



Pharmacogenetics: practices and opportunities for study design and data analysis

Aiden A. Flynn

Exploristics Limited, Cromac Square, Belfast BT2 8LA, United Kingdom

Pharmacogenetics (PGx) is increasingly used as a way to target treatment to patients who are most likely to benefit. To date, PGx has shown clinical significance across a few applications but widespread use has been limited by the need for further technical, methodological and practical advances and for educating clinical researchers on the value of PGx.

Here, I describe the current scope of PGx research, including recent contributions to prospective study design. A case study is included to demonstrate the limitations of current practice and to describe some practical steps for improving the chances of identifying genetic effects. Additionally, I describe some opportunities for the integration and application of disparate data sources in exploratory PGx research.

The role of PGx

The application of PGx within clinical development programmes has experienced unprecedented growth recently [1]. Over the past 15 years, the development and commercialisation of new drugs has become increasingly expensive, with diminishing probability of success [2,3]. In 2004, the US Food and Drug Administration (FDA) launched The Critical Path Initiative (CPI) to promote innovative approaches, such as PGx, to overcome stagnation in drug development (US Food and Drug Administration; <http://www.fda.gov/ScienceResearch/SpecialTopics/CriticalPathInitiative/CriticalPathOpportunitiesReports/ucm077262.htm>). Medical practitioners and healthcare providers have since recognised the importance of PGx in directing treatment to those more likely to benefit and are increasingly using PGx to guide decision making [4–6].

PGx offers additional options in clinical research and development by helping to explain unexpected variability in safety and efficacy outcomes. Previously, this variability might have led to a termination of the development programme but PGx enables the continued research of the effects of treatment within a subgroup of patients. Indeed, there are now several examples where PGx has made an important contribution to the clinical development of drugs to patient populations [7,8].

Although PGx has proved to be extremely valuable for improving the benefits and reducing the risks of some drugs in development and for the targeting of more-effective treatment to patients, the successes remain limited. To some extent, this is owing to the fact that PGx is often used as a tool to re-evaluate development plans when study outcomes are negative or ambiguous rather than being an integral part of an informed personalised medicine plan or strategy. In addition, PGx remains an evolving science that requires further technological and methodological development to optimise the analysis of the data and design of the studies. Indeed, the right combination of prospective planning, data, design and analysis is likely to yield many more successes for PGx.

PGx data analysis

The main objective for PGx analysis is to identify and/or characterise genetic effects. Although there are too many methods to review adequately in this article, there are some general principles that are broadly followed in basic analyses. These include the evaluation and selection of genetic markers that can distinguish between groups of patients who differ in their response to treatment (Fig. 1). The evaluation involves the application of a statistical model comprising the factors that are thought to contribute to the observed variability in response. These models can include two types of effects: main effects where factors make a sole contribution to the observed variability; and interaction effects where

E-mail address: aiden.flynn@exploristics.com.

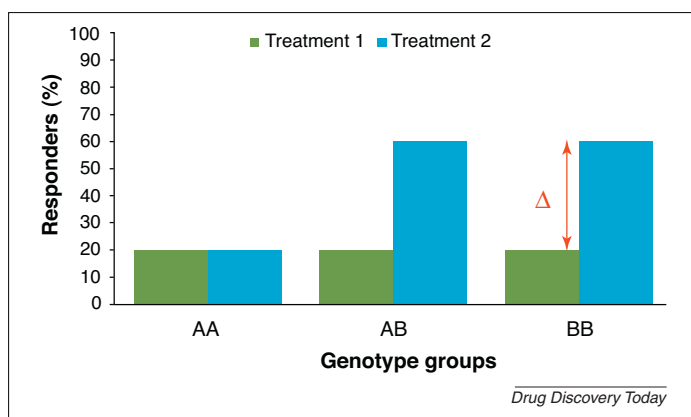


FIGURE 1

Graph illustrating an example of a genetic marker to distinguish between groups of patients. In this case, the percentage of responders in the groups of patients on treatment 2 and with the AB or BB genotype is greater compared with the other groups. In other words, the difference (Δ) between treatments 1 and 2 is greater in AB and BB patients.

two or more factors make a combined contribution. The choice of model is important because it will determine the value and application of selected markers. Models that include genotype as a main effect are useful for identifying markers that are associated with response, regardless of treatment (prognostic markers), whereas markers that are associated with response in the presence of treatment (predictive markers) can be identified using models where there is an interaction between genotype and treatment. Genetic markers are then selected on the basis of their prognostic or predictive utility.

PGx study design

Good study designs ensure a high probability of meeting the study objectives whilst controlling for known sources or variability and bias. In PGx, the study design options depend on how and when PGx is applied [9]. Typically, the early stages of PGx research is exploratory whereby many genes are investigated, often using data that were collected as part of a study designed for another purpose. This is followed by confirmatory research where PGx is the primary objective in a prospectively designed study. To date, most methodological research into study designs for PGx has focused on the prospective, confirmatory applications. To some extent, this focus on confirmatory study design reflects the regulatory preference for prospectively designed studies in all but exceptional cases.

Confirmatory studies

The primary objective of confirmatory studies is to test a hypothesis relating to pre-specified genetic effects. These studies are prospectively designed and involve markers with substantive prior evidence of their function and usefulness in explaining variability in patient response. Several study designs have been proposed and evaluated for prognostic and predictive markers [10–13]. Three common designs that are used in a confirmatory setting for predictive markers are described below.

The targeted or enriched design (Fig. 2a) involves a pre-screening step whereby patients are selected for the study based on genotype. Patients carrying the negative genotype are excluded from the study whereas positive patients are then randomised to

one of the treatment groups. This design has advantages in that it can result in smaller studies when the effect of treatment is greater in the positive group. However, this design does not provide information on treatment effect in the excluded population and can only be used when there is substantive prior knowledge about the impact of a single genetic marker.

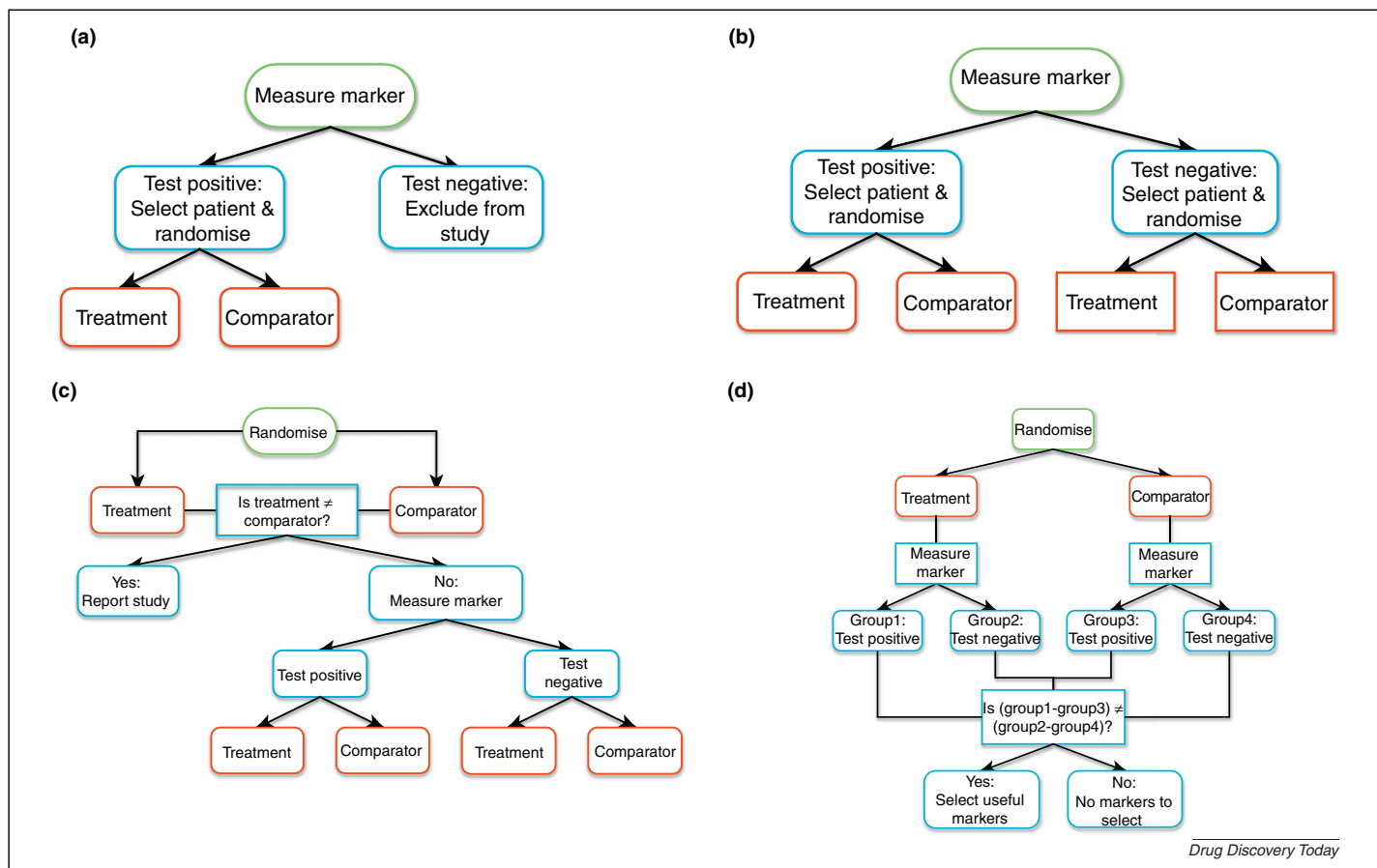
A less restrictive design is the stratification design (Fig. 2b). As with the targeted design, this involves a pre-screening step where all subjects are allocated to groups based on genotype, then randomised to treatment. This design is useful when there is interest in collecting information on the treatment effect in the marker negative group. In addition, this enables the operating characteristics, such as sensitivity and specificity, of a genetic-based test to be estimated. However, this design also requires considerable prior knowledge about the genetic marker.

Both of the above designs are useful when the studies are required to test a single hypothesis relating to the genetic marker. However, it is common for studies to have multiple objectives involving the evaluation of a treatment effect in the entire study population and in sub-populations. A popular approach for this type of study is based on an adaptive design (Fig. 2c) [13]. In this case, enrolled patients are randomised to treatment groups and the treatments are compared as part of a primary objective of the study. If the study fails to meet the primary objective (i.e. there is no difference in the treatments) then patients are sub-divided into groups based on genotype and a comparison of treatments is performed within these groups. Because multiple statistical tests are performed in this approach, the overall false-positive error rate is greater and needs to be controlled. This design can result in larger studies but it is more flexible than the targeted or stratified design because it enables the testing of multiple objectives and can be modified to include the evaluation of several genetic markers.

Exploratory studies

In the exploratory setting, PGx is typically used for identifying potentially useful genetic markers and for generating hypotheses that can be tested in future studies. Depending on the level of prior knowledge of the genetic influence on response, exploratory studies might involve the evaluation of a small number of candidate genes to a complete search of the whole genome [9].

The most common approach to exploratory PGx studies involves the collection of blood samples within a study with the view to use in PGx research if required [14]. In the event that a PGx study is initiated, the study population is retrospectively and iteratively stratified using the genotypes from all markers under investigation (Fig. 2d). This approach is flexible in the sense that it does not compete with the original study objective and enables the evaluation of many markers. However, the limited sample size and need to control the false-positive detection rate severely restricts the ability to detect genetic effects. Furthermore, the lack of randomisation in treatment within the genotype groups introduces bias and imbalance into the genetically defined groups [15]. As a result, the benefit of markers identified from retrospective analysis is limited [16] and there is generally the need to collate further data from prospectively designed studies. Indeed, these issues illustrate why many stakeholders prefer prospectively designed studies.

**FIGURE 2**

Examples of study designs for confirmatory and exploratory pharmacogenetics: Enriched design (a); stratification design (b); adaptive design (c); retrospective design (d).

For these reasons, the retrospective and exploratory approach to PGx has had limited success with the notable exception of a few drugs associated with large genetic effects, such as abacavir [17] and panitumumab [18]. This represents one of the key challenges in PGx research where the lack of sufficient patient data to detect moderate genetic effects makes it difficult to identify markers with clinical utility. Because most PGx research continues to be exploratory, retrospective approaches will continue to have an important role. However, there is a need to improve the probability of success of exploratory PGx studies through the development and application of better design and analysis methods and more consideration of prospective integration into development programmes.

Exploratory PGx Case Study

This case study is used to demonstrate the limitations of exploratory PGx and to illustrate some simple and pragmatic approaches to improve the power in this setting. This case study is based on a Phase IIb study of Alzheimer's disease (AD), comprising 511 subjects randomised into four treatment groups receiving placebo or 2 mg, 4 mg or 8 mg of Rosiglitazone [8]. The subjects were equally divided into each of the groups resulting in 128 subjects per group. The study was powered at 95% to detect a three-point difference between active treatment and placebo in the change from baseline of the Alzheimer's Disease Assessment Scale-cognitive subscale (ADAS-COG), assuming a standard deviation of six units and a 5% level of significance. This study failed to meet its primary

objective. Blood samples were collected for 65% of the enrolled patients and a retrospective PGx analysis was then performed to identify potential predictive markers.

To continue with this illustration, further assumptions were made to assess the likelihood of detecting a predictive marker in this case. Firstly, it was assumed that a clinically useful genetic marker would contain a genetically defined group where the difference between active treatment and placebo was three units and a group where there was no difference. In addition, it was assumed that the case study included the evaluation of 1000 markers with allele frequencies ranging from 10% to 90%. Finally, it was assumed that the pattern of responses reflects a dominant genetic effect and a treatment effect that is observed across all active treatment groups. On the basis of these assumptions, study data were simulated 1000 times. Each simulation was analysed using analysis of covariance with change from baseline ADAS-COG as the dependent variable and genotype, treatment and genotype by treatment as independent variables. The interaction effect was significant at the 5% level. The probability of success for the PGx study was given by the number of significant interactions over all simulations. Six scenarios that describe different study design and analysis options were compared. These were based on the assumptions described above but with the following modifications:

- (i) Original design (i.e. no modification).
- (ii) Original design, significance level set to 0.005% by adjusting for the number of tests (i.e. 5%/1000).

- (iii) Original design is altered with a sample collection rate of 100%.
- (iv) As in (iii), but original design altered further by reducing the number of treatment groups from four to three.
- (v) As in (iv) but analysis uses a non-linear model that accounts for the genetic and dose–response pattern.
- (vi) Original design is altered by increasing the sample size to 1200.

Figure 3 shows the probability of success for the six scenarios. All scenarios show a reduction in the probability of success for high and low allele frequencies reflecting the increased uncertainty around the estimates of response in small genotype groups. The maximum probability of success for the first scenario is just over 20%. By adjusting for multiple testing, as in the second scenario, the likelihood of identifying a clinically useful predictive marker is almost reduced to zero. By increasing the sample collection in the third scenario, the probability of success almost doubles compared with the first scenario. The probability of success can be further increased by a reduction in the number of treatment groups whilst maintaining the total number of subjects (scenario four). Furthermore, the use of a non-linear model that provides a better description of the data can boost the power even further as in scenario five. Finally, an increase in the number of subjects recruited into the study would enhance power to >80% for some of the allele frequencies.

There are several important messages from this case study. Simply collecting samples and taking an entirely retrospective approach to PGx has limitations. Indeed, some prospective planning can enable studies to be designed for another purpose and yet have a marked impact on the likelihood of success for PGx. In addition, the first five scenarios involve the same number of subjects, demonstrating that it is possible to improve the success rate without necessarily increasing cost. Finally, the combination of study design and the analysis method impact the probability of

success. In every exploratory PGx experiment these should be selected to optimise the probability of meeting all the study objectives.

Concluding remarks

PGx offers much promise for the development of safer and more-effective treatments and for presenting additional options within clinical research programmes. To date, the successes have involved the identification and application of genetic markers that have a large effect on treatment response. By contrast, the use of genetic markers with small to moderate effects has proved to be much more challenging. To a large extent, the lack of success is owing to the fact that PGx is an emerging science and researchers are still developing methods to integrate and analyse data from the new sequencing technologies.

There are several ways to improve the success rate of PGx within clinical development programmes. Two important areas that can facilitate the identification and integration of genetic markers are improved study design and data analysis methods. Indeed, there has been a lot of activity in the development of specific study designs that integrate genetic markers with known or well studied effects. These designs have been described in this article and enable the application of PGx in studies where there is a single PGx-related objective and where PGx is one of several objectives. However, there remains the challenge in exploratory PGx of identifying genetic markers in a study that was not specifically designed for PGx research.

The case study presented in this article shows that it is possible to improve the success rate of exploratory studies through better alignment of the study design, analysis methods and study objectives. The case study also describes some pragmatic steps that would enhance the likelihood of success for exploratory PGx without adding to the cost substantially. Indeed, exploratory PGx strategies should go beyond the collection of samples by including prospective planning of design and analysis so that PGx use is optimised.

In future, PGx will be greatly enhanced by improving the success rate of exploratory studies. There are several areas that would help to increase the power to identify useful predictive markers. One obvious approach is to increase the number of available samples by boosting study sizes or by improving access to phenotypic data and tissue samples from other sources. In addition, there is an abundance of relevant knowledge and information that is captured on numerous, disparate sources including biorepositories, biological pathway information, gene ontologies and published literature. This information remains underused when evaluating and interpreting the usefulness of genetic markers; better integration would have a marked impact on the success of exploratory PGx. Furthermore, there is enormous scope for developing and applying statistical models that reflect more closely the underlying biology and patterns of response; using models that describe the data better will increase the power to detect genetic markers. The development of these models will require extensive methodological research and development for the integration and application of disparate data sources. This is likely to involve a multidisciplinary approach in the form of collaboration between clinicians, biologists, geneticists, statisticians, bioinformaticians and computer scientists.

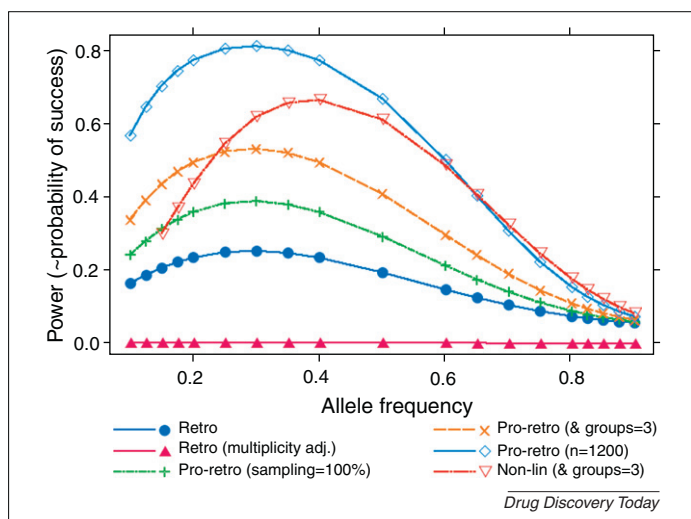


FIGURE 3

Power curves for the six scenarios in the case study. The first two curves represent the entirely retrospective approach to PGx analysis (scenario one and two). The following three curves involve some prospective planning (scenarios three, four and six) and the last curve involves the use of non-linear models to reflect the dose and genetic response pattern (scenario five).

In summary, PGx has undoubtedly improved the efficacy and safety of products on the market and has shown value in clinical development. Indeed, PGx warrants further significant research, particularly in the development of methods with application in the exploratory setting. These methods will result in a marked

improvement in the use of PGx and will enable PGx and personalised healthcare to realise their full potential. Only then will clinical researchers, regulators, healthcare providers and patients understand the true power of PGx.

References

- 1 Jørgensen, J. and Winther, H. (2009) The new era of personalized medicine: 10 years later. *Pers. Med.* 6, 423–428
- 2 Lesko, L. and Woodcock, J. (2004) Translation of pharmacogenomics and pharmacogenetics: a regulatory perspective. *Nat. Rev. Drug Discov.* 3, 763–769
- 3 Woodcock, J. (2008) The FDA CPI and its influence on new drug development. *Annu. Rev. Med.* 59, 1–12
- 4 Burns, D.K. (2010) Developing pharmacogenetic evidence throughout clinical development. *Clin. Pharmacol. Ther.* 88, 867–870
- 5 Frueh, F.W. *et al.* (2008) Pharmacogenomic biomarker information in drug labels approved by the United States food and drug administration: prevalence of related drug use. *Pharmacotherapy* 28, 992–999
- 6 Hughes, A.R. *et al.* (2004) Association of genetic variations in HLA-B region with hypersensitivity to abacavir in some, but not all, populations. *Pharmacogenomics* 5, 203–211
- 7 Risner, M. *et al.* (2006) Efficacy of rosiglitazone in a genetically defined population with mild-to-moderate Alzheimer's disease. *Pharmacogenomics J.* 6, 246–254
- 8 Young, K.Y. *et al.* (2010) The efficiency of clinical trials designs for predictive biomarker validation. *Clin. Trials* 7, 557–566
- 9 Bromley, C.M. *et al.* (2009) Designing pharmacogenetic projects in industry: practical design perspectives from the industry pharmacogenomics working group. *Pharmacogenomics J.* 9, 14–22
- 10 Mandrekar, S.J. and Sargent, D.J. (2009) Clinical trial designs for predictive biomarker validation: one size does not fit all. *J. Biopharm. Stat.* 19, 530–542
- 11 Simon, R. (2010) Clinical trial designs for evaluating the medical utility of prognostic and predictive biomarkers in oncology. *Pers. Med.* 7, 33–47
- 12 Freidlin, B. *et al.* (2010) Randomized clinical trials with biomarkers: design issues. *J. Natl. Cancer Inst.* 102, 152–160
- 13 Jiang, W. *et al.* (2007) Biomarker-adaptive threshold design: a procedure for evaluating treatment with possible biomarker-defined subset effect. *J. Natl. Cancer Inst.* 99, 1036–1043
- 14 Goodsaid, F.M. *et al.* (2010) Voluntary exploratory data submissions to the US FDA and the EMA: experience and impact. *Nat. Rev. Drug Discov.* 9, 435–445
- 15 Wang, S.J. *et al.* (2010) Statistical Considerations in evaluating pharmacogenomics-based clinical effect for confirmatory trials. *Clin. Trials* 7, 525–536
- 16 Burns, D. *et al.* (2010) Designing pharmacogenomic studies to be fit for purpose. *Pharmacogenomics* 11, 1657–1667
- 17 Hughes, A.R. *et al.* (2009) Genetic association studies to detect adverse drug reactions: abacavir hypersensitivity as an example. *Pharmacogenomics* 10, 225–233
- 18 Weber, J. and McCormick, P. (2008) Panitumumab in metastatic colorectal cancer with wild-type KRAS. *BioDrugs* 22, 403–411