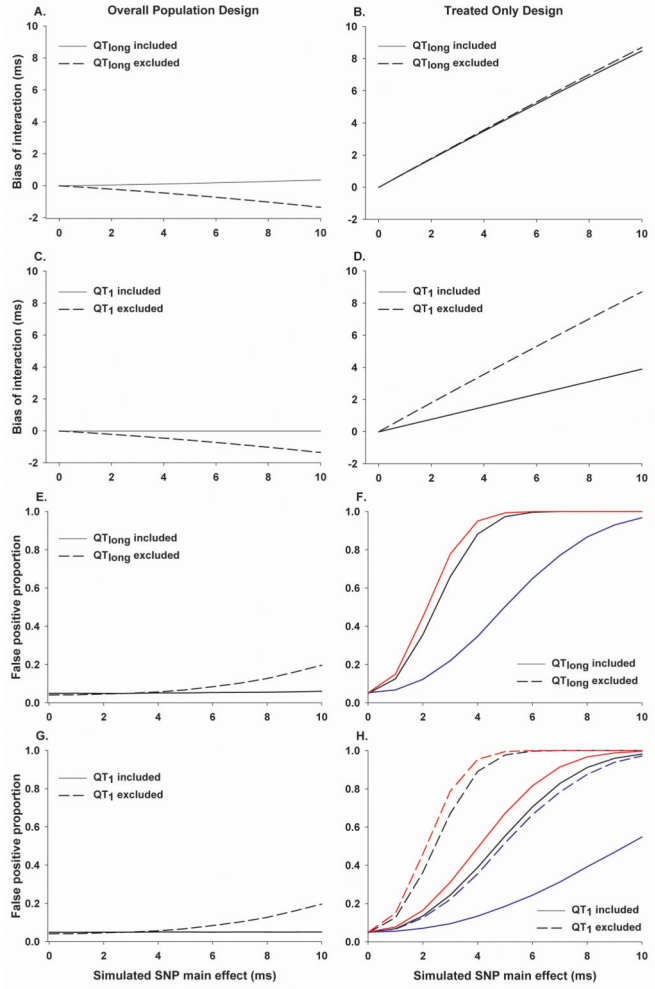

Figure 2.

The bias in (panels A–D) and proportion of false positive associations estimated for (panels E–H) a pharmacogenomic study of QT in among populations unexposed to the drug (i.e. drug-SNP interaction = 0) over a varied SNP main effect (SNP main effect range: 0–10 milliseconds). Left column panes represent the overall population design (OPD) and right column panels denote the treated only design (TOD). The influences of adjusting for confounding by contraindication (QT_{long}) and covariates affected by the SNP (QT_1) are gauged by contrasting models with (—) and without (- - -) adjustment for QT_{long} or QT_1. Blue, black, and red shading represent SNPs with minor allele frequencies (MAF) of 0.05, 0.25, and 0.45, respectively. Simulations were performed assuming a population size of 25,000 participants with approximately 5% of the participants receiving QT-prolonging treatment. A total of 10,000 iterations were performed per scenario.

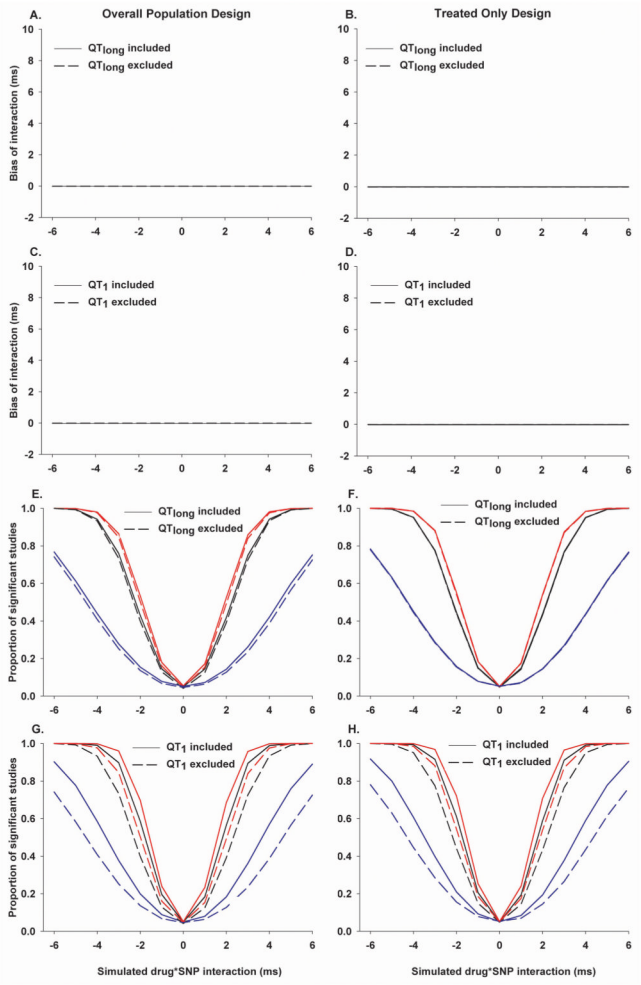

Figure 3.

The bias in (panels A–D) and proportion of significant associations estimated for (panels E–H) a pharmacogenomic study of QT in the absence of a simulated SNP main effect (i.e. SNP = 0), but a varied drug-SNP interaction effect (drug-SNP interaction effect range: −6 – 6 milliseconds). Left column panes represent the overall population design (OPD) and right column panels denote the treated only design (TOD). The influences of adjusting for confounding by contraindication (QT_{long}) and covariates affected by the SNP (QT_1) are gauged by contrasting models with (—) and without (- - -) adjustment for QT_{long} or QT_1. Blue, black, and red shading represent SNPs with minor allele frequencies (MAF) of 0.05, 0.25, and 0.45, respectively. Simulations were performed assuming a population size of 25,000 participants with approximately 5% of the participants receiving QT-prolonging treatment. A total of 10,000 iterations were performed per scenario.

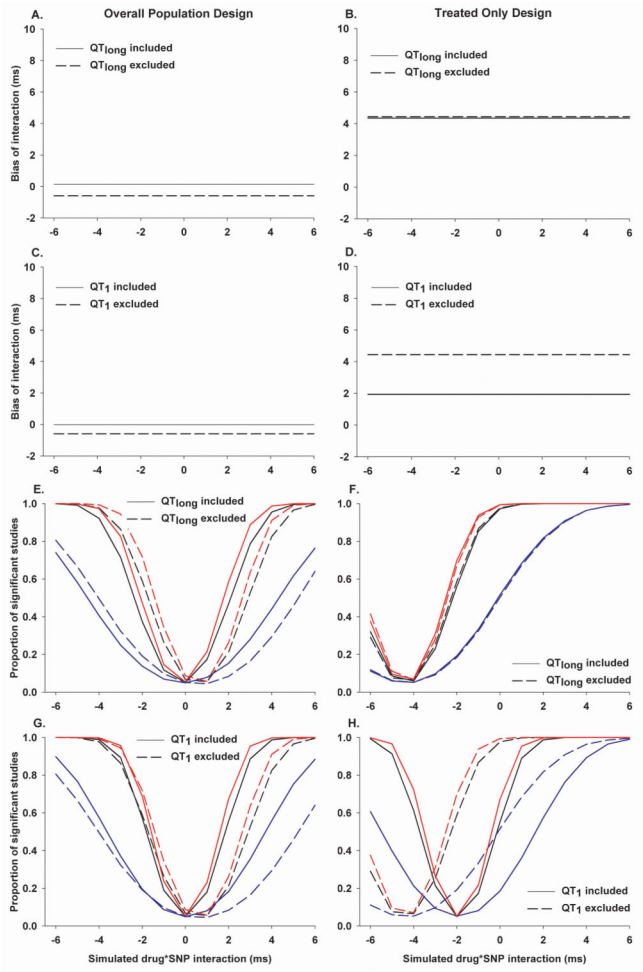

Figure 4.

The bias in (panels A–D) and proportion of significant associations (panels E–H) estimated for a pharmacogenomic study of QT in the presence of a simulated SNP main effect (i.e. SNP = 5) and a varied drug-SNP interaction effect (drug-SNP interaction effect range: −6–6 milliseconds). Left column panes represent the overall population design (OPD) and right column panels denote the treated only design (TOD). The influences of adjusting for confounding by contraindication (QT_{long}) and covariates affected by the SNP (QT_1) are gauged by contrasting models with (—) and without (- - -) adjustment for QT_{long} or QT_1. Blue, black, and red shading represent SNPs with minor allele frequencies (MAF) of 0.05, 0.25, and 0.45, respectively. Simulations were performed assuming a population size of 25,000 participants with approximately 5% of the participants receiving QT-prolonging treatment. A total of 10,000 iterations were performed per scenario.

Table 1

Parameter values for the basic and alternative simulation scenarios.

			Parameter values (ms)	
Parameter Notation	**Equation**	**Meaning**	**Basic scenario**	**Alternative scenarios**
β_0	1,3	Mean QT at Visit 2 for female of mean age with 0 copies of the minor allele and the mean amount of unmeasured confounding	410	
β_{SNP}	1,3	SNP effect	5	0 to 10
β_{Age}	1,3	Age effect	0.3	
β_{Sex}	1,3	Sex effect	−8	
β_U	1,3	Effect of unknown/unmeasured correlates of QT_1 and QT_2	2	
β_{Drug}	3	Drug effect	5	
$\beta_{Drug*SNP}$	3	Drug-SNP interaction effect	3	−6 to 6
α_0	2	Value chosen for 5% prevalence of drug exposure	−3.222	
α_{Age}	2	Log odds of drug exposure for one-year increase in age	0.041	
α_{Sex}	2	Log odds of drug exposure for male versus female	0.693	
α_{QTlong}	2	Log odds of drug exposure for prolonged versus normal QT	−2.303	

Table 2

Descriptive statistics in simulation studies using overall population design versus treated-only design.

	Overall (n=25,000)		Treated only (n=1,313)	
	Mean (SD) or %	**Range**	**Mean (SD) or %**	**Range**
QT_1	407 (26)	296, 506	405 (23)	321, 469
QT_{long}	4.6%		0.76%	
QT_2	407 (26)	300, 504	410 (25)	328, 481
Drug	5.3%		100%	
SNP MAF*	25%		25%	
Age	54 (6)	29, 76	55 (6)	34, 75
Sex	44.9% male		62% male	

*When a MAF of 25% was simulated. MAF, minor allele frequency; QT_1, duration of QT measured at visit 1; QT_2, duration of QT measured at visit 2; QT_{long}, indicator variable (QT_{long}=1 when QT_1>450 ms, = 0 otherwise); SD, standard deviation