



Promises and challenges of pharmacogenetics: an overview of study design, methodological and statistical issues

Stephanie Ross¹ • Sonia S Anand^{1,2,4} • Philip Joseph¹ • Guillaume Paré^{1,3,4}

¹Population Health Research Institute, Hamilton Health Sciences, McMaster University, Hamilton, Ontario L8L 2X2, Canada

²Department of Medicine, McMaster University, Hamilton, Ontario L8N 3Z5, Canada

³Department of Pathology, McMaster University, Hamilton, Ontario L8N 3Z5, Canada

⁴Department of Clinical Epidemiology and Biostatistics, Population Genomics Program, McMaster University, Hamilton, Ontario L8S 4K1, Canada

Correspondence to: Guillaume Paré. Email: pareg@mcmaster.ca

DECLARATIONS

Competing interests

GP reports receiving consulting and speaker fees from Sanofi-Aventis, Bristol-Myers Squibb and Boehringer-Ingelheim, and research grant support from Bristol-Myers-Squibb and Sanofi-Aventis. SSA reports receiving lecture fees from Bristol-Myers Squibb

Funding

GP holds the Canada Research Chair in Genetic and Molecular Epidemiology. SSA holds the Michael DeGrotte and Heart and Stroke Foundation of Ontario Endowed Chair in Population

Summary

Pharmacogenetics is the study of inherited variation in drug response. The goal of pharmacogenetics is to develop novel ways of maximizing drug efficacy and minimizing toxicity for individual patients. Personalized medicine has the potential to allow for a patient's genetic information to predict optimal dosage for a drug with a narrow therapeutic index, to select the most appropriate pharmacological agent for a given patient and to develop cost-effective treatments. Although there is supporting evidence in favour of pharmacogenetics, its adoption in clinical practice has been slow because of sometimes conflicting findings among studies. This failure to replicate findings may result from a lack of high-quality pharmacogenetic studies, as well as unresolved methodological and statistical issues. The objective of this review is to discuss the benefits of incorporating pharmacogenetics into clinical practice. We will also address outstanding methodological and statistical issues that may lead to heterogeneity among reported pharmacogenetic studies and how they may be addressed.

Introduction

It is widely recognized that there is interindividual variability in drug response, where subgroups of patients experience either adverse drug reactions or do not respond properly to treatment.¹ While the definition of individualized response to drug treatment is not yet fully understood² and there is uncertainty as to whether certain patients are consistent non-responders or simply inconsistent responders, this variability may be attributed to biological factors (i.e. age, sex, nature of disease), behavioural factors (i.e. smoking, drug interactions) or genetic

factors (i.e. genetic variants). Furthermore, the lack of patient adherence is also recognized as an important contributor to variability of response. For example, the discontinuation of antiplatelet therapy is the strongest risk factor for stent thrombosis in percutaneous coronary intervention.³ Nonetheless, it is estimated that genetic factors can account for 20–95% of individual variation in drug response;⁴ however, the amount of explained variation depends on the class of drugs.

The wide variability in drug response emphasizes the need for a more 'personalized' approach to medical treatment. It is possible that pharmacogenetics can address this need by providing a

Health and the May
Cohen Eli Lily
Endowed Chair in
Women's Health

Ethical approval
Not applicable

Guarantor
GP

Contributorship
SR and GP wrote the
first draft. SSA and
PJ provided
comments on all
drafts and supplied
additional content
and relevant
references

Acknowledgements
None

better understanding of how genetic variants influence drug response.⁵ This review will focus primarily on pharmacogenetics, which assesses how genetic variants influence drug metabolism and effect. The ultimate goal of pharmacogenetics is to develop novel ways of minimizing harmful drug effects and optimizing care for individual patients. More specifically, a patient's genetic information may be used to predict the optimal dosage for a drug with a narrow therapeutic index, to select the most appropriate pharmacological agent and to develop cost-effective treatment plans.

Despite the promise of personalized medicine, there has been little methodological consistency among pharmacogenetic studies. This may be due to modest effect sizes, heterogeneity among study designs and patient populations, as well as a lack of standardization among biological and phenotypic measures.^{6–8} Holmes *et al.*⁹ performed a systematic review and a field synopsis of pharmacogenetic studies. They reported that the lack of consistency among studies may be a result of the preponderance of reviews over primary research, small sample sizes, mainly a candidate gene approach, surrogate markers, an excess of nominally positive to truly positive associations and paucity of meta-analyses. Therefore, there is an urgent need for properly designed pharmacogenetic studies to advance the discovery and development of medical strategies for individualized treatment. The objective of this review is to discuss the potential benefits of incorporating pharmacogenetics into clinical practice, as well as methodological and statistical challenges faced in pharmacogenetic studies. In this review, we will first identify potential clinical applications of pharmacogenetics and illustrate these with promising contemporary examples. In the second part, we will summarize some of the major methodological challenges facing pharmacogenetic studies.

Potential applications of pharmacogenetics

Personalized medicine has the potential to improve drug safety and efficacy for a specific individual. Adoption of pharmacogenetics in clinical practice promises more effective

decision-making with regard to diagnostic testing, drug selection and dosing. In this section, we describe some future applications of pharmacogenetics and provide contemporary examples that reflect how these topics may be applied to a clinical setting. It should be noted by readers that, in many instances, further research is needed to unequivocally recommend pharmacogenetic testing. It is widely believed that a better understanding of the genetic mechanisms in drug response has the potential to help clinicians predict an individualized drug dosage; however, to date, there are few examples that illustrate this hypothesis with improved clinical outcomes.

Individualized drug dosage

The genetic variants that influence the observed differences in drug response can be classified into two groups: pharmacokinetics (PK) and pharmacodynamics (PD). The genes that influence the PK properties of a drug affect the mechanisms of how the drug is absorbed, distributed, metabolized and excreted by the body. The genes that influence the PD of a drug affect the mechanism of the drug's target and how it affects the body. One underlying principle of individualized drug dosage is that it must be faster and more effective than the use of a PK or PD assay alone. In other words, genetic testing may not be required if the therapeutic level of the drugs or a surrogate can be measured, and it is rapidly available and widespread, such as the case with certain antibiotics (notwithstanding the genetic susceptibility of the pathological agent).

For example, warfarin has a narrow therapeutic index, and inadequate or excessive anticoagulation can lead to an increased risk of adverse cardiovascular events or bleeding complications. Thus, warfarin therapy dosage is complicated by individual variability and requires regular monitoring to achieve proper anticoagulation effects. Initial warfarin therapy is administered by a fixed dosage, or by an estimated regimen based on the patient's clinical characteristics, with further adjustments based on the patient's anticoagulation response measured by laboratory assays, such as the international normalized ratio (INR). However, it

may be more beneficial to use both genetic factors and clinical covariates as opposed to frequent INR monitoring because genetic polymorphisms account for 30–35% of the variability in warfarin metabolism and clinical factors account for 17–21% of variation in warfarin dosing.¹⁰ Therefore, an algorithm that incorporates a combination of these factors would ultimately improve the time required to establish a stable maintenance dose.

The principal genes involved in the metabolism of warfarin are the cytochrome P450 (CYP) 2C9 enzyme and the vitamin K epoxide reductase complex, subunit 1 (*VKORC1*) gene. Carriers of at least one or more variant alleles of the *CYP2C9* genotype are associated with overcoagulation and an increased risk of bleeding while on warfarin therapy,^{11–13} whereas those who possess the variant *VKORC1* genotype experience warfarin treatment resistance and an increased risk of adverse cardiac events.¹⁴ The International Warfarin Pharmacogenetics Consortium developed a pharmacogenetic algorithm for an appropriate warfarin dosage.¹⁵ The study reported that, among 5000 participants, the pharmacogenetic algorithm identified a larger proportion of patients who required a lower dose (≤ 21 mg per week) of warfarin and those who required a higher dose (≥ 49 mg per week) to maintain stable therapeutic anticoagulation. The genetically guided treatment benefited 46.2% of the entire cohort, specifically those for whom the standard dosage of warfarin would not be appropriate. It is important to properly identify this proportion of patients because some (i.e. who require ≤ 21 mg per week) are at risk of excessive anticoagulation, whereas others who require a higher dose of warfarin (i.e. ≥ 49 mg/week) are at risk of inadequate anticoagulation. Data on adverse events such as thromboembolic events or bleeding were not collected for this study.

In another study, patients who were treated using a pharmacogenetic algorithm had 28% less hospitalizations after six months of warfarin therapy compared with a control group (18.5% versus 25.5%, $P < 0.001$).¹⁶ The ability to increase the accuracy of dose prediction may help to enhance drug efficacy and drug safety associated with under-dosing or over-dosing patients. Although promising, it should be emphasized that the implementation of pharmacogenetic

testing ultimately depends on clear evidence of improved clinical outcomes.

Individualized drug selection

Personalized medicine can help to guide individualized treatment when the clinical effect of a drug is expected to vary according to genotype. Under these conditions, the risk–benefit balance of a drug might depend on the variant allele carrier status. This balance can be affected by pharmacogenetic effects on safety, efficacy, or both. For example, the incidence of adverse clinical events may differ according to genotypic groups, if for instance, slow metabolizers accumulate a toxic metabolite. Thus, prior knowledge of a patient's genotype may be used to guide clinical decision-making because the patient may benefit from an alternative pharmacological regimen, such that they receive a reduced dose of a standard therapy or a different drug altogether. Conversely, patients who are classified as fast metabolizers may experience increased drug efficacy from a higher dose if their genetic status results in accelerated clearance of the active metabolite.

Chronic infection with hepatitis C virus (HCV) is treated with a combined therapy of peginterferon- α -2a (PegIFN- α -2a) or PegIFN- α -2b and ribavirin. However, less than half of treated patients achieve a sustained virological response (SVR).¹⁷ A genome-wide association study of 1671 chronic HCV patients reported that a genetic polymorphism in the *IL28B* gene (rs12979860) region was strongly associated with SVR.¹⁸ The authors reported that the polymorphism was associated with a two-fold change in treatment response among Caucasians ($P = 1.06 \times 10^{-25}$) and African-Americans ($P = 2.06 \times 10^{-3}$). Interestingly, the differences in allelic frequency of the *IL28B* genetic variant may explain about half of the difference in treatment response between these two ethnic groups.

Another study assessed whether accounting for the human leukocyte antigen C (HLA-C) and the killer immunoglobulin-like receptors improved the predictive value of the *IL28B* genotype.¹⁹ The authors found that the carriers of the variant *IL28B* genotype were associated with absence of treatment-induced HCV infection clearance and absence of spontaneous HCV infection clearance.

Furthermore, carriers of the variant *HLA-C* genotype were associated with failed treatment but not spontaneous HCV infection clearance. Thus, the prediction of treatment failure among HCV patients was improved from 66% using the *IL28B* genotype to 80% with the use of the *IL28B* and *HLA-C* genotypes. Incorporating this information can help clinicians to improve the clinical management of patients infected with chronic HCV because they will be able to better predict those who will respond the best to PegIFN treatment, which will help to reduce the adverse side-effects associated with this treatment.

Pharmacoeconomy

Pharmacogenetics has the potential to reduce the costs associated with inappropriate drug treatments or serious adverse drug reactions that require hospitalization.²⁰ Pharmacoeconomic considerations are especially important given the moderate effects of genetic determinants typically reported in pharmacogenetic studies. In other words, if a more expensive drug has a slightly decreased benefit in individuals with a certain genotype, then careful evaluation of the costs associated with an alternative therapy or the cost of genotyping is necessary before recommending further pharmacogenetic testing.

One example of utilizing genetic testing to improve cost-effectiveness is the treatment of HIV-positive patients with abacavir, a nucleotide reverse-transcriptase inhibitor. Abacavir hypersensitivity syndrome (AHS) is a potentially lethal side-effect, affecting 5–8% of patients in the first six weeks of treatment.²¹ It presents with a constellation of symptoms such as fever and rash; and rechallenge with abacavir, after initial therapy, may result in worsening AHS symptoms with an increased risk of mortality.²² Patients who experience AHS are strongly associated with the variant histocompatibility complex class I allele (*HLA-B*)*5701 genotype, which is present in 2–6% of Caucasians.²³

Mallal *et al.*²³ observed that, in a double-blind prospective randomized study, Prospective Randomized Evaluation of DNA screening in a Clinical Trial (PREDICT-1), selective abacavir use informed by *HLA-B**5701 testing reduced the risk of AHS. The authors of this study reported that screening eliminated AHS (0% in the

prospective-screening group versus 2.7% in the control group, $P < 0.001$), and had a negative predictive value of 100% and a positive predictive value of 47.9%. This led to the recommendation that prospective *HLA-B**5701 screening should be adopted in clinical care.^{24,25}

Furthermore, several studies have evaluated the cost of prospective *HLA-B**5701 screening.^{26–28} Kauf *et al.*²⁸ analysed the cost-effectiveness of *HLA-B**5701 screening by assessing the cost of prior genetic screening and the cost of using an alternative medication, tenofovir, within short-term and lifetime models. The authors reported that the short-term costs of prospective screening were dependent on the cost of the genetic test, the cost associated with AHS treatment and screening performance. The lifetime models showed that genetically guided abacavir treatment was more effective and less costly than alternative treatment with tenofovir. Furthermore, as of 2009, the patent for abacavir has expired in the United States. Thus, the cost-effectiveness of *HLA-B**5701 screening prior to abacavir-based treatment is now highly dependent on the prevalence of the *HLA-B**5701 genotype, the cost of prescribing a generic medication compared with a non-generic one, screening costs and the method of health-care funding.

Methodological issues in pharmacogenetics

Although pharmacogenetics has the potential to address variability in drug response and improve drug efficacy and safety, the adoption of pharmacogenetics in clinical practice has been slow. This resistance may stem from sometimes conflicting findings among pharmacogenetic studies. The failure to replicate these findings may result from a lack of high-quality studies and unresolved methodological issues. In this section, we address methodological issues pertaining to pharmacogenetic study design and provide specific examples of pharmacogenetic studies that illustrate potential challenges that the reader may encounter.

Study design

Table 1 provides a brief description of each study design.

Table 1
Study designs for pharmacogenetic studies and their main strengths and limitations

Study design	Description	Strength	Limitation
Randomized controlled trial	Participants are stratified by genotype and randomized to treatment groups	Evidence of a drug–gene interaction; evidence of causality; assess multiple outcomes	Requires large sample size; high cost; unable to assess rare events
Prospective cohort study	Participants are followed over time and disease outcome is compared with drug and genotype subgroups	Prospective nature; assess multiple outcomes	Selection bias (loss-to-follow-up); information bias (non-differential); confounding; unable to assess rare events
Case-control study	The genotype frequency and drug response outcome is compared among cases and controls	Requires small sample size; low cost; assess rare events	Selection bias; information bias (differential); confounding; unable to assess rare events

Randomized controlled trials

Randomized controlled trials (RCTs) remain the ‘gold standard’ in epidemiological study design. In the field of pharmacogenetics, there are two ways in which RCTs can be used to establish pharmacogenetic determinants of drug safety and efficacy. First, patients can be randomized to a genetically guided therapy versus standard care. While this design offers the best level of evidence to support the use of genetic data, it may be impractical in some situations. For example, the speed of genotyping may cause delays in treatment or randomization for trials that require known pharmacogenetic determinants. Alternatively, if the genotype of interest is rare and the aim of the study is to compare response between two or more therapeutic regimen among carriers, participants may be stratified based on their genotype and then randomized to the intervention or control group.

Substudies within RCTs can be used to determine the impact of genetic variants in response to drug outcomes. In these studies, stored biological samples from pre-existing clinical trials are genotyped with power comparable with that of a prospectively planned pharmacogenetic cohort study. This appears to be an optimal design to discover and characterize pharmacogenetic determinants prior to an evaluation of gene-guided therapy versus standard care.

Pharmacogenetic RCTs are able to measure the independent effects of the genotype, the drug response and the gene–drug interaction in the active drug and placebo/control groups. With this approach, it is then possible to distinguish the differences between simple markers of disease progression and true pharmacogenetic markers, whose effect on disease progression is only seen in the presence of a drug. This can also be assessed by developing a ‘gene score’ (i.e. combining information from many single nucleotide polymorphisms [SNPs]) and testing for a drug–gene interaction.

One major limitation of pharmacogenetic RCTs is the cost and time required to conduct the study. These studies require a large sample size to be powered enough to detect a modest effect size. Furthermore, a *post hoc* analysis of an RCT may be inappropriate for a pharmacogenetic study because the initial cohort was designed using a specific null hypothesis, estimated effect size and study power, and may underestimate the true gene–drug interaction.

An example of a genetically guided RCT is the Clarification of Optimal Anticoagulation through Genetics (COAG) trial.²⁹ The COAG trial is a randomized, double-blinded clinical trial that compares genotype-guided dosing and clinical-guided dosing for the initiation of warfarin treatment. The objective of the trial is to determine whether genetic information improves drug treatment. This trial is ongoing and final results of the study are yet to be published.

Another example of a pharmacogenetic RCT is the Statin Response Examined by Genetic Haplotype Markers (STRENGTH) study.³⁰ The purpose of the STRENGTH study was to explore the association between genetic polymorphisms and low-density lipoprotein cholesterol (LDLc) lowering in statin-treated patients. The STRENGTH study was a 16-week, randomized, open-label

study of three statins in 509 outpatients with hypercholesterolaemia. Study participants were initially randomized to eight weeks of 10 mg/day atorvastatin, 20 mg/day simvastatin or 10 mg/day pravastatin followed by eight weeks of 80 mg/day atorvastatin, 80 mg/day simvastatin and 40 mg/day pravastatin. Voora *et al.*^{30,31} reported that carriers of the *ABCA1* variant (rs12003906) were associated with a reduced LDLc-lowering effect and carriers of the loss-of-function *SLCO1B1* allele were associated with increased risk of statin therapy discontinuation. The use of pharmacogenetic RCTs will be instrumental in the understanding of how genetic variants contribute to drug therapy and lay a solid foundation for tailored medical therapy.

Another recent RCT example is the effect of the *CYP2C19* genotype on the safety and efficacy of clopidogrel. Dual antiplatelet therapy of clopidogrel and aspirin has been shown to reduce adverse vascular events among patients with acute coronary syndromes.^{32,33} Several studies have observed that carriers of the loss-of-function allele are associated with a reduced response to clopidogrel and an increased risk of adverse cardiovascular outcomes.^{34,35} Based on these findings, in 2010, the Food and Drug Administration put a boxed warning for the prescription of clopidogrel, which may require dose adjustment or use of a different drug.³⁶ However, a genotyped subgroup from the Clopidogrel in Unstable angina to prevent Recurrent Events (CURE) study showed that carrier status of the loss-of-function *CYP2C19* allele did not differ in the safety and efficacy of clopidogrel.³⁷ These findings were also replicated in a subgroup of the Atrial Fibrillation Clopidogrel Trial with Irbesartan for Prevention of Vascular Events (ACTIVE) A trial. While patients in the CURE study were mostly non-invasively managed, another distinguishing feature of the analysis is the inclusion of the placebo group. The addition of the placebo group provides evidence of the efficacy of the experimental treatment. It also helps to reduce sources of confounding, such as potential pleiotropic genetic effects and population stratification. The results of this study have also been confirmed by a systematic review and meta-analysis consisting of 32 studies and 42,016 patients.³⁸ The authors reported a significant association between loss-of-function carrier status and risk of cardiac events using 'treatment-only' studies. However,

the authors failed to report a significant association when using 'effect-modification' studies or studies with more than 200 cardiovascular events. These analyses show the importance of using large RCTs with both placebo and drug arms to guide validated recommendations on pharmacogenetic findings and medication use.

Prospective cohort studies

Prospective cohort studies follow a group of participants who are self-selected into a drug treatment group and assess how genetic distribution corresponds to the risk of developing the study outcome. Prospective cohort studies are able to examine causality through the temporal affects of drug exposure and genetic variants on disease risk.³⁹ However, prospective cohort designs are expensive and time-consuming because they often require a large sample size to detect a relatively modest drug-gene interaction. Moreover, this study design is more subject to confounding because the assignment of drug therapy is subject driven rather than randomly allocated.

Selection bias occurs in prospective cohort studies if loss to follow-up is differential by drug exposure or by genotype. For example, loss to follow-up and drug use may vary by age, and loss to follow-up and genetic polymorphisms may vary by ethnic group. Furthermore, if individuals who were lost to follow-up tended to have different risks associated with the study outcome compared with those who remained for the entire length of the study, then the overall incidence estimates would be biased.

Prospective cohort studies are more subject to non-differential misclassification as compared with case-control studies. Non-differential misclassification occurs when exposure measurement errors are independent of the outcome and result in dilution of the measure of association and bias estimates towards the null. This may occur if drug use is not collected at multiple timepoints throughout the study. During the study, participants may begin a new medication or discontinue their current treatment because of adverse drug events. An increase in data collection over the study period will help to ensure improved accuracy of patient behaviour and improved measurements.

As mentioned previously, subgroups of participants from prospective cohorts can be analysed in

nested case-control studies. These studies select participants who experienced the study outcome, and compare them with randomly selected controls from the original study cohort. The advantages of using this design are that the cases are compared with the same comparison group, which helps to reduce bias and confounding. Furthermore, this design allows researchers to use small sample sizes and allows for a more cost-effective approach.

One such example in pharmacogenetics is the examination of the *CYP2D6**4 allele in tamoxifen-treated patients from the Rotterdam Study.⁴⁰ The *CYP2D6* gene is involved in the formation of endoxifen from tamoxifen, which is used for the treatment of oestrogen receptor-positive breast cancer in post-menopausal women.⁴¹ The objective of the study assessed the association between carriers of the *CYP2D6**4 allele and breast cancer mortality among all incident users of tamoxifen. The study reported that the hazard ratio of breast cancer mortality in patients with the *4/*4 genotype was 4.1 (95% CI 1.1–15.9; $P = 0.041$) compared with those with the wild-type genotype. Although these results are subject to potentially more bias as the exposed and unexposed groups were not randomized, there is greater generalizability in this study as compared with an RCT. This represents a trade-off between optimal internal validity in the RCT design compared with external applicability in the prospective cohort design. It would be wiser to report the more robust estimates of the RCT and use subsequent studies to explore the generalizability of these findings than rely on the estimates from a prospective cohort study.

Case-control studies

Case-control studies are the most common study design in pharmacogenetics. Under this model, cases are defined as those who have had a specific adverse drug event or a poor therapy outcome. The genetic variant frequencies in the cases are compared with the controls who have a comparable level of drug exposure but are also free of the study outcome. These studies are able to measure the effect of the gene-drug interaction but the independent effects of the genotype and drug response cannot be ascertained.

Case-control studies can be performed quickly and they are more cost-effective than large

prospective studies. Case-control studies may be the only feasible study design when it is not possible to conduct an RCT. For example, it may not be possible to use a prospective study design to assess rare adverse drug outcomes or rare variants because they require a very large sample size.⁴² Furthermore, it is unethical to conduct an RCT with *a priori* unequivocal knowledge of severe drug-gene interaction, in which carriers of a variant allele are known to be susceptible to adverse events.

The retrospective design of case-control studies makes it more prone to confounding, selection bias and information bias. Selection bias is the product of inappropriate choice of study controls and differential participation rates between cases and controls. Ideally, controls should represent cases with respect to potential exposures and have the same risk of developing the outcome phenotype. For example, pooled hospital-based controls may include participants whose allelic frequencies correspond to another underlying disease, which will distort the exposure-disease association. Selection bias may also result from differential non-participation among cases and controls if those who failed to participate were related to genotype or drug exposure.

Information bias in case-control studies is most likely to result from differential misclassification. Differential misclassification occurs when there is systematic error in the degree of misclassification between cases and controls, which will distort the true magnitude of association in any direction. One common type of information bias in case-control studies is recall bias. Recall bias occurs when cases remember past exposures differently than controls. For example, cases may recall past drug exposures better than those who did not experience the outcome because they have more motivation to identify possible causes of their disease. It is important to note that there is no recall bias with genetic exposure because participants' genotypes are fixed.

Genetic epidemiology considerations

Phenotype definition

In pharmacogenetic studies, the selection of the study endpoint and the patient response phenotype are crucial for interpreting drug efficacy.⁴³

However, since many pharmacogenetic studies use data from prospective studies, the study endpoints and patient population may not be precise enough to identify functional genes that are associated with the drug response. For example, it may be more appropriate to measure clinical outcomes, such as adverse bleeding events, when studying the association between safety measures and genetic markers. Nevertheless, physiological and biochemical measures may be more appropriate phenotypes to represent the underlying gene function in the drug–gene interaction,⁴⁴ such as platelet count or clotting time. These phenotypes represent stronger biological or causal evidence of the functional activity of the gene or protein in question.

However, across studies, there is great heterogeneity in the biological measurement and definitions of outcomes or phenotypes. For example, the reported prevalence of aspirin resistance ranges from 5% to 45%,⁴⁵ which is thought to result from small sample sizes and heterogeneity within the methodologies used to measure the biochemical and functional components of aspirin resistance.^{46,47} Goodman *et al.*⁴⁸ performed a systematic review of all the genetic studies of aspirin resistance, and observed that the effect of the *PIA1/PIA2* polymorphism in the GPIIIa receptor appears to differ according to the technique used to measure aspirin resistance. The lack of standardization among laboratory tests leads to imprecise effect estimates of the polymorphism and drug response. Therefore, to decrease heterogeneity among studies and for more reliable estimates of pharmacogenetic associations, there must be consistent and functionally relevant phenotypic definitions.

Genetic polymorphisms

The associated genetic variants are either directly functional or they are indirectly correlated with another variant that is the actual cause of the drug response. Linkage disequilibrium (LD) 'is the tendency for a pair of alleles at two linked loci to be associated with each other in the population more than would be expected by chance'.⁴⁹ LD is useful in genetic association studies because high LD allows for a smaller subset of marker SNPs to be genotyped while capturing most of the genetic information. However,

LD varies among ethnic populations and this may affect cross-subpopulation comparisons when causal SNPs are not directly genotyped but rather captured by 'proxy' SNPs.^{50,51}

Population stratification

A source of confounding within-population-based pharmacogenetic studies can result from population stratification.⁵² Population stratification occurs when ethnic subpopulations within the entire study population differ in terms of genotype frequency and risk of disease.⁵³ Population stratification confounds pharmacogenetic associations when differences in the prevalence of an allele parallels the incidence of study outcomes⁵² and may bias both the strength of the association and estimates of precision of the genetic variant–outcome association. In other words, clinical outcomes might vary among genetically distinct populations for reasons other than the variant being tested and thus bias pooled drug–gene interaction effects.⁵⁴ Stratification can also occur in apparently homogeneous populations, for example, Davey Smith *et al.*⁵⁵ observed an increasing north–south gradient in the frequency of the variant allele for lactase persistence across Britain.

One approach to minimizing the confounding effects of population stratification is to match participants based on geographical region and by markers of ethnic origin.⁵⁶ Stratifying the study sample by ethnic groups allows for fair comparisons among homogenous groups; however, depending on the amount of stratification, too many groups will decrease the power able to detect an effect within each stratum. Genetic principal components are also widely used to minimize confounding by stratification. This method corrects for spurious associations in traits that differ among populations and have different allelic frequencies for the genotype of interest. Most differences in allelic frequencies are thought to have occurred because of genetic drift and may not represent functional variants.⁵⁷ Thus, the principal component technique is used to detect and correct for the population heterogeneity to minimize false-positive associations.⁵⁸ Variance component methods have also been recently developed to adjust for population stratification.⁵⁹ Importantly, randomized studies are immune to this bias since equal numbers of individuals of each population strata

will be randomized to the drug of interest or placebo group.

Genetic pleiotropy

Genetic pleiotropy is the phenomenon in which a single gene is responsible for a number of distinct and seemingly unrelated phenotypic traits.⁶⁰ This phenomenon is of special importance to pharmacogenetics because it may confound the pharmacogenetic association. For instance, if the gene of interest is associated with multiple outcomes or intermediate phenotypes, the reported drug–gene interaction may be a result of the underlying gene mechanism and not a product of the drug response.⁶¹ For example, the SH2B3 gene has been associated with multiple phenotypic traits, such as blood pressure,^{62,63} blood eosinophil number,⁶⁴ myocardial infarction,⁶⁴ celiac disease,⁶⁵ type I diabetes,⁶⁶ LDLc,⁶⁷ asthma,⁶⁴ blood platelet number,⁶⁸ haemoglobin concentration⁶⁹ and haematocrit.⁶⁹

Several large trials have observed that lowering LDLc levels decreases the risk of atherosclerosis events, which can be achieved through statin therapy.⁷⁰ The proprotein convertase subtilisin/kexin type 9 (*PCSK9*) gene degrades the LDL receptor, which helps to increase the clearance of LDLc from circulation. Gain-of-function carriers of the *PCSK9* genotype are associated with mild to severe hypercholesterolaemia, while loss-of-function carriers are associated with decreased LDLc and decreased risk of cardiovascular events.^{71,72} The loss-of-function carriers are also associated with more pronounced decrease in LDLc with statin therapy,⁷³ and it is difficult to determine to which extent the observed relationship is driven by a pharmacogenetic effect or by the gene effect. Therefore, to distinguish if there is an independent relationship and true effect modification, it is essential to use an RCT design with a control group to see if the effect occurs in the treatment group alone.

Statistical issues in pharmacogenetics

A major issue in pharmacogenetics is the lack of replication among population-based studies. Possible explanations for the sometime

inconsistent findings are modest effect sizes, small sample sizes and multiple hypothesis testing. In this section, we will discuss sample size and multiple testing issues, and how to address them.

Sample size

The ability to determine whether there is a clinically significant difference between groups is dependent on the study sample size. Pharmacogenetic studies must be large in order to have enough statistical power to detect a gene effect, a treatment effect and a drug–gene interaction.⁷⁴ The power to detect a statistical interaction depends on the number of SNPs, the allelic frequencies of each SNP and the type of study design.⁷⁵ It is unlikely that a common genetic variant will have a large effect in a complex trait, such as drug response.⁷⁶ Studies should thus be powered to detect a common or rare variant with a modest or very large effect size, respectively.^{76,77}

Table 2 shows the approximate sample sizes needed to detect a significant gene–drug interaction (assuming 80% power and $\alpha = 0.05$) by effect size and allelic frequency (among controls). Under these conditions, it is assumed that the genetic variant is causal; however, it is possible that the variant allele is in LD with the actual causal variant, which may require a larger sample size.^{79,80} If a rare genetic variant is anticipated with a small or modest effect, a sample size of more than 900,000 participants would be required. However, if a common variant with a large effect was expected, then a sample size of approximately 900 participants is needed. These results suggest that the majority of pharmacogenetic studies are underpowered, which may give rise to false-negative or false-positive estimates.

For some pharmacogenetic questions, the required sample sizes may be difficult to obtain. The need for large datasets has led to the creation of international consortia where data between investigators are pooled or analysed together, or large population-based biobanks, which store biological materials (i.e. blood or DNA) and demographic information, including drug use.⁸¹ In addition, RCTs now incorporate genetic add-on studies which have the same high internal and external validity and large sample size of the parent RCT, while remaining cost-effective.⁸²

Table 2**Sample size required to detect a drug–gene interaction in a pharmacogenetic study based on minor allele frequency***

Prevalence of variant allele carriers among controls	Odds ratio*					
	1.10	1.20	1.30	1.40	1.50	2.00
0.01	906,222	240,782	113,415	67,443	45,528	14,464
0.05	189,628	50,434	23,785	14,164	9577	3072
0.10	100,570	26,779	12,647	7543	5109	1657
0.15	71,330	19,013	8991	5370	3644	1193
0.20	57,106	15,237	7213	4314	2931	968

*Sample sizes have been calculated based on a drug–gene interaction assuming an additive genetic model. These estimates assume a type-I error rate of 0.05, a power of 80% and a baseline risk of an adverse drug reaction among exposed subjects to be 10%. Sample sizes were calculated using QUANTO⁷⁸

Multiple testing

Multiple testing refers to the repeated use of a statistical test and the risk of an overall type I error.⁸³ Multiple testing arises when there are multiple comparisons in statistical models that contain multiple genes, multiple exposures and multiple interactions.⁸⁴ Within these models, it is inappropriate to use the standard *P* value of 0.05 because as the number of tests increases so does the frequency of type I errors.

The most common approach to correct for multiple testing is to use the Bonferroni correction, in which the *P* value that is used for one test is divided by the total number of tests in the analysis. However, the use of the Bonferroni correction may be considered too conservative because many SNPs are in LD, which may mask their effects and increase type II errors. Furthermore, since many of the pharmacogenetic studies are underpowered to detect a drug–gene interaction, the Bonferroni correction may null the study results.⁸⁵ Another possible approach to adjust for multiple testing is to use the false discovery rate (FDR), which is less conservative than the Bonferroni correction.⁸⁶ The FDR estimates the expected proportion of false-positives among associations that are declared significant, which is expressed as a *q* value.

Conclusions

Pharmacogenetic studies offer a promising future yet have a challenging present. Personalized

medicine has the potential to maximize drug efficacy and minimize the toxic effects; however, there are many issues in study design and analysis that need to be addressed. Large collaborative efforts across biostatisticians, epidemiologists, pharmacologists and clinicians are needed to provide robust evidence to support individualized treatment for improved drug efficacy and safety.

References

- 1 Eichler HG, Abadie E, Breckenridge A, *et al.* Bridging the efficacy-effectiveness gap: a regulator's perspective on addressing variability of drug response. *Nat Rev Drug Discov* 2011;**10**:495–506
- 2 Senn S. Individual response to treatment: is it a valid assumption. *BMJ* 2004;**329**:966–8
- 3 Lemesle G, Delhaye C, Bonello L, de LA, Waksman R, Pichard A. Stent thrombosis in 2008: definition, predictors, prognosis and treatment. *Arch Cardiovasc Dis* 2008;**101**:769–77
- 4 Kalow W, Tang BK, Endrenyi L. Hypothesis: comparisons of inter- and intra-individual variations can substitute for twin studies in drug research. *Pharmacogenetics* 1998;**8**:283–9
- 5 Weinshilboum R. Inheritance and drug response. *N Engl J Med* 2003;**348**:529–37
- 6 Ioannidis JP, Bernstein J, Boffetta P, *et al.* A network of investigator networks in human genome epidemiology. *Am J Epidemiol* 2005;**162**:302–4
- 7 Ioannidis JP, Gwinn M, Little J, *et al.* A road map for efficient and reliable human genome epidemiology. *Nat Genet* 2006;**38**:3–5
- 8 Khoury MJ, Little J, Higgins J, Ioannidis JP, Gwinn M. Reporting of systematic reviews: the challenge of genetic association studies. *PLoS Med* 2007;**4**:e211
- 9 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:e7960
- 10 Manolopoulos VG, Ragia G, Tavridou A. Pharmacogenetics of coumarinic oral anticoagulants. *Pharmacogenomics* 2010;11:493–6
 - 11 Higashi MK, Veenstra DL, Kondo LM, et al. Association between CYP2C9 genetic variants and anticoagulation-related outcomes during warfarin therapy. *JAMA* 2002;287:1690–8
 - 12 Aithal GP, Day CP, Kesteven PJ, Daly AK. Association of polymorphisms in the cytochrome P450 CYP2C9 with warfarin dose requirement and risk of bleeding complications. *Lancet* 1999;353:717–9
 - 13 Limdi NA, McGwin G, Goldstein JA, et al. Influence of CYP2C9 and VKORC1 1173C/T genotype on the risk of hemorrhagic complications in African-American and European-American patients on warfarin. *Clin Pharmacol Ther* 2008;83:312–21
 - 14 Johnson JA, Gong L, Whirl-Carrillo M, et al. Clinical pharmacogenetics implementation consortium guidelines for CYP2C9 and VKORC1 genotypes and warfarin dosing. *Clin Pharmacol Ther* 2011;90:625–9
 - 15 Klein TE, Altman RB, Eriksson N, et al. Estimation of the warfarin dose with clinical and pharmacogenetic data. *N Engl J Med* 2009;360:753–64
 - 16 Epstein RS, Moyer TP, Aubert RE, et al. Warfarin genotyping reduces hospitalization rates results from the MM-WES (Medco-Mayo Warfarin Effectiveness study). *J Am Coll Cardiol* 2010;55:2804–12
 - 17 Feld JJ, Hoofnagle JH. Mechanism of action of interferon and ribavirin in treatment of hepatitis C. *Nature* 2005;436:967–72
 - 18 Ge D, Fellay J, Thompson AJ, et al. Genetic variation in IL28B predicts hepatitis C treatment-induced viral clearance. *Nature* 2009;461:399–401
 - 19 Suppiah V, Gaudieri S, Armstrong NJ, et al. IL28B, HLA-C, and KIR variants additively predict response to therapy in chronic hepatitis C virus infection in a European Cohort: a cross-sectional study. *PLoS Med* 2011;8:e1001092
 - 20 Chung WK. Implementation of genetics to personalize medicine. *Gen Med* 2007;4:248–65
 - 21 Hetherington S, McGuirk S, Powell G, et al. Hypersensitivity reactions during therapy with the nucleoside reverse transcriptase inhibitor abacavir. *Clin Ther* 2001;23:1603–14
 - 22 Clay PG. The abacavir hypersensitivity reaction: a review. *Clin Ther* 2002;24:1502–14
 - 23 Mallal S, Phillips E, Carosi G, et al. HLA-B*5701 screening for hypersensitivity to abacavir. *N Engl J Med* 2008;358:568–79
 - 24 Gazzard BG, Anderson J, Babiker A, et al. British HIV Association Guidelines for the treatment of HIV-1-infected adults with antiretroviral therapy 2008. *HIV Med* 2008;9:563–608
 - 25 Hammer SM, Eron JJ Jr, Reiss P, et al. Antiretroviral treatment of adult HIV infection: 2008 recommendations of the International AIDS Society-USA panel. *JAMA* 2008;300:555–70
 - 26 Schackman BR, Scott CA, Walensky RP, Losina E, Freedberg KA, Sax PE. The cost-effectiveness of HLA-B*5701 genetic screening to guide initial antiretroviral therapy for HIV. *AIDS* 2008;22:2025–33
 - 27 Hughes DA, Vilar FJ, Ward CC, Alfirevic A, Park BK, Pirmohamed M. Cost-effectiveness analysis of HLA B*5701 genotyping in preventing abacavir hypersensitivity. *Pharmacogenetics* 2004;14:335–42
 - 28 Kauf TL, Farkouh RA, Earnshaw SR, Watson ME, Maroudas P, Chambers MG. Economic efficiency of genetic screening to inform the use of abacavir sulfate in the treatment of HIV. *Pharmacoeconomics* 2010;28:1025–39
 - 29 French B, Joo J, Geller NL, et al. Statistical design of personalized medicine interventions: the Clarification of Optimal Anticoagulation through Genetics (COAG) trial. *Trials* 2010;11:108
 - 30 Voora D, Shah SH, Reed CR, et al. Pharmacogenetic predictors of statin-mediated low-density lipoprotein cholesterol reduction and dose response. *Circ Cardiovasc Genet* 2008;1:100–6
 - 31 Voora D, Shah SH, Spasojevic I, et al. The SLCO1B1*5 genetic variant is associated with statin-induced side effects. *J Am Coll Cardiol* 2009;54:1609–16
 - 32 Connolly SJ, Pogue J, Hart RG, et al. Effect of clopidogrel added to aspirin in patients with atrial fibrillation. *N Engl J Med* 2009;360:2066–78
 - 33 Yusuf S, Zhao F, Mehta SR, Chrolavicius S, Tognoni G, Fox KK. Effects of clopidogrel in addition to aspirin in patients with acute coronary syndromes without ST-segment elevation. *N Engl J Med* 2001;345:494–502
 - 34 Shuldiner AR, O'Connell JR, Bliden KP, et al. Association of cytochrome P450 2C19 genotype with the antiplatelet effect and clinical efficacy of clopidogrel therapy. *JAMA* 2009;302:849–57
 - 35 Mega JL, Close SL, Wiviott SD, et al. Cytochrome p-450 polymorphisms and response to clopidogrel. *N Engl J Med* 2009;360:354–62
 - 36 Daly AK, King BP. Pharmacogenetics of oral anticoagulants. *Pharmacogenetics* 2003;13:247–52
 - 37 Pare G, Mehta SR, Yusuf S, et al. Effects of CYP2C19 genotype on outcomes of clopidogrel treatment. *N Engl J Med* 2010;363:1704–14
 - 38 Holmes MV, Perel P, Shah T, Hingorani AD, Casas JP. CYP2C19 genotype, clopidogrel metabolism, platelet function, and cardiovascular events: a systematic review and meta-analysis. *JAMA* 2011;306:2704–14
 - 39 Hill AB. The environment and disease: association or causation. *Proc R Soc Med* 1965;58:295–300
 - 40 Bijl MJ, van Schaik RH, Lammers LA, et al. The CYP2D6*4 polymorphism affects breast cancer survival in tamoxifen users. *Breast Cancer Res Treat* 2009;118:125–30
 - 41 Desta Z, Ward BA, Soukhova NV, Flockhart DA. Comprehensive evaluation of tamoxifen sequential biotransformation by the human cytochrome P450 system *in vitro*: prominent roles for CYP3A and CYP2D6. *J Pharmacol Exp Ther* 2004;310:1062–75
 - 42 Amstutz U, Carleton BC. Pharmacogenetic testing: time for clinical practice guidelines. *Clin Pharmacol Ther* 2011;89:924–7
 - 43 Bromley CM, Close S, Cohen N, et al. Designing pharmacogenetic projects in industry: practical design perspectives from the Industry Pharmacogenomics Working Group. *Pharmacogenomics J* 2009;9:14–22
 - 44 Dempfle A, Scherag A, Hein R, Beckmann L, Chang-Claude J, Schafer H. Gene-environment interactions for complex traits: definitions, methodological requirements and challenges. *Eur J Hum Genet* 2008;16:1164–72

- 45 Gum PA, Kottke-Marchant K, Poggio ED, *et al.* Profile and prevalence of aspirin resistance in patients with cardiovascular disease. *Am J Cardiol* 2001;**88**:230–5
- 46 Eikelboom JW, Hirsh J, Weitz JI, Johnston M, Yi Q, Yusuf S. Aspirin-resistant thromboxane biosynthesis and the risk of myocardial infarction, stroke, or cardiovascular death in patients at high risk for cardiovascular events. *Circulation* 2002;**105**:1650–5
- 47 Hankey GJ, Eikelboom JW. Aspirin resistance. *Lancet* 2006;**367**:606–17
- 48 Goodman T, Ferro A, Sharma P. Pharmacogenetics of aspirin resistance: a comprehensive systematic review. *Br J Clin Pharmacol* 2008;**66**:222–32
- 49 Thomas DC. *Statistical Methods in Genetic Epidemiology*. New York: Oxford University Press, 2004
- 50 Hirschhorn JN, Lohmueller K, Byrne E, Hirschhorn K. A comprehensive review of genetic association studies. *Genet Med* 2002;**4**:45–61
- 51 Ardlie KG, Kruglyak L, Seielstad M. Patterns of linkage disequilibrium in the human genome. *Nat Rev Genet* 2002;**3**:299–309
- 52 Peters BJ, Rodin AS, de BA, Maitland-van der Zee AH. Methodological and statistical issues in pharmacogenomics. *J Pharm Pharmacol* 2010;**62**:161–6
- 53 Thomas DC, Witte JS. Point: population stratification: a problem for case-control studies of candidate-gene associations. *Cancer Epidemiol Biomarkers Prev* 2002;**11**:505–12
- 54 Umbach DM. Invited commentary: on studying the joint effects of candidate genes and exposures. *Am J Epidemiol* 2000;**152**:701–3
- 55 Smith GD, Lawlor DA, Timpson NJ, *et al.* Lactase persistence-related genetic variant: population substructure and health outcomes. *Eur J Hum Genet* 2009;**17**:357–67
- 56 Cardon LR, Palmer LJ. Population stratification and spurious allelic association. *Lancet* 2003;**361**:598–604
- 57 Pare G. Genome-wide association studies – data generation, storage, interpretation, and bioinformatics. *J Cardiovasc Transl Res* 2010;**3**:183–8
- 58 Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D. Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet* 2006;**38**:904–9
- 59 Kang HM, Sul JH, Service SK, *et al.* Variance component model to account for sample structure in genome-wide association studies. *Nat Genet* 2010;**42**:348–54
- 60 Lobo I. Pleiotropy: one gene can affect multiple traits. *Nature Education* 2008;**1**(1). See <http://www.nature.com/scitable/topicpage/pleiotropy-one-gene-can-affect-multiple-traits-569> (last checked 4 January 2012)
- 61 Lewis SJ. Mendelian randomization as applied to coronary heart disease, including recent advances incorporating new technology. *Circ Cardiovasc Genet* 2010;**3**:109–17
- 62 Newton-Cheh C, Johnson T, Gateva V, *et al.* Genome-wide association study identifies eight loci associated with blood pressure. *Nat Genet* 2009;**41**:666–76
- 63 Levy D, Ehret GB, Rice K, *et al.* Genome-wide association study of blood pressure and hypertension. *Nat Genet* 2009;**41**:677–87
- 64 Gudbjartsson DF, Bjornsdottir US, Halapi E, *et al.* Sequence variants affecting eosinophil numbers associate with asthma and myocardial infarction. *Nat Genet* 2009;**41**:342–7
- 65 Hunt KA, Zhernakova A, Turner G, *et al.* Newly identified genetic risk variants for celiac disease related to the immune response. *Nat Genet* 2008;**40**:395–402
- 66 Todd JA, Walker NM, Cooper JD, *et al.* Robust associations of four new chromosome regions from genome-wide analyses of type 1 diabetes. *Nat Genet* 2007;**39**:857–64
- 67 Talmud PJ, Drenos F, Shah S, *et al.* Gene-centric association signals for lipids and apolipoproteins identified via the HumanCVD BeadChip. *Am J Hum Genet* 2009;**85**:628–42
- 68 Soranzo N, Spector TD, Mangino M, *et al.* A genome-wide meta-analysis identifies 22 loci associated with eight hematological parameters in the HaemGen consortium. *Nat Genet* 2009;**41**:1182–90
- 69 Ganesh SK, Zakai NA, van Rooij FJ, *et al.* Multiple loci influence erythrocyte phenotypes in the CHARGE Consortium. *Nat Genet* 2009;**41**:1191–8
- 70 Baigent C, Keech A, Kearney PM, *et al.* Efficacy and safety of cholesterol-lowering treatment: prospective meta-analysis of data from 90,056 participants in 14 randomised trials of statins. *Lancet* 2005;**366**:1267–78
- 71 Abifadel M, Varret M, Rabes JP, *et al.* Mutations in PCSK9 cause autosomal dominant hypercholesterolemia. *Nat Genet* 2003;**34**:154–6
- 72 Peterson AS, Fong LG, Young SG. PCSK9 function and physiology. *J Lipid Res* 2008;**49**:1152–6
- 73 Cohen JC, Boerwinkle E, Mosley TH Jr, Hobbs HH. Sequence variations in PCSK9, low LDL, and protection against coronary heart disease. *N Engl J Med* 2006;**354**:1264–72
- 74 Kelly PJ, Stallard N, Whittaker JC. Statistical design and analysis of pharmacogenetic trials. *Stat Med* 2005;**24**:1495–508
- 75 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**:503–10
- 76 Bansal V, Libiger O, Torkamani A, Schork NJ. Statistical analysis strategies for association studies involving rare variants. *Nat Rev Genet* 2010;**11**:773–85
- 77 Davey SG, Ebrahim S, Lewis S, Hansell AL, Palmer LJ, Burton PR. Genetic epidemiology and public health: hope, hype, and future prospects. *Lancet* 2005;**366**:1484–98
- 78 Gauderman W, Morrison M. *QUANTO 1.1: A Computer Program for Power and Sample Size Calculations for Genetic-Epidemiology Studies*. 2006. See <http://hydra.usc.edu/gxe/> (last checked 4 January 2012)
- 79 Singer C, Grossman I, Avidan N, Beckmann JS, Pe'er I. Trick or treat: the effect of placebo on the power of pharmacogenetic association studies. *Hum Genomics* 2005;**2**:28–38
- 80 Elston RC, Idury RM, Cardon LR, Lichter JB. The study of candidate genes in drug trials: sample size considerations. *Stat Med* 1999;**18**:741–51
- 81 Hattersley AT, McCarthy MI. What makes a good genetic association study. *Lancet* 2005;**366**:1315–23
- 82 Kirchheiner J, Fuhr U, Brockmoller J. Pharmacogenetics-based therapeutic recommendations – ready for clinical practice. *Nat Rev Drug Discov* 2005;**4**:639–47

- 83 Cobos A, Sanchez P, Aguado J, Carrasco JL. Methodological quality in pharmacogenetic studies with binary assessment of treatment response: a review. *Pharmacogenet Genomics* 2011;**21**:243–50
- 84 Hunter DJ. Gene-environment interactions in human diseases. *Nat Rev Genet* 2005;**6**:287–98
- 85 Little J, Burke W. *Human Genome Epidemiology: A Scientific Foundation for Using Genetic Information to Improve Health Prevent Disease*. Oxford: Oxford University Press, 2004
- 86 Storey JD, Tibshirani R. Statistical significance for genomewide studies. *Proc Natl Acad Sci USA* 2003;**100**:9440–5

© 2012 Royal Society of Medicine Press

This is an open-access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by-nc/2.0/>), which permits non-commercial use, distribution and reproduction in any medium, provided the original work is properly cited.