Received: 30 December 2020 | Revised: 23 March 2021 | Accepted: 25 March 2021

DOI: 10.1002/psp4.12624

ARTICLE

Impact of study design and statistical model in pharmacogenetic studies with gene-treatment interaction

Camille Couffignal[1,2] | **France Mentré**[1,2] | **Julie Bertrand**[1]

[1]INSERM, IAME, Université de Paris, Paris, France

[2]Clinical Research, Biostatistics and Epidemiology Department, AP-HP, Hôpital Bichat, Paris, France

Correspondence
Camille Couffignal, Université de Paris, INSERM, IAME, F-75006 Paris, France.
Email: camille.couffignal@aphp.fr

Funding information
No funding was received for this work.

Abstract

Gene-treatment interactions, just like drug-drug interactions, can have dramatic effects on a patient response and therefore influence the clinician decision at the patient's bedside. Crossover designs, although they are known to decrease the number of subjects in drug-interaction studies, are seldom used in pharmacogenetic studies. We propose to evaluate, via realistic clinical trial simulations, to what extent crossover designs can help quantifying the gene-treatment interaction effect. We explored different scenarios of crossover and parallel design studies comparing two symptom-modifying treatments in a chronic and stable disease accounting for the impact of a one gene and one gene-treatment interaction. We varied the number of subjects, the between and within subject variabilities, the gene polymorphism frequency and the effect sizes of the treatment, gene, and gene-treatment interaction. Each simulated dataset was analyzed using three models: (i) estimating only the treatment effect, (ii) estimating the treatment and the gene effects, and (iii) estimating the treatment, the gene, and the gene-treatment interaction effects. We showed how ignoring the gene-treatment interaction results in the wrong treatment effect estimates. We also highlighted how crossover studies are more powerful to detect a treatment effect in the presence of a gene-treatment interaction and more often lead to correct treatment attribution.

Study Highlights

WHAT IS THE CURRENT KNOWLEDGE ON THE TOPIC?

When pharmacogenetic effects are suspected for drugs of the same therapeutic area, they should be explored in order to choose the best treatment and dose for each patient to avoid rejecting a new drug.

WHAT QUESTION DID THIS STUDY ADDRESS?

Investigating if pharmacogenetic effects differ between treatments is important to attribute the best therapeutic option (treatment or regimens) in each genetic subgroup.

WHAT DOES THIS STUDY ADD TO OUR KNOWLEDGE?

To capture adequately the gene-treatment interaction, a crossover design is more powerful than a parallel design. Ignoring an existing gene-treatment interaction results in incorrect treatment effect estimates.

HOW MIGHT THIS CHANGE DRUG DISCOVERY, DEVELOPMENT, AND/OR THERAPEUTICS?

Assuming a gene-treatment interaction and using a crossover design seems the best strategy for the pharmacogenetic study of concurrent drug treatments.

INTRODUCTION

The development of personalized medicine should lead to improved safety and efficacy of drug use,[1,2] even more for drugs with a narrow therapeutic margin and a high interindividual variability.[3] It is now well-established that pharmacokinetic and pharmacodynamic studies enable to quantify the interindividual variability in treatment response and its genetic component when it exists.[4,5]

Indeed, genetic variability has been described in the metabolism and effect of drugs, and gene modulators of the response to drug treatment have been identified.[6] Precisely, pharmacogenetic studies investigate the influence of genetic polymorphism on drug response,[7–9] thereby providing a tool for treatment personalization.[10,11]

Currently, during the development of a new drug, associations with specific polymorphisms are routinely explored, nonetheless the potential influence of the metabolizer status for certain enzymes (e.g., CYP450 cytochrome) is hardly reported. Attia et al. highlighted four conceptual objectives (i) to identify a polymorphism with a key role on the drug efficacy, (ii) to avoid rejecting wrongly a drug candidate because of an unidentified gene effect, (iii) to increase the consistence of result across populations, and (iv) to help the clinician choice over a drug panel.[12] The development of a new drug, historically confined to the "one-size-fits-all" approach, must now tend toward personalized medicine to increase its chances at providing a superior efficacy, a more convenient dosing regimen or route of administration, or a lower risk of adverse effects.[5] Many studies report associations between genotype and efficacy in patients treated with a given drug, ignoring information from untreated patients.[13] The benefits of incorporating pharmacogenetic into clinical practice is now well-established, however, high-quality findings are lacking due to unresolved methodological and statistical issues of pharmacogenetic studies.[1,6,14,15] Briefly, as well illustrated by Holmes et al.[8] in a systematic review on the methodological pitfall of pharmacogenetic studies, the lack of consistency of these studies may be a result of the small sample sizes, use of candidate gene approaches, with paucity of reproducibility and paucity of meta-analyses. Therefore, there is still an urgent need for individualized treatments.[16] Better-designed and analyzed pharmacogenetic studies could provide alternative medications and better response through a change of regimens. An important aspect of the design of pharmacogenetic studies, as of any clinical trial, is to have sufficient power to detect a clinically significant difference between genotypes.[1,7] Evaluation of the power to detect gene-treatment interaction is complicated because it depends not only on the treatment effect size within each genotype, but also on the number of genotypes, their size, and the gene effect size.[7]

The crossover design is an alternative to the parallel design regularly used for drug-drug interaction evaluation, consisting in randomly allocating patients to treatment sequences so as to capture within individual differences between treatments especially relevant for chronic and stable disease with rapid and reversible treatment effect. A standard crossover design is the two-treatment, two-period crossover, in which each patient receives both treatments but is randomly allocated to one of two sequences whereas in the parallel design patients are randomly allocated one of the two treatments.[17,18] Its nature permits each patient to act as their own control, exploiting the fact that in most instances the variability between measurements from different subjects in a study will be far greater than that from the same subject on different occasions. Therefore, crossover trials are often more powerful than parallel group trials. However, it is seldom used in pharmacogenetic studies,[18,19] which preferably use parallel designs where every patient is given only one drug, the reference, or the test.

The aim of the present work is to show the impact of the study design and statistical model on the power of a pharmacogenetic study evaluating two treatments (candidate and reference) when a genetic polymorphism does or does not increase the benefit of the candidate treatment. Although the strength of a crossover study is recognized by some in the statistical community, we used simulations to help demonstrate its advantages to detect a gene-treatment interaction and to study the impact of the choice of the statistical model.

METHODS

Statistical model

Let us consider a reference ($T = 0$) and a test treatment ($T = 1$) and one genetic polymorphism with $G = 1$ for the genotype "rare allele homozygotes" and $G = 0$ otherwise (i.e., two alleles of the rare variant are required for the polymorphism to have an effect; this corresponds to a recessive genetic model). G affects the outcome of interest directly and through an interaction with the test treatment.

Let Y_{ij} be a continuous outcome of interest for patient $i = 1, \ldots, N$ receiving treatment T_{ij} at occasion $j = 1$ or 2 within a standard two-treatment two-sequences two-periods crossover design (D_{xo}) as follows:

$$Y_{ij} = \mu + \beta_T T_{ij} + \beta_G G_i + \beta_I I_{ij} + b_i + k_{ij}, \quad (1)$$

where μ is the intercept, β_T the treatment effect, β_G the gene effect, and I the gene-treatment interaction (i.e., $I_{ij} = T_{ij} \times G_i$) such that $I = 1$ when the test treatment is given to a patient genotype $G = 1$ with β_I the gene-treatment interaction effect. Here, the random effects b_i and k_{ij} capture the between and within subjects' variability and follow normal distributions of mean 0 and variances ω^2 and γ^2, respectively. In addition, we define R_G, the gene component coefficient based on the ratio of the within and between subject variances, such as $R_G = 1 - \gamma^2/\omega^2$ [20] that varies from 0 (weak gene component to the variability) to 1 (strong gene component to the variability).

For a parallel design (D_p), (1) simplifies in:

$$Y_i = \mu + \beta_T T_i + \beta_G G_i + \beta_I I_i + b_i, \quad (2)$$

where the random effect b_i follows a normal of mean 0 and variance σ^2 $(=\omega^2 + \gamma^2)$, and every patient receive only one treatment (i.e., N/2 patients receive the reference treatment and N/2 the test).

Simulation study

We choose to set the simulation study in the context of a chronic disease with a rapid and reversible treatment effect (symptoms modifying drug) to enable the assumption of no carry-over, sequence or period effects in the crossover study. The simulated values for the intercept $\mu = 8$ and the residual error standard deviation $\sigma = 14$ were based on the Ideal trial study.[21] In Table 1 for D_{xo} and D_p, we display the set of all simulated values for N (N for $D_{xo} = 2 \times N$ for D_p), F (the frequency of $G = 1$), β_T, β_G, β_I, ω^2, γ^2, and the corresponding R_G and σ^2. For the treatment, the gene and the interaction effect size we considered a 50% change in the continuous outcome (i.e., a medium magnitude according to ref. 22) For the gene and the interaction effect size, we in addition considered a 100% change for illustration purposes. To explore the impact of the number of subjects, the minimum and maximum values are rather typical for crossover (from 50 to 200) and parallel (from 100 to 400) designs. For instance, the IDEAL study, which was a crossover study included $N = 112$ patients. Then, to explore the impact of the percentage of mutant homozygotes, we considered 20% (typical of CYP2A6 or CYP2B6 variant homozygotes in White patients) and 4% (typical of CYP2C8 or CYP2C9 homozygotes variants in White patients).[23] For D_{xo}, all the possible combinations of

TABLE 1 Set of values used for the various scenarios for each parameter for the crossover D_{XO} and parallel D_P designs

Design	Crossover D_{XO}		Parallel D_P
Number of subjects	50		100
	100		**200**
	200		400
Allelic frequency	0.04		0.04
	0.20		**0.20**
Gene component coefficient and standard deviation	R_G	$\omega;\gamma$	σ
	0.5	11.43;8.08	**14**
	0.7	**12.28;6.73**	
	0.9	13.35;4.22	
Treatment effect β_T	0		0
	−5		−5
	5		5
Gene effect β_G	0		0
	−5		−5
	−10		−10
Gene-treatment interaction effect β_I	0		0
	−5		−5
	−10		−10

Note: In bold are the values used to illustrate the main results.

Abbreviations: β_T, treatment effect; β_G, gene effect and β_I, gene-treatment interaction effect. F, frequency of mutant homozygotes; N, number of subjects; ω and γ, standard deviation of the between and within subjects' variability, respectively; R_G, gene component coefficient; σ, standard deviation for D_P.

N (3 values) × F (2 values) × R_G (3 values) × β_T (3 values) × β_G (3 values) × β_I (3 values) were simulated = 486 scenarios. Similarly, for D_p, all the possible combinations of N (3 values) × F (2 values) × β_T (3 values) × β_G (3 values) × β_I (3 values) were simulated = 162 simulation scenarios. Therefore, in total, we simulated 648 scenarios and for each scenario, we simulated one thousand datasets with the R software.

Each simulated dataset was analyzed using three models: (i) M_T estimating only the treatment effect β_T and assuming no gene effect, no gene-treatment interaction, (ii) M_{TG} estimating the treatment effect β_T and the gene effect β_G but assuming no gene-treatment interaction, and (iii) M_{TGI} estimating the treatment effect β_T, the gene effect β_G, and the gene-treatment interaction effect β_I. In all three models, between and within subject variances on D_{xo} and between subject variances on D_p were estimated. We used the R package nlme to fit the simulated datasets.[24]

In all scenarios, we evaluated the type I error and the power of the bilateral Wald tests at the level 0.05 to detect (i) a treatment effect $H_0 : \beta_T = 0$, when the data were fitted with M_T, M_{TG} or M_{TGI}, (ii) a gene effect $H_0 : \beta_G = 0$, when the data were fitted with M_{TG} or M_{TGI}, and (iii) a gene-treatment interaction effect $H_0 : \beta_I = 0$, when the data were fitted with

M_{TGI}. The 95% prediction interval around 0.05 for 1000 simulated datasets is (0.037; 0.065). We calculated the estimation errors: $\widehat{\beta_k} - \beta_k^*$ where β_k^* is the true simulated value, for all parameters on all scenarios.

Treatment attribution error

We also explored the treatment attribution error. For each scenario, the correct treatment attribution could be determined according to the patient genotype. For example, for the scenario $\beta_T = 5$; $\beta_G = 0$ and $\beta_I = -10$, for a patient with $G = 0$, if $T = 0$ the predicted outcome $Y_{\text{pred T = 0 and G = 0}} = 8$ and if $T = 1$ then $Y_{\text{pred T = 1 and G = 0}} = 13$, so $Y_{\text{pred T = 0 and G = 0}} < Y_{\text{pred T = 1 and G = 0}}$, the patient should be assigned the treatment test. However, for a patient with $G = 1$, if $T = 0$ the outcome $Y_{\text{pred T = 0 and G = 1}} = 8$ and if $T = 1$ then $Y_{\text{pred T = 1 and G = 1}} = 3$, so $Y_{\text{pred T = 0 and G = 1}} > Y_{\text{pred T = 1 and G = 1}}$, the patient should not be assigned the treatment test. Table 2 illustrates the decision rules based on the model fitted (M_T, M_{TG} or M_{TGI}), the test result on β_T and the sign of $\beta_T + \beta_I$ highlighting the attribution error cases.

The error was thereafter calculated as the percentage of simulated datasets selecting a model leading to the wrong treatment attribution for each genotype and in the whole population (i.e., averaging over the genotype frequency in the population, F).

RESULTS

Parameter estimation

The estimation errors on all parameters for all scenarios are presented in Supplementary Material Figures S1–S4.

For the mean effect μ, ignoring the gene effect (using M_T) resulted in a downward trend in the estimation errors for μ driven by the value of F and β_G. Conversely ignoring the gene-treatment interaction (using M_{TG}) resulted in an upward trend in the estimation errors for μ driven by the value of F and β_I. Using M_{TGI}, no trend was observed.

For β_T, using M_T or M_{TG} led to a downward trend in the estimation errors for β_T driven by F and β_I, especially for D_P and a decreasing R_G. Using M_{TGI}, there was no trend with greater uncertainty for D_P and a decreasing R_G.

For β_G, using M_{TG}, a downward trend in the estimation errors for β_G was observed driven by F and β_I, with greater uncertainty for D_{XO} and a low F. Using M_{TGI}, there was no trend and also greater uncertainty for D_{XO} and a low F.

For β_I, using M_{TGI}, no trend was detected with greater uncertainty for D_P and a decreasing R_G.

Type I error and power

In the following, we only detail the results for D_{xo} ($N = 100$ with $R_G = 0.7$) and D_p ($N = 200$) when $F = 0.2$, and varying β_T; β_G and β_I. Of note, all the results, for all other simulated values of F, N, and R_G, are in Supplementary Material Figures S5–S7.

Figure 1 illustrates the type I error and power to reject $H_0 : \beta_T = 0$ when fitting the data with M_T, M_{TG}, or M_{TGI} for D_{xo} or D_p. When the data were fitted with M_{TGI}, no type I error inflation was observed for both D_{xo} and D_p, whatever the values of β_G or β_I (i.e., even when $\beta_I = 0$). When ignoring the gene-treatment interaction (i.e., fitting the data with M_T or M_{TG}), a type I error inflation was observed with increasing values of β_I for D_{xo} (up to 50% for $\beta_I = -10$) and to a lesser extend for D_p. Fitting

TABLE 2 Treatment prediction ($T_{pred} = 1$ for the test treatment and 0 for the reference treatment) according to fitted model and Wald test on the treatment effect β_T, the gene-treatment interaction effect β_I and the sign of the sum $\beta_T + \beta_I$, for a patient with $G = 0$ and a patient with $G = 1$

			$G = 0$ ($T_{true} = 1$)			$G = 1$ ($T_{true} = 0$)		
			Tests on βT					
Fitted model			**Not significant**	**Significant with $\beta_T > 0$**	**Significant with $\beta_T < 0$**	**Not significant**	**Significant with $\beta_T > 0$**	**Significant with $\beta_T < 0$**
M_T or M_{TG}			$T_{pred} = 0$	$T_{pred} = 1$	$T_{pred} = 0$	$T_{pred} = 0$	$T_{pred} = 1$	$T_{pred} = 0$
M_{TGI}	Tests on β_I	Not significant	$T_{pred} = 0$	$T_{pred} = 1$	$T_{pred} = 0$	$T_{pred} = 0$	$T_{pred} = 1$	$T_{pred} = 0$
		Significant with $\beta_T + \beta_I > 0$	$T_{pred} = 0$	$T_{pred} = 1$	$T_{pred} = 0$	$T_{pred} = 1$	$T_{pred} = 1$	$T_{pred} = 1$
		Significant with $\beta_T + \beta_I < 0$	$T_{pred} = 0$	$T_{pred} = 1$	$T_{pred} = 0$	$T_{pred} = 0$	$T_{pred} = 0$	$T_{pred} = 0$

Notes: The gray boxes correspond to attribution error cases, for example, when a patient with $G = 0$ is attributed the reference treatment ($T_{pred} = 0$ whereas $T_{true} = 1$) or a patient with $G = 1$ is attributed the test treatment ($T_{pred} = 1$ whereas $T_{true} = 0$).

Abbreviations: β_T treatment effect, β_I gene-treatment interaction effect. M_T only treatment effect, M_{TG} only treatment and gene effects, and M_{TGI} treatment, gene and gene-treatment interaction effects.

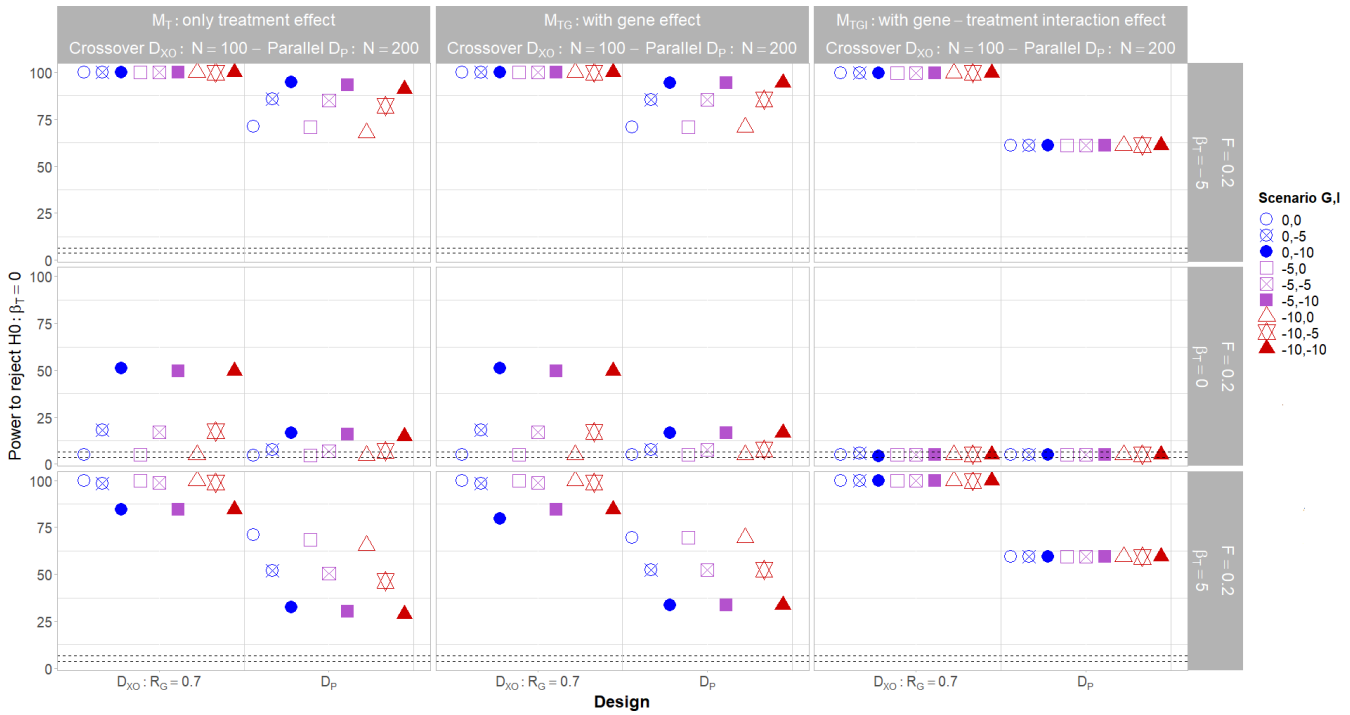

FIGURE 1 Type-I-error (when $\beta_T = 0$, second line) and power (when $\beta_T = -5$ or $+5$) to reject $H_0 : \beta_T = 0$ according to the three fitted models (M_T only treatment effect, M_{TG} only treatment and gene effects, and M_{TGI} treatment, gene, and gene-treatment interaction effects) for the scenarios where the frequency of $G = 1$, $F = 0.2$ and the number of subjects $N = 100$ for crossover trials (D_{xo}) (with the ratio of variability $R_G = 0.7$), and $N = 200$ for parallel trials (D_p)

the data with M_{TGI}, a 100% power to reject H_0 $\beta_T = 0$ was achieved for D_{xo} versus 60% for D_p, whatever the value of β_G or β_I (i.e., even when $\beta_I = 0$). Fitting the data with M_T or M_{TG}, the power dropped when β_T and β_I were not of the same sign, down to about 80% for D_{xo} and 30% for D_P when $\beta_T = 5$ and $\beta_I = -10$. Conversely, when $\beta_I = 0$ or was of the same sign as β_T the power actually increased (up to 90% for D_p without adjusting for the type I error inflation).

Figure 2 illustrates the type I error and power to reject H_0 $\beta_G = 0$ when fitting the data with M_{TG} or M_{TGI} for D_{xo} or D_p. Similar plots were obtained for other simulated values of β_T (Figure S6). The design had little impact on the type I error and power to reject H_0 $\beta_G = 0$. However, using M_{TGI}, no inflation of the type I error was observed whatever the simulated value of β_I or β_T, and the power was driven by β_G only. Using M_{TG}, an inflation of the type I error was observed, driven by β_I and the power was driven by both β_G and β_I to a lesser extent.

Figure 3 illustrates the type I error and power to detect a gene-treatment interaction effect ($H_0 : \beta_I = 0$) for D_{xo} or D_p. Similar plots were obtained for other simulated values of β_T (Figure S7). No inflation of the type I error was observed whatever the simulated value of β_T or β_G and the power to detect a gene-treatment interaction effect was three times higher for D_{xo} compared to D_p for a strong interaction and twice higher for a mild interaction.

Influence of N, F, and R_G

When the data were fitted with M_{TGI}, the power to reject H_0 $\beta_T = 0$ or H_0 $\beta_I = 0$ increased with N and R_G(for D_{XO}) (Figures S5 and S7). F only influenced the power to detect a gene-treatment interaction effect. The power to reject H_0 $\beta_G = 0$ was not affected by R_G or the study design.

Whereas ignoring the gene-treatment interaction (i.e., fitting the data with M_{TG}), an inflation of the type I error was observed driven by F and largely for D_p. The power to reject H_0 $\beta_G = 0$ increased with N and F, largely for D_P (Figure S6).

When ignoring the gene effect and the gene-treatment interaction (i.e., fitting the data with M_T or M_{TG}), the type I error and the power to reject H_0 $\beta_T = 0$ were driven by N, F, and R_G. Of note, R_G had an opposite effect on the power to reject H_0 $\beta_T = 0$ according to the positive or negative simulated value of β_T. The power to reject H_0 $\beta_T = 0$ increased with increasing R_G when β_T and β_I had the same sign conversely.

Treatment attribution error

Figure 4 and Table S2 illustrate the treatment attribution error (in percentage) per genotype and in the whole population, for scenario $\beta_T = 5$; $\beta_G = 0$ and $\beta_I = -10$, as a function of N, F, the design and R_G. In that scenario, the test treatment ($T = 1$) should

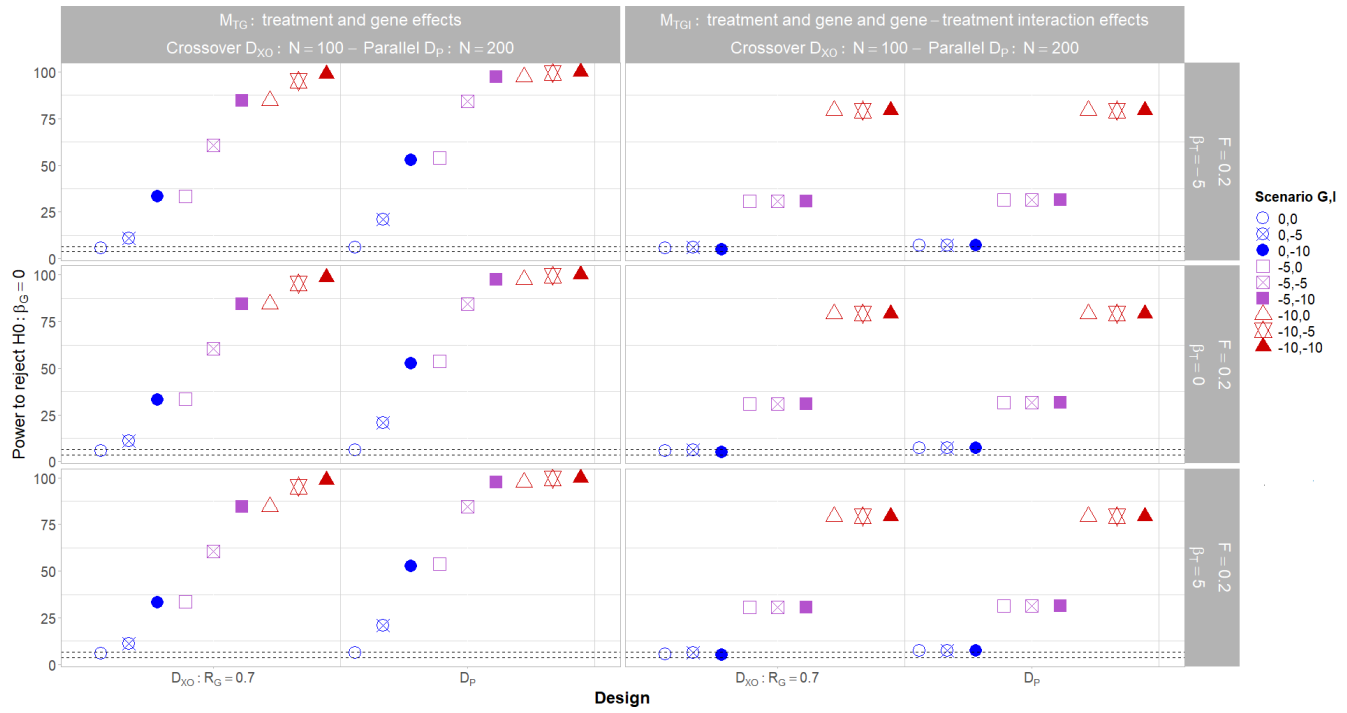

FIGURE 2 Type-I-error (when $\beta_G = 0$, circle symbols) and power (when $\beta_G = -5$, square symbols or -10, triangle symbols) to reject H_0 $\beta_G = 0$ according to the two fitted models (M_{TG} treatment and gene effects, and M_{TGI} treatment, gene, and gene-treatment interaction effects) for the scenarios where the frequency of $G = 1$, $F = 0.2$ and the number of subjects $N = 100$ for crossover trials (D_{xo}) (with the ratio of variability $R_G = 0.7$), and $N = 200$ for parallel trials (D_p)

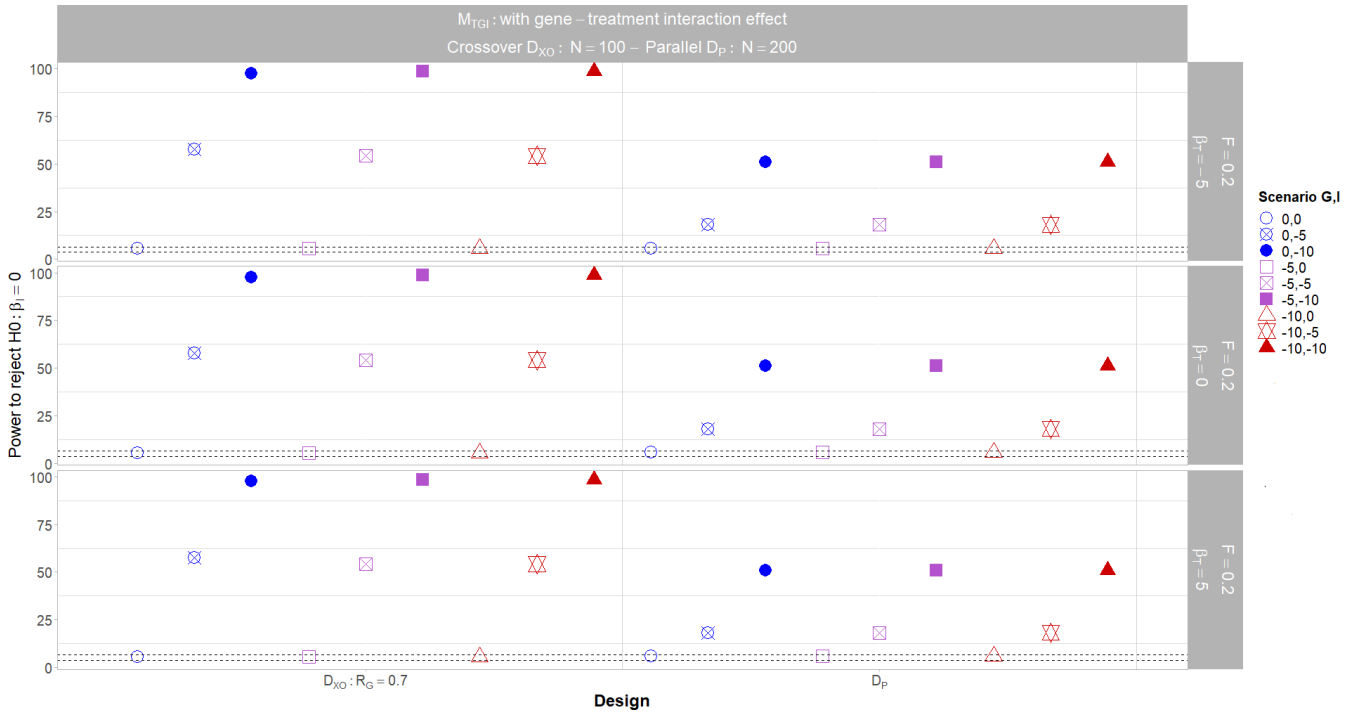

FIGURE 3 Type-I-error (when $\beta_I = 0$, empty symbols) and power (when $\beta_I = -5$, cross symbols or -10, full symbols) to reject H_0 $\beta_I = 0$ according to the fitted model M_{TGI} (treatment, gene, and gene-treatment interaction effects) for the scenarios where the frequency of $G = 1$, $F = 0.2$, and the number of subjects $N = 100$ for crossover trials (D_{xo}) (with the ratio of variability $R_G = 0.7$), and $N = 200$ for parallel trials (D_p)

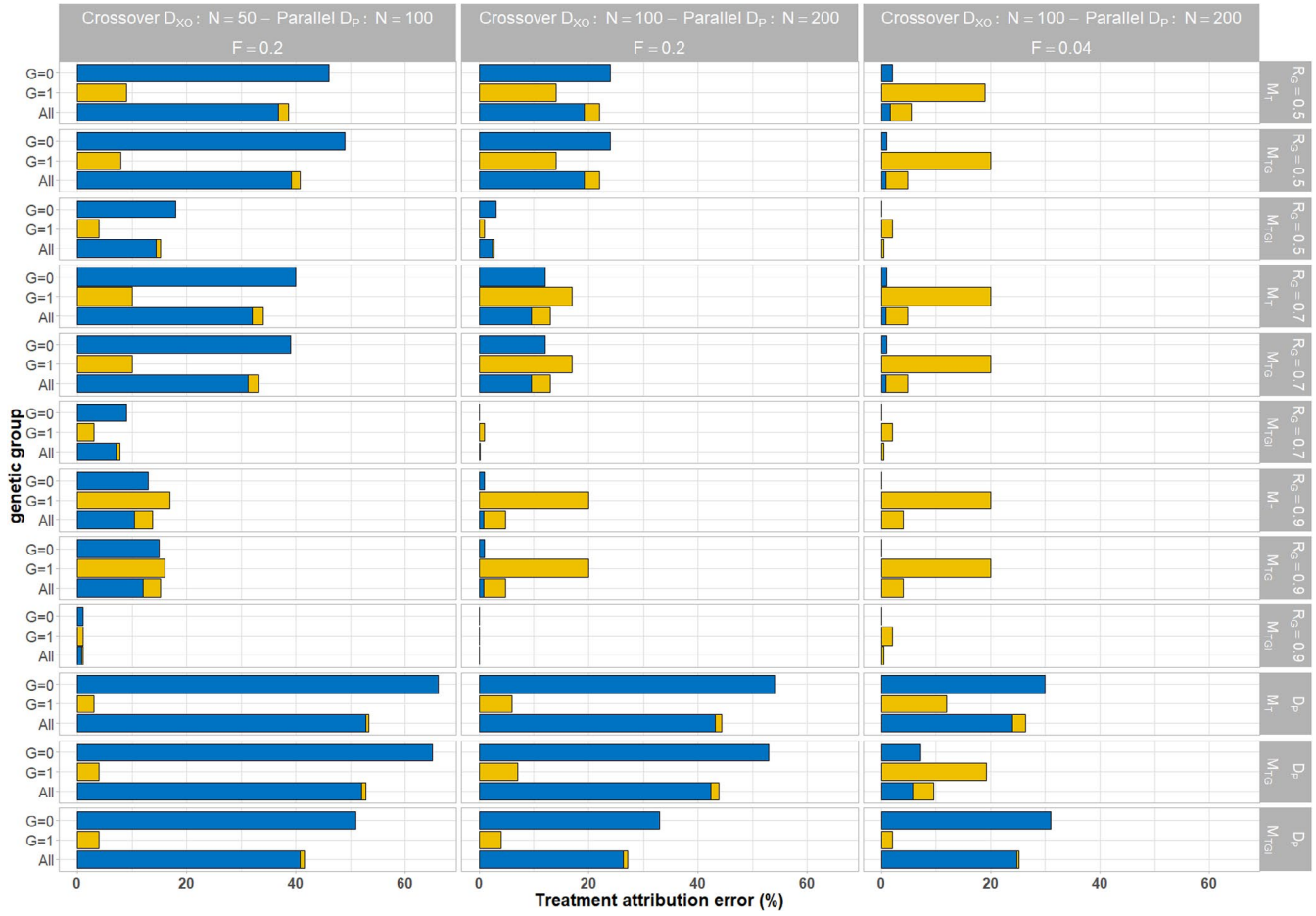

FIGURE 4 Treatment attribution error ($T = 0$ or 1) (in %) per genetic group ($G = 0$ or 1 and in whole population according to the three fitted model (M_T only treatment effect, M_{TG} only treatment and gene effects, and M_{TGI} treatment, gene, and gene-treatment interaction effects). Results are displayed as a function of the frequency F of $G = 1$, the design (crossover D_{xo} and parallel D_P), the number of subjects N, and the ratio of variability R_G (not applicable for D_p). Results are presented for the scenario with treatment effect $\beta_T = 5$, gene effect $\beta_G = 0$ and gene-treatment interaction effect $\beta_I = -10$ (i.e., the test treatment [$T = 1$] should be assigned to a patient with $G = 0$ and the reference treatment [$T = 0$] should be assigned to the patient genotype $G = 1$)

be assigned to a patient with $G = 0$ and the reference treatment ($T = 0$) should be assigned to the patient genotype $G = 1$. For example, using M_T for the scenario with $F = 0.2$, $R_G = 0.5$ and $N = 100$ for D_{XO}, we have 24% of attribution error in patients $G = 0$, 14% in patients $G = 1$, and 22% in the whole population.

The treatment attribution error in the whole population was lower for D_{xo} (13% for $N = 100$ and $R_G = 0.7$ when $F = 0.2$) than for D_p (44% for $N = 200$ when $F = 0.2$) using M_T or M_{TG}. The treatment attribution error in the whole population was even lower using M_{TGI} with 1% for D_{xo} ($N = 100$ and $R_G = 0.7$ when $F = 0.2$) versus 27% for D_p ($N = 200$ when $F = 0.2$). The treatment attribution error in the whole population decreased with increasing N and/or R_G and decreasing F. Of note, D_p led to consistently higher treatment attribution errors in patients with genotype $G = 0$ with, for example, an estimate of 54% versus 6% in patients with genotype $G = 1$ using M_T on the scenario with $F = 0.2$ and $N = 200$. Conversely, D_{xo} led to higher treatment attribution errors in patients with genotype $G = 1$ when $R_G > 0.5$ and $N > 100$ with, for example, an estimate of 17% versus 12% in patients with genotype $G = 0$ using M_T on the scenario with $F = 0.2$, $N = 100$ and $R_G = 0.7$.

DISCUSSION

As shown with this simulation study, first the choice of the model and second the choice of the trial design strongly affects not only the statistical type I and power to detect a gene-treatment interaction in pharmacogenetic studies but also the correct treatment attribution. Indeed, ignoring a true gene-treatment interaction in the model led, notably, to biased treatment effect estimates and inflated type I, whereas no penalty is paid when accounting for a nonexistent gene-treatment interaction. Further, to capture adequately the gene-treatment interaction, a crossover design is more powerful than a parallel design.

First, we note that the gene-treatment interaction effect size strongly affects the power to detect a treatment effect, whereas

the gene effect size has little influence but on the standard error. Indeed, in agreement with our study, the sample size simulation studies by Cardon et al.[25] and Puangpetch et al.[26] highlighted the association between sample size ratios and the genetic model, frequency, and effect size. In a crossover design, baseline covariates not impacting the within-subject variability, have limited impact on the power to detect a treatment effect. However, polymorphisms can sometimes impact the within subject variability. Indeed, Alfaro et al.[27] observed an increase in power to detect a treatment difference, when accounting for the CYP2D6 polymorphism in a crossover design due to decreased within variability between genotypes. More specifically, Gonzalez-Vacarezza et al.[28] and Cabaleiro et al.[29] have shown how selecting patients on the basis of their CYP2D6 their CYP2D6 metabolizer status could lower the sample size of bioequivalence studies thanks to a decreased within-subject variance in extreme metabolizer groups.

If the gene-treatment interaction effect is opposite to the treatment effect, the latter is completely masked except when accounting for the interaction in the model.[4,19,27] Good estimates of treatment effect size are only a means to the treatment attribution end.[1,16] We focused on a scenario with a positive treatment effect and a strong opposite gene-treatment interaction to illustrate how the model or design choice affected the treatment attribution, according to the genotypes. Our study confirms that neglecting the gene-treatment interaction effect had a real impact on the attribution treatment for all genotypes, as shown in Figure 4. Therefore, in such specific cases, quantifying the gene-treatment interaction would be essential to attribute the best therapeutic option (treatment or regimens).[14] The gene effect size had little impact on treatment attribution compared to the effect size of the gene-treatment interaction, which only applies when administering the new treatment. Whereas the gene effect size applies whatever the given treatment (reference or new).[25,26]

This work has limitations. Our example mimics symptoms modifying drugs for a chronic disease enabling us to assume no sequence or carry-over effects. It corresponds, for example, to the study by Reichert et al. whom identified an interaction of sleep pressure and the ADA rs73598374 polymorphism on sleepiness using a crossover design.[30] Similarly, Lopez-Minquez et al. identified an interaction of physiological melatonin and the MTNR1B rs10830963 polymorphism on glucose tolerance in a crossover study.[31] In a more pharmacological context, Park et al. explored the effects of itraconazole and CYP2D6*10 genetic polymorphism on the pharmacokinetics and pharmacodynamics of haloperidol in a crossover study.[32] Of note in these studies, the magnitude of the treatment, gene, and interaction effects varied from 10% to 81%. However, pharmacogenetic studies cannot always ignore the disadvantages of crossover designs (e.g., the carry-over effect), the handling of drop-outs and their unsuitability for disease modifying treatments. The carry-over effect can be anticipated at the design stage with an appropriate washout period (for example 1 month) and dropouts will require sensitivity analyses, but for disease-modifying treatments, only a parallel design can be considered.

We also considered the effect of only one recessive polymorphism. A perspective work would be to explore two polymorphisms with opposite or synergic effects and to explore an additive polymorphism model. Further, the simulated ranges of genotype frequencies (4% and 20%) may appear too high and/or our sample sizes too low. However, Cardon et al.[25] in their simulation study showed that, in the presence of a polymorphism advantageous for the treatment under study, a pharmacogenomic trial requires a smaller sample size than a traditional trial (not adjusting on the polymorphism) to detect the same effect. Further, Katara et al.[5] argued that polymorphisms could be responsible for around 30%–40% of the overall functional variability and significantly impacts drug response differences. In this context, one may expect a common polymorphism to have an intermediate-to-strong effect and not necessarily a combination of many common polymorphisms with a small effect.

To conclude, based on realistic simulations, we highlighted how ignoring an existing gene-treatment interaction results in incorrect treatment effect estimates. Several pharmacogenetic studies have acknowledged that small sample size was their main limitation.[15,26,33] However, we showed that a study design and analysis plan based on a full model with a gene-treatment interaction term could overcome such a limitation. Indeed, crossover designs proved to be more powerful to detect a treatment effect in the presence of a gene-treatment interaction, here, in a simulation framed in the context of a chronic disease with a quick rapid and reversible treatment effect (a short period of washout maximum of 1 month) and no carry-over effect. Therefore, assuming a gene-treatment interaction and using a crossover design seems the best strategy for the pharmacogenetic study of concurrent drug treatments. Well conducted clinical trials to explore efficacy and/or tolerance accounting for candidate polymorphisms appears an inevitable step for the development of personalized medicine.[16,26] Our study partly addresses the challenge of developing pharmacogenetic tools in the context of well-known candidate polymorphisms as framed by Claassens et al.[34] Here, we demonstrated the advantages of crossover designs and of accounting for gene-treatment interaction in the analysis. We hope our results will help improve future pharmacogenetic studies.

DISCLAIMER

As Editor-in-Chief of *CPT: Pharmacometrics & Systems Pharmacology*, France Mentré was not involved in the review or decision process for this paper.

ACKNOWLEDGEMENTS

The authors thank Professor François Gueyffier, principal investigator of IDEAL study.

CONFLICT OF INTEREST

The authors declared no competing interests for this work.

AUTHOR CONTRIBUTIONS

C.C. and J.B. wrote the manuscript. C.C., F.M., and J.B. designed the research. C.C. performed the research. C.C. analyzed the data.

ORCID

Camille Couffignal https://orcid.org/0000-0002-5006-6249
France Mentré https://orcid.org/0000-0002-7045-1275
Julie Bertrand https://orcid.org/0000-0002-6568-1041

REFERENCES

1. Ross S, Anand SS, Joseph P, Paré G. Promises and challenges of pharmacogenetics: an overview of study design, methodological and statistical issues. *JRSM Cardiovasc Dis*. 2012;1(1):1-13.
2. Collins FS, Varmus H. A new initiative on precision medicine. *N Engl J Med*. 2015;372(9):793.
3. Attia J, Ioannidis JPA, Thakkinstian A, et al. How to use an article about genetic association: B: are the results of the study valid? *JAMA*. 2009;301(2):191-197.
4. Lesko LJ, Salerno RA, Spear BB, et al. Pharmacogenetics and pharmacogenomics in drug development and regulatory decision making: report of the first FDA-PWG-PhRMA-DruSafe workshop. *J Clin Pharmacol*. 2003;43(4):342-358.
5. Katara P, Yadav A. Pharmacogenes (PGx-genes): current understanding and future directions. *Gene*. 2019;15(718):144050.
6. Kirchheiner J, Nickchen K, Bauer M, et al. Pharmacogenetics of antidepressants and antipsychotics: the contribution of allelic variations to the phenotype of drug response. *Mol Psychiatry*. 2004;9(5):442-473.
7. Kelly PJ, Stallard N, Whittaker JC. Statistical design and analysis of pharmacogenetic trials. *Stat Med*. 2005;24(10):1495-1508.
8. Holmes MV, Shah T, Vickery C, Smeeth L, Hingorani AD, Casas JP. Fulfilling the promise of personalized medicine? Systematic review and field synopsis of pharmacogenetic studies. *PLoS One*. 2009;4(12):e7960.
9. Ventola CL. Role of pharmacogenomic biomarkers in predicting and improving drug response: part 1: the clinical significance of pharmacogenetic variants. *P T*. 2013;38(9):545-560.
10. Sluiter RL, Janzing JGE, van der Wilt GJ, Kievit W, Teichert M. An economic model of the cost-utility of pre-emptive genetic testing to support pharmacotherapy in patients with major depression in primary care. *Pharmacogenomics J*. 2019;16:1.
11. Radhakrishnan A, Kuppusamy G, Ponnusankar S, Shanmukhan NK. Pharmacogenomic phase transition from personalized medicine to patient-centric customized delivery. *Pharmacogenomics J*. 2019;10:1-18.
12. Attia S, Egger M, Müller M, Zwahlen M, Low N. Sexual transmission of HIV according to viral load and antiretroviral therapy: systematic review and meta-analysis. *AIDS*. 2009;23(11):1397-1404.
13. Ioannidis JPA. To replicate or not to replicate: the case of pharmacogenetic studies: have pharmacogenomics failed, or do they just need larger-scale evidence and more replication? *Circ Cardiovasc Genet*. 2013;6(4):413-418.
14. Lauschke VM, Ingelman-Sundberg M. The importance of patient-specific factors for hepatic drug response and toxicity. *Int J Molec Sci*. 2016;17(10):1714.
15. Nelson MR, Johnson T, Warren L, et al. The genetics of drug efficacy: opportunities and challenges. *Nat Rev Genet*. 2016;17(4):197-206.
16. Nandal S, Burt T. Integrating pharmacoproteomics into early-phase clinical development: state-of-the-art, challenges, and recommendations. *Int J Mol Sci*. 2017;18(2):448.
17. Curtin F, Altman DG, Elbourne D. Meta-analysis combining parallel and cross-over clinical trials. I: continuous outcomes. *Stat Med*. 2002;21(15):2131-2144.
18. Senn SS. *Cross–over Trials in Clinical Research*, 2nd edn. Chichester, Eng.; New York: Wiley-Blackwell; 2002:364 p.
19. Senn S. Cross-over trials in statistics in medicine: the first '25' years. *Stat Med*. 2006;25(20):3430-3442.
20. Bertrand J, Balding DJ. Multiple single nucleotide polymorphism analysis using penalized regression in nonlinear mixed-effect pharmacokinetic models. *Pharmacogenet Genomics*. 2013;23(3):167-174.
21. Bejan-Angoulvant T, Baguet J-P, Erpeldinger S, et al. The IDEAL study: towards personalized drug treatment of hypertension. *Therapie*. 2012;67(3):195-204.
22. Sawilowsky SS. New effect size rules of thumb. *J Mod App Stat Meth*. 2009;8(2):597-599.
23. de Winter BCM, van Gelder T, Sombogaard F, Shaw LM, van Hest RM, Mathot RAA. Pharmacokinetic role of protein binding of mycophenolic acid and its glucuronide metabolite in renal transplant recipients. *J Pharmacokinet Pharmacodyn*. 2009;36(6):541-564.
24. Pinheiro J, Bates D, DebRoy S, Sarkar D, R Core Team. nlme: Linear and Nonlinear Mixed Effects Models [Internet]; 2020. Available from: https://CRAN.R-project.org/package=nlme.
25. Cardon LR, Idury RM, Harris TJ, Witte JS, Elston RC. Testing drug response in the presence of genetic information: sampling issues for clinical trials. *Pharmacogenetics*. 2000;10(6):503-510.
26. Puangpetch A, Srisawasdi P, Unaharassamee W, et al. Association between polymorphisms of LEP, LEPR, DRD2, HTR2A and HTR2C genes and risperidone- or clozapine-induced hyperglycemia. *Pharmgenomics Pers Med*. 2019;12:155-166.
27. Alfaro CL, Lam YW, Simpson J, Ereshefsky L. CYP2D6 inhibition by fluoxetine, paroxetine, sertraline, and venlafaxine in a crossover study: intraindividual variability and plasma concentration correlations. *J Clin Pharmacol*. 2000;40(1):58-66.
28. González-Vacarezza N, Abad-Santos F, Carcas-Sansuan A, et al. Use of pharmacogenetics in bioequivalence studies to reduce sample size: an example with mirtazapine and CYP2D6. *Pharmacogenomics J*. 2013;13(5):452-455.
29. Cabaleiro T, Ochoa D, Román M, et al. Polymorphisms in CYP2D6 have a greater effect on variability of risperidone pharmacokinetics than gender. *Basic Clin Pharmacol Toxicol*. 2015;116(2):124-128.
30. Reichert CF, Maire M, Gabel V, et al. Insights into behavioral vulnerability to differential sleep pressure and circadian phase from a functional ADA polymorphism. *J Biol Rhythms*. 2014;29(2):119-130.
31. Lopez-Minguez J, Saxena R, Bandín C, Scheer FA, Garaulet M. Late dinner impairs glucose tolerance in MTNR1B risk allele carriers: a randomized, cross-over study. *Clin Nutr*. 2018;37(4):1133-1140.

32. Park J-Y, Shon J-H, Kim K-A, et al. Combined effects of itraconazole and CYP2D6*10 genetic polymorphism on the pharmacokinetics and pharmacodynamics of haloperidol in healthy subjects. *J Clin Psychopharmacol*. 2006;26(2):135-142.
33. Faria NR, Rambaut A, Suchard MA, et al. The early spread and epidemic ignition of HIV-1 in human populations. *Science*. 2014;346(6205):56-61.
34. Claassens DMF, Vos GJA, Bergmeijer TO, et al. A genotype-guided strategy for oral P2Y12 inhibitors in primary PCI. *N Engl J Med*. 2019;381(17):1621-1631.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Couffignal C, Mentré F, Bertrand J. Impact of study design and statistical model in pharmacogenetic studies with gene-treatment interaction. *CPT Pharmacometrics Syst. Pharmacol.* 2021;10:340–349. https://doi.org/10.1002/psp4.12624