diseases — information that could be misused. However, the risk that is associated with such information is much less with AE prediction than with disclosing the results of genetic disease diagnoses[33]. A PGx prediction could lead to the suggestion that the patient should not be treated with a new medication, presumably after the disease is already diagnosed. The result would be that a few more patients might not be treated with a predictably inappropriate medicine. Moreover, the screening SNP panels for PGx might be designed to exclude most known disease-associated SNPs (those that underlie variation in both the disease and the response to drugs that are administered to treat it are the notable exceptions). It could also be possible to design a separate panel for genetic disease screening, but it would be distinct from a standardized SNP panel for PGx[46]. The ethical considerations would clearly support the early screening of individuals before filling the prescription to avoid a second wave of patients with AEs[57,63–65].

The prohibitive cost of 100,000 SNP assays, when measured by yesterday's price is daunting, just as chromosome sequencing was a decade ago. However, today, there are commercially available prototype chips for 100,000 SNPs, and the cost of SNP assays is decreasing over time, again analogous to the DNA-sequencing experience. For a pharmaceutical company to perform a screen on 30 patients who experience an AE, at a penny an assay, the total cost would be US $30,000; at US $100 per chip, it would be US $3,000. When millions of people use the same standardized panels, the cost would decrease considerably. Comparison with data on large control ethnic cohorts placed into the public domain that had also been screened with the standardized SNP panel would be essentially free[7]. Given that it costs US $1 billion to develop a successful new drug, the technology would certainly be used.

The purpose of this review was to introduce current realities of genome-wide PGx to the broader genetic community that is generally more concerned with disease genes, animal models of diseases or broader issues of cross-species gene mapping. The constrained linear nature of the genome sequence provides the opportunity for detailed accurate mapping with relatively few individuals[9,10]. The usual random associations and the statistical methods that have been commonly used need to be re-assessed in light of the power of current methodologies. A new age for the treatment of diseases with safer and more targeted medicines is beginning.

1. Cardon, L. R. & Bell, J. I. Association study designs for complex diseases. *Nature Rev. Genet.* **2**, 91–99 (2001).
   An informative review of the application of genetic association methods used to identify genes associated with complex diseases.
2. Tabor, H. K., Risch, N. J. & Myers, R. M. Candidate-gene approaches for studying complex genetic traits: practical considerations. *Nature Rev. Genet.* **3**, 391–397 (2002).
3. Weiss, K. M. & Clark, A. G. Linkage disequilibrium and the mapping of complex human traits. *Trends Genet.* **18**, 19–24 (2002).
   An overview of the applications of LD mapping for diseases, but equally applicable to safety (adverse-event) PGx.
4. Goldstein, D. B., Tate, S. K. & Sisodiya, S. M. Pharmacogenetics goes genomic. *Nature Rev. Genet.* **4**, 937–947 (2003).
5. Carlson, C. S. *et al.* Selecting a maximally informative set of single-nucleotide polymorphisms for association analyses using linkage disequilibrium. *Am. J. Hum. Genet.* **74**, 106–120 (2004).
6. Dawson, E. *et al.* A first-generation linkage disequilibrium map of human chromosome 22. *Nature* **418**, 544–548 (2002).
7. Lai, E. *et al.* Medical applications of haplotype-based SNP maps: learning to walk before we run. *Nature Genet.* **32**, 353 (2002).
   Comment on the difference between a strategy of whole-genome mapping, limited to a full panel of small, well-defined LD blocks, versus the use of an initial screening set to determine regions of extended LD for fine mapping (for examples see references 10, 16 and 47).
8. Roses, A. D. Pharmacogenetics. *Hum. Mol. Genet.* **10**, 2261–2267 (2001).
9. Roses, A. D. Genome-based pharmacogenetics and the pharmaceutical industry. *Nature Rev. Drug Discov.* **1**, 541–554 (2002).
10. Xu, C. F. *et al.* Identification of a pharmacogenetic effect by linkage disequilibrium mapping. *Pharmacogenomics J.* (in the press).
    A proof of principle analysis of the application of whole-genome LD to map a locus (loci) associated with an adverse effect of a medicine.
11. Food and Drug Administration. *Innovation or Stagnation: Challenge and Opportunity on the Critical Path to New Medical Products* [online], <http://www.fda.gov/oc/initiatives/criticalpath/whitepaper.html> (2004).
12 Lindsay, M. A. Target discovery. *Nature Rev. Drug Discov.* **2**, 831–838 (2003).
13. Chanda, S. K. & Caldwell, J. S. Fulfilling the promise: drug discovery in the post-genomic era. *Drug Discov. Today* **8**, 168–174 (2003).
14. Collins, F. S. Genetics: an explosion of knowledge is transforming clinical practice. *Geriatrics* **54**, 410–417 (1999).
15. Subramanian, G., Adams, M. D., Venter, J. C. & Broder, S. Implications of the human genome for understanding human biology and medicine. *JAMA* **286**, 2296–2307 (2001).
16. McCarthy, L. C. *et al.* Single-nucleotide polymorphism alleles in the insulin receptor gene are associated with typical migraine. *Genomics* **78**, 135–149 (2001).
    An early example of the use of high density SNP mapping in a region of putative genetic linkage, in this case for the identification of specific SNP variants of the insulin receptor gene that are associated with migraine.
17. Hakonarson, H., Gulcher, J. R. & Stefansson, K. deCODE genetics, Inc. *Pharmacogenomics* **4**, 209–215 (2003).
18. Kristjansson, K. *et al.* Linkage of essential hypertension to chromosome 18q. *Hypertension* **39**, 1044–1049 (2002).
19. Nathan, D. G. Clinical research: perceptions, reality, and proposed solutions. National Institutes of Health Director's Panel on Clinical Research. *JAMA* **280**, 1427–1431 (1998).
20. Boston Consulting Group. *A Revolution in R&D: How Genomics and Genetics Will Affect Drug Development Costs and Times* in Parexel's Pharmaceutical R&D Statistical Sourcebook, 2003/2003 (Parexel International, Waltham, USA, 2003).
21. Harris, S. & Foord, S. M. Transgenic gene knock-outs: functional genomics and therapeutic target selection. *Pharmacogenomics* **1**, 433–443 (2000).
22. Brenner, S. E. Target selection for structural genomics. *Nature Struct. Biol.* **7** (Suppl.), 967–969 (2000).
23. DeFife, K. M. & Wong-Staal, F. Integrated approaches to therapeutic target gene discovery. *Curr. Opin. Drug Discov. Dev.* **5**, 683–689 (2002).
24. Searls, D. B. Pharmacophylogenomics: genes, evolution and drug targets. *Nature Rev. Drug Discov.* **2**, 613–623 (2003).
    An analysis of the importance of placing genomic variation in an evolutionary context, taking advantage of phylogenetic relationships and comparisons, to better understand function and in particular selective pressures.
25. Debouck, C. & Metcalf, B. The impact of genomics on drug discovery. *Annu. Rev. Pharmacol. Toxicol.* **40**, 193–207 (2000).
26. Drews, J. Drug discovery: a historical perspective. *Science* **287**, 1960–1964 (2000).
27. Thorisson, G. A. & Stein L. D. The SNP Consortium website: past, present and future. *Nucleic Acids Res.* **31**, 124–127 (2003).
28. Vaschetto, M., Weissbrod, T., Bodle, D. & Guner, O. Enabling high-throughput discovery. *Curr. Opin. Drug Discov. Dev.* **6**, 377–383 (2003).
29. Duyk, J. Attrition and translation. *Science* **302**, 603–605 (2003).
30. Twyman, R. M. & Primrose, S. B. Techniques patents for SNP genotyping. *Pharmacogenomics* **4**, 67–79 (2003).
31. Renegar, G., Rieser, P. & Manasco, P. Family consent and the pursuit of better medicines through genetics research. *J. Contin. Educ. Health Prof.* **21**, 265–270 (2001).
32. Horrobin, D. F. Modern biomedical research: an internally self-consistent universe with little contact with medical reality? *Nature Rev. Drug Discov.* **2**, 151–154 (2003).
33. Roses, A. D. High through-put disease-associated target identification. *Drug Discov. Today* (in the press).
    Description of high throughput multi-gene target testing to determine variants that are highly associated with human disease.
34. Lynch, T. J. *et al.* Activating mutations in the epidermal growth factor receptor underlying responsiveness of non-small-cell lung cancer to gefitinib. *N. Engl. J. Med.* **350**, 2129–2139 (2004).
35. Paez, J. G. *et al.* *EGFR* mutations in lung cancer: correlation with clinical response to gefitinib therapy. *Science* **304**, 1497–1500 (2004).
36. Dos Santos, C. *et al.* A common polymorphism of the growth hormone receptor is associated with increased responsiveness to growth hormone. *Nature Genet.* **36**, 720-724 (2004).
37. Tantisira, K. G. *et al.* Corticosteroid pharmacogenetics: association of sequence variants in *CRHR1* with improved lung function in asthmatics treated with inhaled corticosteroids. *Hum. Mol. Genet.* **13**, 1353–1359 (2004).
38. Sesti, G. *et al.* The Arg^{972} variant in insulin receptor substrate-1 is associated with an increased risk of secondary failure to sulfonylurea in patients with type 2 diabetes. *Diabetes Care* **27**, 1394–1398 (2004).
39. Macedo, A., Farre, M. & Banos, J. E. Placebo effect and placebos: what are we talking about? Some conceptual and historical considerations. *Eur. J. Clin. Pharmacol.* **59**, 337–342 (2003).
40. Goetz, C. G., Janko, K., Blasucci, L. & Jaglin, J. A. Impact of placebo assignment in clinical trials of Parkinson's disease. *Mov. Disord.* **18**, 1146–1149 (2003).
41. Anonymous. *The Fruits of Genomics: Drug Pipelines Face Indigestion Until the New Biology Ripens.* (Lehman Brothers, McKinsey and Co., New York, 2001).

©2004 Nature Publishing Group

42. Tufts Center for the Study of Drug Development. *Backgrounder: How New Drugs Move Through the Development and Approval Process* <http://csdd.tufts.edu/NewsEvents?RecentNews.asp?newsid=4> (2001).
43. Gilbert, J., Henkse, P. & Singh, A. Rebuilding big pharma's business model. *In vivo*, the business and medicine report. Windhover Information, 21, No. 10, (2003).
44. Noble M. E., Endicott, J. A. & Johnson, L. N. Protein kinase inhibitors: insights into drug design from structure. *Science* **303**, 1800–1805 (2004).
45. Vogel C. L. & Franco S. X. Clinical experience with trastuzumab (herceptin). *Breast J.* **9**, 452–462 (2003).
46. Goldstein, D. B., Ahmadi, K. R., Weale, M. E. & Wood NW. Genome scans and candidate gene approaches in the study of common diseases and variable drug responses. *Trends Genet.* **19**, 615–622 (2003).
47. Lai, E., Riley, J., Purvis, I. & Roses, A. A 4-Mb high-density single nucleotide polymorphism-based map around human APOE. *Genomics* **54**, 31–38 (1998).
An early example of LD mapping of a region of genetic linkage to demonstrate the polymorphisms of SLC12A8, a gene coding for a member of the solute carrier family 12 proteins that are associated with familial psoriasis.
48. Hewett, D. *et al.* Identification of a psoriasis susceptibility candidate gene by linkage disequilibrium mapping with a localized single nucleotide polymorphism map. *Genomics* **79**, 305–314 (2002).
49. deCODE Genetics. deCODE and Merck & Co., Inc. form broad drug development alliance. deCODE News [online], <http://www.decode.com/main/view.jsp?branch=5026&e342RecordID=1901&e342DataStoreID=3917> (2004).
50. Merck Press Release. Drug development: pharmaceutical company joins drug development alliance *Genomics & Genetics Weekly* **38**, (2004).
51. Melzer, D., Detmer, D. & Simmern, R. Pharmacogenetics and public policy: expert views in Europe and North America. *Pharmacogenomics* **4**, 689–691 (2003).
52. Lee S. S. Race, distributive justice and the promise of pharmacogenomics: ethical considerations. *Am. J. Pharmacogenomics* **3**, 385–392 (2003).
53. Hapgood, R. The potential and limitations of personalized medicine in the doctor-patient relationship. *Pharmacogenomics* **4**, 685–687 (2003).
54. Drazen, J. M. *et al.* Pharmacogenetic association between ALOEX5 promotor genotype and the response to anti-asthma treatment. *Nature Genet.* **22**, 168–170 (2003).
55. Wechsler, M. E. & Israel, E. Pharmacogenetics of treatment with leukotriene modifiers. *Curr. Opin. Allergy Clin. Immunol.* **2**, 395–401 (2002).
56. Spear, B. B., Heath-Chiozzi, M. & Huff, J. Clinical application of pharmacogenetics. *Trends Mol. Med.* **7**, 201–204 (2001).
57. Roses A. D. 2025: the practice of neurology: back from the future. *Arch. Neurol.* **58**, 1766–1767 (2001).
58. Ahmadi, K. R. *et al.* A single nucleotide polymorphism tagging set for human drug metabolism and transport. *Science* (in the press).
A study and analysis of the use of LD mapping to determine a tagged set of SNPs for 56 human drug metabolizing enzymes.
59. Griffiths, J. D., Stark, R. J., Ding, J. C. & Cooper, I. A. Vincristine neurotoxicity in Charcot–Marie–Tooth syndrome. *Med. J. Aust.* **143**, 305–306 (1985).
60. Danoff, T. M. *et al.* A Gilbert's syndrome UGT1A1 variant confers susceptibility to tranilast-induced hyperbilirubinemia. *Pharmacogenomics J.* **4**, 49–53 (2004).
61. Gibbs, R. A. *et al.* The International HapMap Project. *Nature* **426**, 789–796 (2003).
62. Bowman, C. Classification using SNP profiles. <http://www.luc.ac.be/censtat/RSS2003> (2003).
63. Norbert P. W. & Roses A. D. Pharmacogenetics and pharmacogenomics: recent developments, their clinical relevance and some ethical, social, and legal implications. *J. Mol. Med.* **81**, 35–40 (2003).
64. Lesko L. J. *et al.* Pharmacogenetics and pharmacogenomics in drug development and regulatory decision making: report of the first FDA-PWG-PhRMA-DruSafe Workshop. *J. Clin. Pharmacol.* **43**, 342–358 (2003).
A view from the FDA, the agency that regulates pharmaceutical licenses in the US, on decision making using pharmacogenetics and pharmacogenomics during drug development, including definition of key terms generically referred to as "biomakers."
65. Buchanan, A. *et al.* Pharmacogenetics: ethical issues and policy options. *Kennedy Inst. Ethics J.* **12**, 1–15 (2002).
66. Welsh-Bohmer K. A. *et al.* Apolipoprotein E genotypes in a neuropathological series from the Consortium to Establish a Registry for Alzheimer's Disease. *Ann. Neurol.* **42**, 319-325 (1997).

Acknowledgements

This research was fully supported from the sales of medicines. I especially thank my successive Presidents of Research and Development, both physician-scientists, James Niedel and Tachi Yamada, for having the foresight and commitment to support and apply strategic genetic research in 1997 when human linkage research was considered as peaked, and since 2000 when medical research investments for the future were still unfashionable for pharmaceutical R&D organizations. The ongoing pharmacogenetics work that is described in this review has been contributed to by hundreds of GlaxoSmithKline (GSK) employees over a seven year period. I also thank the Genetic Executive Teams, who have shared a vision of safer and more effective medicines, the thousands of patients who consented to participate in pharmacogenetic research protocols, Ronald Krall who provided access to all patients participating in GSK drug development, my wife, Ann Saunders, who provided intellectual, emotional and laboratory support, and my daughters — Maija, Stephanie and Joanna — who created a place to live, to work and to love.

Competing interests statement

The author declares competing financial interests: see Web version for details.

Online links

DATABASES

The following terms in this article are linked online to:
Entrez: http://www.ncbi.nih.gov/Entrez/
CYP2D6 | *UGT1A1*

FURTHER INFORMATION

Food and Drug Administration Science Board meeting (22 April 2004):
http://www.fda.gov/ohrms/dockets/ac/04/transcripts/4039t1.htm
Food and Drug Administration Science Board presentation (22 April 2004):
http://www.fda.gov/ohrms/dockets/ac/04/slides/4039s1.htm
International HapMap Project: http://www.hapmap.org/
Access to this links box is available online.

©2004 Nature Publishing Group

PERSPECTIVES

OPINION

Translation of pharmacogenomics and pharmacogenetics: a regulatory perspective

Lawrence J. Lesko and Janet Woodcock

Abstract | Pharmacogenomics and pharmacogenetics provide methodologies that can lead to DNA-based tests to improve drug selection, identify optimal dosing, maximize drug efficacy or minimize the risk of toxicity. Rapid advances in basic research have identified many opportunities for the development of 'personalized' treatments for individuals and/or subsets of patients defined by genetic and/or genomic tests. However, the integration of these tests into routine clinical practice remains a major multidisciplinary challenge, and even for well-established biomarkers there has been little progress. Here, we consider this challenge from a regulatory perspective, highlighting recent initiatives from the FDA that aim to facilitate the integration of pharmacogenetics and pharmacogenomics into drug development and clinical practice.

The promise of pharmacogenomics (PGx) lies in its potential to identify sources of inter-individual variability in drug response that affect drug efficacy and drug safety. The identification of PGx BIOMARKERS (see Glossary) can lead to the development of PGx tests that can be used to individualize therapy with the intent of maximizing effectiveness, minimizing risks and optimizing doses in therapeutic applications.

In this article, our definition of PGx is very broad, and includes the study of inter-individual variations in whole-genome or candidate gene SINGLE-NUCLEOTIDE POLYMORPHISM (SNP) MAPS, HAPLOTYPE markers and alterations in gene expression or inactivation that might be correlated with pharmacological function and therapeutic response. Pharmacogenetics (PGt), by contrast, is narrower in definition and refers to the study of inter-individual variations in DNA sequence related to drug absorption and disposition (pharmacokinetics) or drug action (pharmacodynamics), including polymorphic variation in genes that encode transporters, drug-metabolizing enzymes, receptors and other proteins. We will not consider proteomics in this article, although gene-driven proteomic patterns in serum ('protein signatures') show promise, for example, as prognostic or screening biomarkers for staging cancer or for identifying high-risk subgroups in a disease population. We acknowledge that there is overlap between the definitions of PGx and PGt, and we will use the terms 'pharmacogenomic test' or 'pharmacogenetic test' to refer to an assay to study these inter-individual variations in conjunction with drug therapy.

Translating PGx from bench to bedside (or from discovery to marketability) is a multidisciplinary problem that involves addressing philosophical, societal, cultural, behavioural and educational differences between the private and public sector, as well as issues unique to drug development, extent of scientific expertise, interdisciplinary communication and clinical practice. However, we will focus on a regulatory science perspective of PGx and PGt that will cover three broad areas: first, the views of the Food and Drug Administration regarding the value and challenges of integrating PGx and PGt into the continuum of drug research and development and regulatory decision making; second, the major, structured approach that the FDA has undertaken to encourage the use of PGx and PGt both in drug development and clinical practice; and third, selected examples of how PGx and PGt have been used both in new drug development and in updating the labels of approved drugs. Within the context of these three areas, we will point out various challenges that drug developers, regulatory agencies, health-care providers and others will have to address in order to attain the benefits of PGx and PGt more fully.

Drug R&D: what is the problem?

By and large, drug development and private- and public-sector research has been reasonably successful during the past 15–20 years, and there is, in fact, much to celebrate. However, as indicated by analysis and metrics provided by regulatory agencies in the United States and Europe, we are now facing a major challenge: it is essential to improve the success of pharmaceutical research and development (R&D). Although the productivity of drug discovery and early development has increased over the years (as measured by the upward trend in the identification of new molecules, drug targets and INVESTIGATIONAL NEW DRUG APPLICATIONS (INDs) filed with the FDA), the number of major drug and biological product NEW DRUG APPLICATIONS (NDAs) and BIOLOGIC LICENSE APPLICATIONS (BLAs) for new molecular entities that have been submitted to the FDA has steadily decreased during this period. The pharmaceutical industry submitted almost 50% fewer applications to the FDA in 2002–2003 than it did in 1996–1997 (FIG. 1). During the same timeframe, investment in biomedical research spending for the private and public sectors increased almost 2.5-fold (FIG. 2). So, it is clear that many biomedical discoveries have not been transformed into marketable products in the United States and worldwide.

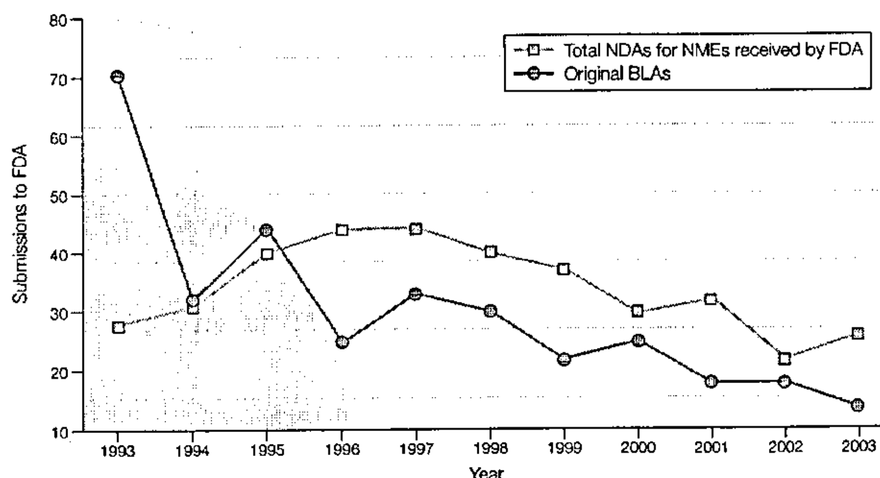

Figure 1 | **Ten-year trends in major drug and biological product submissions to the FDA.** Based on data from REF. 6.

The existing model of drug development in the pharmaceutical industry faces daunting challenges. More than 80% of potential products that enter the development pipeline with the filing of an IND fail to make it to market because of fatal flaws in one or more of the three dimensions of product development: first, drug safety (a high incidence of adverse events or unexpected toxicity); second, drug efficacy (no strong signal of effectiveness over placebo and/or active comparator); and third, industrialization (the product cannot be manufactured at a commercial scale with consistently high quality). Furthermore, it has been estimated that to develop a single, successful new chemical entity now costs in excess of US $800 million[1] (a figure that includes 'opportunity costs'), and the average time taken to do so is 8–10 years. The clinical component of the overall cost of new drug development is ~58% or US $400 million[1]. A significant proportion of these dollars goes towards supporting the Phase III randomized controlled trials (RCT) that provide the most convincing evidence of the safety and efficacy of a drug product. However, from a recent report, one can estimate the failure rate in Phase III trials to be ~50%[2].

Variability in drug response

Variability in drug response is a major barrier to successful drug development. As Sir William Osler said in 1892 about the practice of medicine, "If it were not for the great variability among individuals, medicine might as well be a science and not an art". PGx and PGt provide the scientific tools that enable us to explore the pathophysiological mechanisms underlying these differences in drug response at the molecular level. We expect that there will be an increase in the public demand for more science and less art in the search for better and more effective therapies to reduce the morbidity and mortality of chronic diseases such as hypertension and cancer. In order to improve the 'art' and the productivity of the drug development process, PGx and PGt can improve the predictability of preclinical safety studies, and clinical safety and efficacy trials.

A key to increasing R&D success is identifying drugs that are likely to either succeed or fail late in the process early in the drug development process and thereby reduce attrition in Phase III trials — that is, before the high costs of these trials are incurred by a sponsor. This is an important achievement because the average size of a Phase III clinical trial has nearly tripled in the past 20 years. It does not make much sense to wait until a Phase III trial fails to try to establish why the drug did not provide evidence of efficacy or lack of toxicity, and how to design the next trial; that approach is expensive and time consuming.

It is typical that each Phase III trial is preceded by a much longer preclinical and early clinical work-up of the drug, so what is needed is an increased ability to predict Phase III success or failure, aimed at the preclinical and early clinical time period. For example, in terms of cost, a 10% improvement in predicting failure before large-scale Phase III clinical trials begin could save ~US $100 million in development costs. Other opportunities for saving US $12–21 million dollars in direct development costs can be attained by shifting just 5% of clinical failures from Phase III to Phase I, or by shifting 25% of failures from Phase II to Phase I[3].

The major causes of attrition of drugs in late-phase clinical trials are lack of efficacy or concerns about safety. To achieve increases in productivity and success, effective scientific development tools, such as those provided by PGx and PGt, are needed to predict product performance — whether it be success or failure — with a high degree of certainty, and this needs to occur both early and reliably in the development process. For example, PGx biomarkers can be used to identify potential responders. By stratifying patients by biomarker status in Phase II clinical trials, populations with a high probability of responding can be identified, thereby simplifying Phase III trials and increasing their probability of success.

Clearly, modern innovative tools are needed to predict the performance and manufacturing quality of twenty-first century products. Although it seems that everyone agrees with this premise, the problem is that the drug development process is no longer able to keep pace with the rate and scope of discoveries in basic science. For example, although imaging-based biomarkers are presently being used to develop drugs for Alzheimer's disease, there has not been a successful strategy for correlating anatomical imaging with primary clinical endpoints of cognition and function to enable the identification of new drug candidates that can modify disease progression. The tools currently used in drug discovery and development — the so-called 'critical path' tools — have not incorporated either the latest advances in biomarker technologies (with links to clinical outcomes), the basic and information sciences (such as the new knowledge and technologies provided by the rapid development of genomic research), or innovations in clinical sciences (for example, adaptive trial designs) to substantially affect the success of drug development and improve the quality of public health. Although the reasons underlying the failures of drugs in development (especially those failing in late-phase clinical trials), and inefficiencies in the development process in general, are not well understood, many suspect that a lack of understanding of variability in drug response between patients is a key part of the problem. Recent and rapidly accumulating evidence is beginning to point toward genetic and genomic factors, alone and taken together with environmental factors, as being of considerable importance in determining inter-individual variability in drug responses.

An example of the power of PGx is evident from recent publications regarding gefitinib (Iressa; AstraZeneca)[4,5]. Gefitinib is one member of a new class of targeted cancer therapies that inhibit the tyrosine kinase activity of the epidermal growth factor receptor (EGFR), which is important in many cancers. Gefitinib

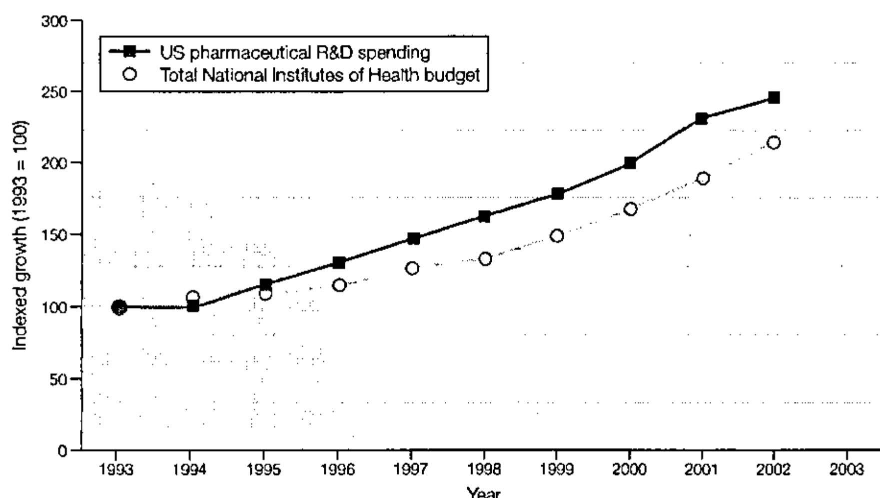

Figure 2 | **Ten-year trends in biomedical research spending.** Source: Parexel's Pharmaceutical R&D Statistical sourcebook 2002/2003.

was approved by the FDA for advanced non-small-cell lung cancer (NSCLC) in May 2003. The overall response rate at approval, as measured by significant tumour shrinkage, was less than optimal and occurred in only about 10% of patients who were administered gefitinib. However, clinician reports indicated that the drug works rapidly and amazingly well in some patients. In addition, higher response rates were noted in Japanese subjects, women and patients with adenocarcinoma.

PGx provides a molecular explanation for why gefitinib is so much more effective in some patients, whereas others seem insensitive to it. For example, Lynch *et al.* identified somatic mutations in the tyrosine kinase domain of the *EGFR* gene in eight out of nine patients with lung cancer characterized as 'responders', and in none of seven patients who had no response. A genomic approach to identifying responder subsets would clearly be advantageous given the potential safety consequences (for example, interstitial lung disease) in patients who have a small chance of benefiting from gefitinib treatment. Screening for these mutations in lung-cancer patients earlier could possibly identify responders and facilitate earlier treatment and thereby reduce disease progression.

If such findings are generalizable, they could markedly improve development of further EGFR tyrosine kinase inhibitors. For example, future clinical trials for such drugs intended to treat NSCLC might include screening for EGFR mutations in Phase II (hypothesis-generating) trials to identify patients who would then have a greater likelihood of a beneficial response. This drug development strategy would lead to Phase III trials enriched with patients with EGFR mutations that would have a higher probability of successfully demonstrating efficacy. This approach could reduce the risk of treatment failure, and could decrease the size and cost of subsequent Phase III trials, thereby bringing greater efficiency to the development process. At a minimum, this type of genomic information could also help understand the drug better by identifying the root cause of variability in responsiveness. This is not to suggest that a new drug should not be tested in patients who test negative for the mutation of interest, unless it is obvious that the drug could not work in this group. If a pharmacogenomic test is not intended to be available in clinical practice to direct drug treatment to the patients demonstrated to be responders based on a mutation, then data in the subgroup that tests negative will be needed to assess the benefit/risk ratio in the overall population during the drug development process.

The Critical Path

The FDA released a white paper on 16 March 2004 entitled *Innovation or Stagnation? Challenge and Opportunity on the Critical Path to New Medical Products*[6]. This white paper is a serious attempt by the FDA to bring attention and focus to the need for targeted scientific efforts to modernize the tools, techniques and methods used to evaluate the safety, efficacy and quality of drug products. It describes the urgent need to build bridges among constituencies such as the FDA, the National Institutes of Health and the private sector to modernize the development process for medical products — the Critical Path — to make product development more predictable and less costly.

The critical path is defined as the path from candidate selection to product launch and it defines the potential bottlenecks in bringing a product to market. The focus of the critical path initiative is to identify ways to update the product development infrastructure for drugs, biologics and devices, and the evaluative tools currently used to assess the safety and efficacy of new medical products. Examples of evaluative tools include: better pathophysiological cell and/or animal disease state models for preclinical screening of new molecules; new and innovative scientific approaches, such as the use of BAYESIAN STATISTICS; the use and verification of pathophysiological and/or descriptive biomarkers for patient selection for clinical trials and/or use as surrogate endpoints; the use of modelling and computer simulation to design clinical trials and/or predict failures of medical devices; and improvement in processes for post-market reporting of adverse events related to implanted devices. In addition, an important example of a scientific opportunity for improving the critical path is the use of PGx and PGt, or, more specifically, the identification of DNA-based biomarkers or RNA-expression profiles that can provide insights into the stage of a disease, disease progression, drug response and drug-dosing requirements, and thereby lead to the development of tests to predict clinical outcomes more reliably.

The FDA's aim to advance PGx

The FDA's mission includes protecting and advancing public health, and encouraging innovations that make medicines and foods more effective, safer and more affordable. Beginning in earnest in June 2001, the FDA took the lead with several key initiatives in PGx and PGt that are intended to stimulate the use of PGx and PGt technologies in drug development, and to foster improvements in drug product safety and efficacy. After publication of a forward-looking paper that provides a regulatory perspective on the opportunities and challenges of integrating pharmacogenomics into drug development and regulatory decision-making[7], the FDA has coordinated its efforts with the pharmaceutical and biotechnology industries to convene a series of public PGx and PGt workshops. These workshops are a structured effort to bring together stakeholders from industry and academia with FDA scientists to openly discuss the status of PGx and PGt technology, the use of PGx and PGt in drug development and therapeutics, and the specific strategies that are most needed for using PGx and PGt as a tool to facilitate more efficient and effective research along the critical path of drug development.

Publications of the proceedings of these workshops are valuable references that describe the current status of PGx and PGt in drug development, and what is needed to continue to advance this critical path tool[8–10].

New drug development. The culmination of many individual efforts within the FDA, and the public input derived from the synergistic FDA–industry co-sponsored workshops, led to a significant milestone in the advancement of PGx: the November 2003 publication of the *Draft Guidance for Industry: Pharmacogenomic Data Submission* (see link to document in further information). This guidance was timely, in that there was considerable uncertainty and fear about what the FDA would do with exploratory genomic data obtained during the new drug development process, a fear that was a stumbling block for many pharmaceutical companies. The major concern was that the FDA would overreact to non-validated, exploratory genomic biomarkers, take them out of context, misinterpret them, cause delays in drug development, request additional clinical trials and/or put clinical trials on hold. This concern led to a reluctance of the industry to introduce genomic studies into their drug development plans.

The FDA wanted to break down these real or perceived barriers and motivate drug developers to consider PGx and PGt strategies seriously in their drug development portfolios. The PGx data guidance proposed a new pathway for industry and others for submitting non-clinical and exploratory clinical genomic data during the IND period without it undergoing formal regulatory review, and describes the submission format and regulatory review of such data by the Interdisciplinary PGx Review Group (IPRG). It introduced some new concepts related to genomic biomarkers and defined categories of biomarkers; that is, exploratory biomarkers, valid biomarkers, probable valid biomarkers and known valid biomarkers. By design, the guidance shied away from presenting very specific recommendations for biomarker validation and formats for submitting genomic data to avoid hindering progress in the field — the FDA recognized that the science is still evolving.

Important components of the guidance are three decision algorithms or decision trees based on the categories of biomarkers and the stage of drug development. Generally, most genomic data submitted to the FDA to date has been exploratory and not suitable for regulatory decision making. Such data — for example, those derived from gene-expression microarrays — have either no clear pathophysiological correlates, and/or are not crucial for entering patients into clinical trials or supporting claims about safety, efficacy and/or dosing. Valid biomarkers are defined as those biomarkers measured in an analytical test system with well-established performance characteristics and with an established scientific framework or body of evidence that explains the physiological, pharmacological, toxicological or clinical significance of the test results. Known valid biomarkers are those broadly accepted in the scientific community, whereas probable valid biomarkers are those that seem to have predictive value for clinical outcomes, but which have not yet been widely accepted or independently replicated. The decision trees can be used to determine when genomic data can be submitted voluntarily, and when submissions of the data are required by FDA regulations. In addition, the guidance describes the format (for example, full report, abbreviated report, synopsis or voluntary submission report) for submitting such data.

An example of one of the decision trees from the guidance that illustrates the process for submitting PGx data to an IND, as either a required submission or as a voluntary genomic data submission (VGDS), is shown in FIG. 3. It should be noted that the process by which industry submits VGDSs to the FDA uses the existing path for IND (or as a pre-IND in some cases) or NDA submissions, which assures the sponsor of the confidentiality of their data.

The FDA hopes that voluntary submissions will benefit both the industry and the Agency, and will provide a rational scientific basis for future data standards and genomic policies. Information and knowledge gained from voluntary submissions will be shared publicly across submissions in a way that protects the proprietary interests of companies. The FDA is currently in the process of finalizing the *Draft Guidance* on pharmacogenomic data submissions, and is writing two other internal documents that will describe the process for sponsors submitting VGDS and the roles and responsibilities of the IPRG.

Improving approved drugs. The FDA has a long-standing interest in 'individualization factors', such as those defined by intrinsic factors (for example, age, gender, race, renal dysfunction and genetics) and extrinsic factors (for example, food, co-administered drugs, smoking and alcohol). The Agency believes that an appreciation of controllable sources of variability in drug action and potential injury to patients should be achieved before the marketing of new pharmaceutical products[11]. Information on these important co-variates influencing drug safety and efficacy are generally reported in various sections of the product package insert. PGt, or more specifically the patient genotype, has been shown to be a clinically relevant co-variate for drugs approved recently, as well as those approved decades ago. Understanding the PGt of a drug is the first step towards developing a predictive test to optimize therapeutics.

A recent example of the role that PGt played in the labelling of a new drug is the case of atomoxetine (Strattera; Eli Lilly). This drug was approved by the FDA in November 2002 for attention-deficit/hyperactivity disorder with a fixed dose of 0.5 mg per kg to be titrated up to 1.2 mg per kg. The drug is metabolized by cytochrome P450 2D6 (CYP2D6) with a clearance of 0.35 l per h per kg in extensive metabolizers (EM) and 0.03 l per h per kg in poor metabolizers (PM). The ratio (PM/EM) of the AREA-UNDER-CURVE (AUC) for plasma atomoxetine was ~10. The sponsor did a sensible analysis of adverse events in clinical trials by looking at a post-facto stratification of patient subsets defined by genotype. The frequency of adverse drug reactions (ADRs) — primarily insomnia and irritability — was 9% in PMs and 6% in EMs. There were no major differences in serious ADRs between PMs and EMs.

The label of atomoxetine mentions CYP2D6 in seven different sections, including those describing pharmacokinetics, drug–drug interactions, adverse events and laboratory tests. However, the evidence did not warrant recommending that a pharmacogenetic test for *CYP2D6* status be done before prescribing the drug, but it did provide descriptive information that could be used along with other observations (for example, an adverse event) to guide clinician decisions about an individual's need for dosing adjustment. This example demonstrates the value that pharmacogenetic information in a package insert can bring to the use of a drug, including knowledge related to genotype (for example, *CYP2D6*3*), phenotype (for example, poor metabolizers) and clinical outcomes (for example, adverse events) that can increase the quality of a clinician's decision about individualizing drug treatment.

The atomoxetine example also brings to mind several challenges that face sponsors, regulatory agencies and clinicians in translating genotype information from research to the clinic. First, what is the best way to define PMs in a research setting? The PM phenotype can be determined by the urinary metabolic ratio, the observed AUC or plasma clearance of the drug in different genotype subsets. There are more than 40 ALLELES of *CYP2D6*, and about 25% of these have greatly decreased

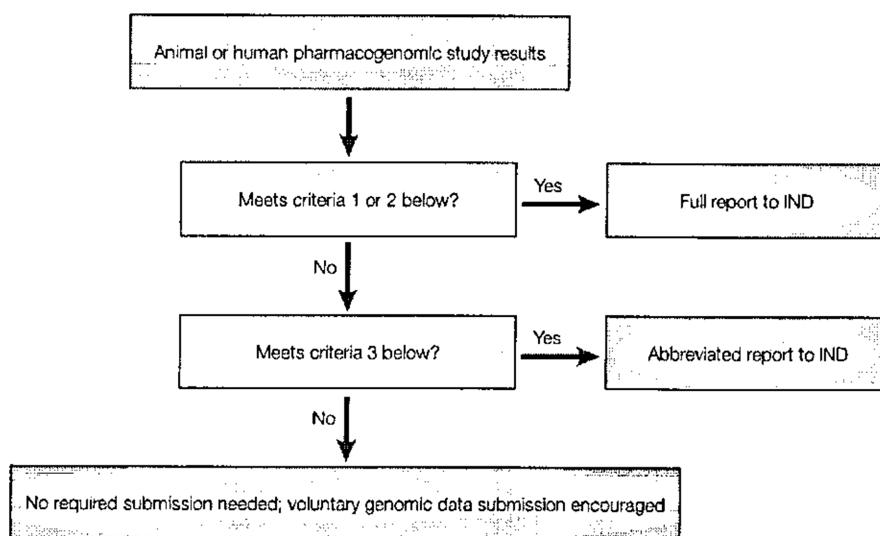

Figure 3 | **An example of a decision tree for submitting pharmacogenomic data during the IND period as a required submission or as a voluntary genomic data submission.** Pharmacogenomic data must be submitted in the Investigational New Drug application (IND) under CFR 312.23 if *any* of the following criteria apply: first, the test results are used for making decisions pertaining to a specific clinical trial, or in an animal trial used to support safety (for example, the results will affect dose selection, entry criteria into a clinical trial, safety monitoring or subject stratification); second, the test results are being used to support scientific arguments pertaining to, for example, the pharmacological mechanism of action, the selection of drug dosing or the safety and effectiveness of the drug; and third, the test results constitute a known valid biomarker for physiological, pathophysiological, toxicological or clinical states and/or outcomes in humans, or is a known valid biomarker for a safety outcome in an animal study. If the information on the biomarker (for example, human *CYP2D6* status) is *not* being used for the purposes described in the first two points above, the information must be submitted to the IND as an abbreviated report. Adapted from Appendix A in the *Guidance for Industry: Pharmacogenomic Data Submissions*.

or null activity. There is also significant variability in the frequency of null alleles of *CYP2D6* in different racial or ethnic groups. So, an open question remains: what alleles should be studied in drug development, and how should this information be translated into a product's package insert?

Second, how should PGt information be reported in the label? This raises two sub-issues: whether or not to report only phenotype data (for example, PMs and EMs), or specific alleles of CYP2D6 (for example, *3, *4 and *5); and the question of who will interpret the significance of these data with respect to dosing, safety and efficacy.

Third, if PGt information is included in the label of a drug product in a way that gives physicians and patients an option to have a genomic test done as part of therapy, this raises translational issues that include public knowledge that the test is available, the quality of the test results, its cost and the proper interpretation of test results.

Despite the high expectations that have surrounded the Human Genome Project, and the frequent reports of the discovery of genes that control a variety of diseases and variability in drug response, there has been relatively little translation of this information into drug development and even less into clinical practice. The FDA believes that there is value in applying long-established PGt to older, marketed drugs in the post-marketing period to improve their risk/benefit ratio by optimizing or individualizing dosing. Examples of older drugs that could benefit from PGt are 6-mercaptopurine (6-MP); azathioprine and 6-thioguanine (6-TG), each of which are substrates for thiopurine methyltransferase (TPMT); irinotecan (a substrate for uridine diphosphate glucoronosyltransferase (UGT1A1)); and warfarin (a substrate for CYP2C9). Each of these drugs has a narrow therapeutic range, wide inter-individual variability in dosing requirements, and frequent and serious safety problems. The genes encoding each of the enzymes mentioned above can exist in one of several isoforms (for example, *TPMT*2*, *UGT1A1*28* and *CYP2C9*3*) and these enzymes are mostly found in either red blood cells (in the case of TPMT) or the liver (for UGT1A1 and CYP2C9). Certain mutations in these isoforms, or gene variants, produce different phenotypes, but the most important factor for drug dosing is the PM phenotype that results in heightened exposure to either the parent drug or a major metabolite, or reduced exposure to an active metabolite (for example, morphine from codeine administration).

In July 2003, the FDA Pediatric Subcommittee of the Oncology Drug Advisory Committee (ODAC) discussed whether or not the package insert of 6-MP should be updated to include information on *TPMT* genotypes. 6-MP was approved decades ago for use in children with acute lymphoblastic leukaemia (ALL) and, taken orally together with methotrexate and/or other chemotherapeutic agents, is the backbone of continuation therapy. The dose intensity of 6-MP is a major determinant of both event-free survival (efficacy) and NEUTROPENIA (safety). The clearance of 6-MP, and therefore exposure to active moieties, is dependent on its conversion to 6-MMP (inactivation via the TPMT pathway) and 6-TG (active nucleotides). More than 11% of individuals in Caucasian populations are heterozygous or homozygous carriers of TPMT null alleles (intermediate or poor metabolizers), which results in the excess accumulation of 6-TG at the expense of 6-MMP formation. There are three major genotypes in the population, each with a range of TPMT activity (high, intermediate and low), and each with a different relative risk of developing neutropenia when administered the standard dose of 6-MP (50 mg per m^2). The PM genotype, which has an incidence of 1 in 300, accumulates excess 6-TG that is nearly certain to lead to severe and potentially fatal bone-marrow toxicity. It has been recommended that the usual dose of 6-MP be reduced by 80–90% for the PM genotype to reduce the risk of neutropenia.

On the basis of the evidence presented in July 2003, the Subcommittee considered the consequences of a label revision thoroughly, and in the end recommended that the label of 6-MP should be updated with current information on *TPMT* genotypes, but stopped short of recommending that testing for *TPMT* status be mandatory before prescribing 6-MP. The experts on the subcommittee considered many factors in making their recommendation, some of which follow: first, the scarcity of prospective clinical trials to support specific recommendations about dose reduction in patients who were either heterozygous or homozygous for null alleles; second, the wide inter-individual variability in TPMT activity, in particular for patients with one variant TPMT allele, and the subsequent risk of reducing effectiveness if doses are reduced erroneously; third, the potential benefit and cost of TPMT genotyping as compared with current phenotyping based on TPMT activity in red blood cells

Glossary

ALLELES
Different or alternative forms of the same gene that can occupy a particular locus on a specific chromosome. Humans have two alleles at that location, one on each chromosome of a homologous pair.

AREA-UNDER-CURVE
(AUC). A metric that summarizes serum or plasma drug concentrations measured over time (for example, 24 hours) in a given individual following the administration of a drug. The AUC is interpreted as the total systemic exposure and is an index of how much of a drug reaches the bloodstream in a set period of time. AUC is also a means to compare the bioavailability of drug from a drug product.

BAYESIAN STATISTICS
A statistical method of analysis that incorporates prior knowledge (for example, on safety and efficacy parameters), specifications of prior distributions and accumulated clinical data experience into making probability calculations and designing future clinical trials.

BIOLOGIC LICENSE APPLICATION
A formal application analogous to a New Drug Application, but for biotechnology-derived pharmaceuticals (for example, complex, large molecules).

BIOMARKER
A characteristic that is objectively measured and evaluated as an indicator of normal biological processes, pathogenic processes or pharmacological responses to a therapeutic intervention.

HAPLOTYPE
A set or combination of alleles or linked genetic markers found on a single chromosome, which tend to be inherited together in a given individual.

INVESTIGATIONAL NEW DRUG APPLICATION
Before initiating any clinical trials of a new drug in humans, a drug sponsor must submit an Investigational New Drug application (IND) to the FDA. The IND contains three broad categories of information: data from animal pharmacology and toxicology studies, manufacturing information, and clinical protocols and investigator information.

NEW DRUG APPLICATION
A formal application that serves as a vehicle through which sponsors propose that the FDA approve a new pharmaceutical (for example, traditional small molecules) for sale and marketing in the United States. When the investigational phase of a drug is completed, the manufacturer submits the results of all the studies to the FDA in a New Drug Application (NDA) for review by FDA officials. The purpose of the NDA is for the FDA to decide if the drug meets the statutory standards for safety, effectiveness and benefit/risk for its intended use, and labelling and manufacturing quality.

NEUTROPENIA
An abnormal decrease in the number of white blood cells in the blood (as measured by an absolute neutrophil count), which increases the risk of infection and fever. It usually occurs as a result of chemotherapy.

SINGLE-NUCLEOTIDE POLYMORPHISM MAPS
A diagram or overview of a stretch of DNA containing single-nucleotide polymorphisms (SNPs). SNPs are DNA sequence variations that occur when a single nucleotide (A, T, C or G) in the genome sequence is altered in different individuals. A map of SNPs across the genome allows genetic traits to be localized by statistical association with the specific region of the genome that is marked by the SNP or multiple nearby SNPs.

coupled with observations of early neutropenia; and fourth, the widespread availability of TPMT testing.

This illustrative example demonstrates that PGt can, in fact, make a contribution to drug safety by guiding doctors towards appropriate dosing. However, translating PGt information from research to the clinic for older drugs is in some ways more challenging than for newer drugs, for the reasons cited above. The three categories of issues or questions raised as challenges following the atomoxetine example also apply to older drugs. However, there are, in addition, other issues and questions that need to be resolved.

First, what is the best way to educate clinicians about the advantages and limitations of adopting a PGt test for a drug that they have been using, albeit not optimally, for decades? This is particularly pertinent for cases such as 6-MP, for which the assessment of neutropenia or another test (for example, TPMT activity in red blood cells) has been used phenotypically as a rough guide to reduce the intensity of dosing.

Second, how should the dosing of a drug such as 6-MP be adjusted, based on genotype, when there is an absence of prospective clinical trials to demonstrate the efficacy of the reduced dose? This is a relevant question in the case of 6-MP, for which the success rate of event-free survival in childhood ALL is nearly 80–85% and evidence supporting the reduction of dose in patients with intermediate TPMT activity is not substantial. Patients with high TPMT activity relative to a given dose might not receive the maximum benefit from the drug because of rapid clearance.

Third, when is the best time for genotyping of patients being administered 6-MP for their TPMT activity status? Options include routinely genotyping TPMT before initiation of 6-MP, genotyping TPMT within the first week of receiving 6-MP or genotyping TPMT only in the case of overt neutropenia.

However, as the Pediatric Subcommittee of ODAC pointed out, genotyping TPMT activity is not a substitute for careful monitoring of white-blood-cell counts in patients receiving 6-MP, but an adjunct. TPMT testing, when combined with other tests and observations, can lead to higher-quality decisions about drug selection and drug dosing that will further decrease the risk of severe and preventable bone-marrow suppression. The FDA is in the process of revising the 6-MP label on the basis of the recommendations of the Subcommittee and is deliberating all of these challenges in translating PGt data into useful information for practitioners and their patients.

Conclusion

The FDA has become a proactive and thoughtful advocate of PGx and PGt, and believes that as a public health Agency it has a responsibility to play a leading role in bringing about the translation of PGx and PGt from bench to bedside. The FDA also realizes that it can hinder innovation and become a regulatory barrier in the translational process if it is not careful with its guidance, policies and procedures. The Agency hopes that pharmaceutical companies view advances in PGx and PGt as an opportunity and one kind of investment in R&D that can help bring a fresh approach to addressing the 'pipeline' problem outlined in the FDA Critical Path white paper.

We believe that PGx and PGt have the potential to revolutionize the drug development process, making it more efficient and bringing value to patient care, including more diagnostic or test products to individualize therapy. This could, in retrospect, seem to have taken much longer than was anticipated but we feel that progress is being made. Regulatory agencies, pharmaceutical companies, the clinical community, third-party payers and patient-advocacy groups are all interested in strategies that can improve the cost, quality and time of drug development, and reduce the risks associated with drug therapy in patients. We do not expect that big changes in these areas will happen overnight with one seminal event or be straightforward to implement, but rather will occur in a more evolutionary or iterative manner, such that progress builds on one successful application of PGx or PGt after another — which now seem to be occurring more rapidly.

We acknowledge that there are, and will continue to be, many different kinds of challenges in translating PGx and PGt from bench to bedside, ranging from issues of historical practices, cost, test availability and reimbursement, to issues of science, biomarker validation, education and adoption of PGx and PGt tests into clinical practice. But, as has been highlighted by the promising results with gefitinib, and the tried and true examples of atomoxetine and 6-MP, these challenges are being met and overcome to benefit both the science of drug development and the quality

of public health. The FDA can be influential and will play an important role in collaborating with others in translating the important discoveries of PGx and PGt from bench to bedside.

But we also need to be on guard. We are aware that drug development is a global enterprise, and therefore international collaboration between regulatory agencies must continue to grow further to harmonize guidance and policies in a way that facilitates and not complicates the drug development process. We must also strive harder to engage the various stakeholders and constituencies, in both the private and public sectors, in conversation regarding effective strategies to advance PGx and PGt. It is clearly in the interest of everyone to streamline the pre-approval drug development process (in terms of cost, time, early attrition, and late-phase success) and reduce the likelihood of toxicity in the post-approval period. We hope that others view the key initiatives and strategies adopted by the FDA — the Critical Path white paper, and its advocacy of PGx and PGt — as a willingness to work together to link bench discoveries to bedside benefits, and we look forward to continued involvement.

Lawrence J. Lesko is Director of the Office of Clinical Pharmacology and Biopharmaceutics, Center for Drug Evaluation and Research, Food and Drug Administration, Rockville, Maryland, USA.
Janet Woodcock is Acting Deputy Commissioner of Operations, Food and Drug Administration, 5,600 Fishers Lane, Rockville, Maryland 20857, USA. Correspondence to L. J. L.
e-mail: leskol@cder.fda.gov

doi:10.1038/nrd1499

1. DiMasi, J., Hansen, R. W. & Grabowski, H. G. The price of innovation: new estimates of drug development costs. *J. Health Econ.* **22**, 151–185 (2003).
2. Gilbert, J. Rebuilding big PhRMA's business model. *In Vivo* November, 73–80 (2003).
3. Tufts CDSS quantifies savings from boosting new drug R&D efficiency. *Tufts CSDD Impact Report* **4**, 1–4 (2002).
4. Lynch, T. J. *et al.* Activating mutations in the epidermal growth factor receptor underlying responsiveness of non-small cell lung cancer to gefitinib. *N. Eng. J. Med.* **350**, 2129–2139 (2004).
5. Guillermo Paez, J. *et al.* EGFR mutations in lung cancer: correlation with clinical response to gefitinib. *Science* **304**, 1497–1500 (2004).
6. *Innovation or Stagnation? Challenge and Opportunity on the Critical Path to New Medical Products* [online], <http://www.fda.gov/oc/initiatives/criticalpath/whitepaper.html> (2004).
7. Lesko, L. J. & Woodcock, J. Pharmacogenomic-guided drug development: regulatory perspective. *Pharmacogenomics J.* **2**, 20–24 (2002).
8. Lesko, L. J. *et al.* Pharmacogenetics and pharmacogenomics in drug development and regulatory decision making: report of the first FDA-PWG-PHRMA-DRUSAFE workshop. *J. Clin. Pharmacol.* **43**, 342–358 (2003).
9. Salerno, R. A. & Lesko, L. J. Pharmacogenomics in drug development and regulatory decision-making: the genomic data submission (GDS) proposal. *Pharmacogenomics* **5**, 25–30 (2004).
10. Salerno, R. A. & Lesko, L. J. Pharmacogenomic data: FDA voluntary and required submission guidance. *Pharmacogenomics* **5**, 503–505 (2004).
11. Peck, C. C., Temple, R. T. & Collins, J. Understanding consequences of concurrent therapies. *JAMA* **269**, 1513–1518 (1993).

Acknowledgements
We would like to acknowledge the helpful comments of our FDA colleagues F. Frueh, S.-M. Huang and A. Rudman.

Competing financial interest
The authors declare no competing financial interests.

Online links

DATABASES
The following terms in this article are linked online to:
Entrez Gene:
http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=gene
CYP2D6 | EGFR | TPMT | UGT1A1
National Cancer Institute Cancer Topics:
http://www.cancer.gov/cancer_information/
ALL | NSCLC
Online Mendelian Inheritance in Man:
http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=OMIM
Alzheimer's disease

FURTHER INFORMATION
Center for Drug Evaluation and Research:
http://www.fda.gov/cder/
Draft Guidance for Industry: Pharmacogenomic Data Submission: http://www.fda.gov/cder/guidance/5900dft.pdf
Food and Drug Administration: http://www.fda.gov/
National Institutes of Health: http://www.nih.gov/
Access to this interactive links box is free online.

The European Agency for the Evaluation of Medicinal Products
Evaluation of Medicines for Human Use

London, 21 November 2002
EMEA/CPMP/3070/01

COMMITTEE FOR PROPRIETARY MEDICINAL PRODUCTS (CPMP)

POSITION PAPER ON TERMINOLOGY IN PHARMACOGENETICS

DISCUSSION AT THE PHARMACOGENETICS EXPERT GROUP	April/September 2001
SUBMISSION TO CPMP	November 2001
RELEASE FOR CONSULTATION	December 2001
DEADLINE FOR COMMENTS	June 2002
DISCUSSION AT THE PHARMACOGENETICS EXPERT GROUP	September 2002
ADOPTION BY THE CPMP	November 2002
DATE FOR COMING INTO OPERATION	June 2003

7 Westferry Circus, Canary Wharf, London, E14 4HB, UK
Tel. (44-20) 74 18 84 00 Fax (44-20) 74 18 86 13
E-mail: mail@emea.eu.int http://www.emea.eu.int
©EMEA 2002 Reproduction and/or distribution of this document is authorised for non commercial purposes only provided the EMEA is acknowledged

1. Introduction

Pharmacogenetic research started from the observations that not all subjects respond in the same way to the same medicine and that these differences between individuals may be caused partially by their genetic profile.

Today the drug development programmes consider (usually for practical reasons) the subjects as coming from a rather homogenous population since it is not possible to accommodate fully in the drug development programme the whole range of inter-individual variability observed within a population. When differences in drug response are anticipated, e.g. in subjects with renal or hepatic disease, or with age-related differences, then studies are requested in the specific subgroup identified.
The contribution of genetic influences to variability in drug response often far exceeds that of any other variable and is what the science of pharmacogenetics aims to unravel. The analysis of a broad set of genetic variations may show that a genotypically defined subgroup of subjects may have a higher probability of responding to a certain drug differently from others in the population. The overall genetic profile may vary according to ethnicity.

As a result of the development within the areas of genetics and genomics, changes are likely to occur in the way drug development is currently being conducted and the way medicines will be used.

The use of terms that are harmonised and widely accepted by the stakeholders would contribute greatly to clarity in the dialogue. At present there is not an agreed set of working definitions crucial for pharmacogenetic clinical research. This is urgently required for protocols and guidelines addressing pharmacogenetic testing to ease communication particularly between ethics committees, investigators and subjects.

Following extensive consultation, the CPMP has agreed on a specific set of definitions directly relevant to the current practices in clinical research, with the understanding that they may have to be revisited in the light of future scientific advance and taking into account emerging legislation. The definitions discussed hereafter are highly relevant to the scenario of individual clinical protocols including pharmacogenetic testing; the principles might however be relevant also for trials involving testing other than pharmacogenetics.

The terms "pharmacogenetics" and "pharmacogenomics" as well as the terms used in the handling of samples and data for pharmacogenetic testing have been defined from the scientific-technical point of view.
The same definitions, following appropriate consultation will then be written in lay-terms and made available in all EU official languages to constitute a useful technical asset for regulatory authorities, ethics committees, health professionals and subjects when confronted with pharmacogenetic testing protocols and consent documents for medicinal product clinical trials.

©EMEA 2002

2. Scope

This position paper focuses on a specific set of critical terms that are frequently used in protocols for pharmacogenetic testing and that are relevant to define appropriate levels of protection for the privacy of the subjects when describing how the results and samples will be used in clinical trials.

The choice of the level depends on the extent to which it is desired or considered possible to link the data and samples to an identifiable subject and corresponds to the defined category of sample linkage.

The most appropriate level for a particular study depends on the nature of the research, the intended use of the data, the regulatory and legal environment and the specific concerns of the investigator and study sponsor. This choice must respect the needs for the privacy of subjects participating in a clinical study.

Generally, the greater the subject privacy in a study, the less are the opportunities for the subject after sample collection and pharmacogenetic testing have been performed to withdraw the individual samples from further analyses or to receive individual results from the study.

Privacy of information, control over the use of samples, and knowledge of study results may all contribute to a subject's willingness to take part in a study, and as a consequence the choice of process may significantly affect enrolment in a clinical trial in which pharmacogenetic testing is planned.

Sample coding procedures should be documented according to Good Clinical Practices (GCPs) and as provided for by relevant EU directives and accompanying guidance documents. Primary study data and original study-related records should be accessible to the competent regulatory authority in order to validate the evidence that is reported. While the regulatory authority can accept different levels of documentation, depending on the particulars of the study and the availability of other evidence or records, there may be times when it is necessary to link a clinical outcome to a particular patient. In principle, there is a framework for protecting patients enrolled in clinical trials now, and this framework may be adequate, perhaps with small changes, to apply to clinical pharmacogenetic trials.

Complete anonymity of the subject without any possibility of linking the samples/data to an individual will have great impact on the usefulness of the results and on what aspects might be verified during a GCP inspection from a competent authority or a sponsor audit. The individual subject record is an important component of data for submission to regulatory agencies and so the use of data from a study involving anonymised samples might not be acceptable for the submission of a claim to be included in the label of a drug or clinical diagnostic assay.

In designing clinical trials, investigators and sponsors should attempt, in consultation with competent authorities and ethics committees, to find the optimum balance between achieving the aims of the study and protecting the subject's safety or right to privacy.

It is recognised that DNA data unique to a subject could potentially be used to reconstruct a link between a subject's medical record and genotype information. Procedures should ensure that in order to respect the subject's wishes and privacy, such links are not reconstructed. For the same reasons, it is further recommended that the code should comprise randomly assigned numbers/letters and should not be based on protocol and site number (and perhaps gender) because if a particular site has included only a few subjects, it might be theoretically possible to reconstruct a link to individual subjects.

3. Pharmacogenetics and Pharmacogenomics

There is at present no consensus in the literature on the definitions of "pharmacogenetics" and "pharmacogenomics". Actually the terms are frequently used interchangeably. The achievement of widely accepted working definitions of the two would be a useful first approach to applying pharmacogenetics and pharmacogenomics in clinical trials. It is important to single out pharmacogenetics and pharmacogenomics from the wider field of genetic testing as the latter encompasses different level of concerns especially in terms of sensitivity of sample handling, data and trial results management.

Pharmacogenetics is the study of interindividual variations in DNA sequence related to drug response.

Pharmacogenomics is the study of the variability of the expression of individual genes relevant to disease susceptibility as well as drug response at cellular, tissue, individual or population level. The term is broadly applicable to drug design, discovery, and clinical development

4. Definitions applicable to DNA samples and data in clinical trials including pharmacogenetic testing

Different terminologies relate to the collection of human samples for pharmacogenetic research and the management of the data therefrom. The set of terms described in this paper are a key to correct handling of the samples and the data and to transparency of communication among industry, ethics committees, regulatory authorities and subjects about the pharmacogenetic approach in clinical research, regulatory assessment of medicinal products and clinical practice.
The processes by which samples and data are collected, labelled and stored have a direct effect on how the samples and the results obtained can be used in the future and on the obligations of the investigator and sponsor to the sample subject. This pertains particularly to situations when a subject withdraws his or her consent to further participation in a study and affects the possibility to return information to the subject or his/her physician, the possibility to withdraw a sample from future analyses and verification of data ascribed to a subject in reports and regulatory submissions. Additionally, the readiness and willingness with which a subject would or would not want to take part in a study may be affected by such factors as the uses of the results, the nature of the information the subject might receive, and the perceived risk resulting from disclosure of genetic information to third parties.

Five definitions (See table 1) for the labeling and coding of pharmacogenetic samples and data are proposed describing direct implications for the handling methodology of samples for pharmacogenetic testing and corresponding consequences for the level of privacy protection and use of the information for regulatory purposes. Duration of retention of the sample or its destruction needs to be defined in the protocol and in the consent form. Otherwise, if and when relevant, the timepoint and the procedure for anonymisation of the sample itself should be defined in these documents.

4.1 Identified samples and data

are those labeled with personal identifiers such as Name or Social Security Number.

Identified samples and data are treated in much the same way as those acquired in everyday medical practice. Because the sample and the data generated from it are directly traced to the subject, it is easy to withdraw the sample or the data from the study, update subject information, and return results to the subject. Also, at an inspection of the study it will be possible to verify the connection between the subject and the reported results. On the other hand, since a subject's genotyping results are directly linked to the subject's identity, the use of identified samples offers no extra privacy protection in addition to those generally provided.

Identified samples and relevant data might be coded at the given point in time in order to provide for extra long-term privacy protection following the closure of the trial.
The protocol should also specify when and whether the samples and data might be destroyed or anonymised.

4.2 Single coded samples and data

are those to which a single specific code is attributed for protecting individuals. It is recommended that the code should comprise randomly assigned numbers/letters

The investigator stores the key connecting the code of the sample to the individual's data. This step separates the subject's identity from the results of the pharmacogenetic analysis. The researcher with knowledge of the pharmacogenetic data would not have ready access to the identity of the subject.
Only breaking the code can reveal the subject's identity.
It is possible to withdraw a subject's sample for prospective use or return individual results to the subject or physician if desired.
The maintenance of a link between the subject and the pharmacogenetic information by a single code allows verification of data ascribed to an individual subject. Because the investigator who has coded the sample might also have access to the pharmacogenetic data, the safeguards of the subject's privacy, including doctor-subject confidentiality, are equivalent to those in current clinical trials practice.

4.3 Double-coded samples and data

have an additional privacy safeguard imposed by the use of a second coding system. Adding an additional code to the samples and data provides further protection.

The investigator who only knows the first code does not know this second code. In this way, anyone with knowledge of the pharmacogenetic results can only trace a subject identity to a coded identifier but no further, unless a key is used to link the codes between the data set with subject identifiers and the data set containing the pharmacogenetic information.
The code key linking the double coded pharmacogenetic samples and information is kept by a third party. This should not be the investigator in possession of the key linking coded sample and/or information to the subject.
The key to the double code might be maintained by the sponsoring organisation, in areas entrusted with maintaining confidential information (e.g. legal, quality assurance, clinical statistics) under strict operating procedures. Alternatively, the key might be held by an external entity, such as governmental agency, legal counsel, or other qualified third party not involved with the research.

The individual can only be linked with the sample or data obtained from it by bringing the two code keys together. Although the samples do not carry any information on the identity of

the subject, it is still considered to be possible to identify the subject as long as both code keys exist.

As with single coded samples, the existence of a link between the pharmacogenetic data and the subject's identity makes it possible to withdraw a sample or data (up to the time the results stemming from that data are reported), update subject information, return results and inspect the process. However, the conditions under which the pharmacogenetic information might be linked back to the subject's identity for any purpose are determined strictly by the specifics of the research protocol.
These conditions should be explicitly described in each protocol, and included within the subject's informed consent.

4.4 Anonymised samples and data

are for practical purposes double coded samples where the key linking the first and/or second code is deleted. They may be also previously single coded samples where the single code key is destroyed or even previously identifiable samples where the name/identifier is removed.

Anonymised samples and data do not carry any longer personal identifiers. Once the linking key has been deleted, information related to the subject's identity is no longer linked to data related to the pharmacogenetic results. This offers an additional level of security to the individual's data.

After anonymisation it is not possible to withdraw a subject's sample from analyses, to update subject information for further use, or to return any individual results to the subject or the subject's physician. Similarly, it also is not possible to inspect the study to determine that pharmacogenetic data is accurately correlated to a specific subject.

There will be times when stored samples may provide a regulatory agency additional information related to clinical outcome. The ability to link individual data to a patient will be essential in some circumstances and anonymised samples would be a problem.

In general, anonymised samples are well suited to research studies in which hypotheses are generated, but may be less so for clinical trials on which label claims are based.

4.5 Anonymous samples and data

are those that do not have any link whatsoever between the sample and the individual identity.

Anonymous samples may have population information (e.g., the samples may come from subjects with diabetes) but no individual data that might allow the identity of the subject to be traced. The clinical information is limited to broad categories of data, such as "male, age 50-55, cholesterol > 240 mg / dl". In many instances, the sample has no clinical data at all.
This situation is applicable in cases where the population is large enough and measures are taken in building up the code (see recommendations on page 3 on reconstructing a link).
Anonymous samples are useful in some types of pharmacogenetic studies.

Table 1. Summary table of the five terms of sample labelling.

:

Sample Labelling Category	Link Between Subject Identity and Pharmacogenetic Data	Records Identifiable for Clinical Monitoring	Actions Possible if subject withdraws Consent	Return of Individual Results to Subject	Scope of Subject Privacy protection
Identified	Yes, directly	Yes	Sample can be withdrawn with immediate effect for any prospective use	Possible	Similar to general healthcare confidentiality
Single coded	Indirectly, via code key	Yes, via protocol-specified procedures	Sample can be withdrawn with immediate effect for any prospective use	Possible	Standard for clinical research Conforms to principles of GCP
Double-coded	Very indirectly, via two sets of code keys	Yes, via protocol-specified procedures	Sample can be withdrawn with immediate effect for any prospective use	Possible	Double code offers added privacy protection over single code
Anonymised	No. Key(s) identifying the link between pharmacogenetic data and the identity of the subject is deleted	No	Sample and data are not identifiable. Sample cannot be withdrawn once key is deleted	Not possible	Pharmacogenetic data not linked to individuals
Anonymous	No	No	None	Not possible	Complete

©EMEA 2002

Reprinted from

The Pharmacogenomics Journal

VOLUME 2, SUPPLEMENT 5 **AUGUST 2002**

Elements of informed consent for pharmacogenetic research; perspective of the pharmacogenetic working group

DC Anderson[1], B Gomez-Mancilla[1], BB Spear[2], DM Barnes[2], K Cheesman[3], PM Shaw[4], J Friedman[4], A McCarthy[4], A McCarthy[5], C Brazell[5], SC Ray[6], D McHale[7], L Hashimoto[8], R Sandbrink[9], ML Watson[10], RA Salerno[11], N Cohen[12], and CE Lister[13], on behalf of The Pharmacogenetics Working Group

The Pharmacogenomics Journal (2002) 2, 284–292
© 2002 Nature Publishing Group All rights reserved 1470-269X/02 $25.00
www.nature.com/tpj

CLINICAL IMPLICATION

Elements of informed consent for pharmacogenetic research; perspective of the pharmacogenetics working group

DC Anderson[1], B Gomez-Mancilla[1], BB Spear[2], DM Barnes[2], K Cheeseman[3], PM Shaw[4], J Friedman[4], A McCarthy[5], C Brazell[5], SC Ray[6], D McHale[7], L Hashimoto[8], R Sandbrink[9], ML Watson[10], RA Salerno[11], N Cohen[12] and CE Lister[13]on behalf of The Pharmacogenetics Working Group

[1]*Pharmacia Corporation, Pharmacogenomics, Kalamazoo, MI, USA;* [2]*Abbott Laboratories, Pharmacogenetics, Abbott Laboratories, Abbott Park, IL, USA;* [3]*AstraZeneca, Experimental Medicine, Loughborough, UK;* [4]*Bristol-Myers Squibb, Pharmacogenomics & Human Genetics, Princeton, NJ, USA ;* [5]*GlaxoSmithKline, Genetics Research, Greenford, Middlesex, UK;* [6]*GlaxoSmithKline, Genetics Research, Research Triangle Park, NC, USA;* [7]*Pfizer, Global Research & Development, Sandwich Laboratories, Sandwich, Kent, UK;* [8]*F Hoffmann-La Roche Ltd., Roche Genetics, Basel, Switzerland;* [9]*Schering AG, SBU Specialized Therapeutics, Berlin, Germany;* [10]*Merck & Co., Clinical Genomics, West Point, PA, USA;* [11]*Wyeth Research, Wyeth Pharmaceuticals, Inc., Worldwide Regulatory Affairs, St. Davids, PA, USA;* [12]*Johnson & Johnson Pharmaceutical Research and Development, Pharmacogenomics, Raritan, NJ, USA;* [13]*Covington and Burling, London, UK*

Members of the Pharmacogenetics Working Group include: Abbott Laboratories, Amgen, AstraZeneca, Aventis, Bristol-Myers Squibb, Eli Lilly, GlaxoSmithKline, Johnson & Johnson, Merck, Novartis, Organon, Pfizer, Pharmacia Corporation, F Hoffmann-La Roche Ltd., Schering AG, Schering-Plough, and Wyeth Research.

The Pharmacogenomics Journal (2002) **2**, 284–292. doi:10.1038/sj.tpj.6500131

INTRODUCTION

The explosion of new human genetic data and an increasing awareness that genetic variation influences drug responsiveness has signaled a new era of pharmacogenetic research. Rich opportunities now exist to dissect the variable pharmacokinetic and pharmacodynamic effects of new and marketed drugs as well as the molecular-genetic basis for disease heterogeneity. Pharmacogenetic investigations should provide important advances in the design, development and delivery of safe and efficacious pharmaceuticals, including opportunities to provide customized drugs for specific patient populations as defined by their unique genotypes. As such, pharmacogenetics represents an important new approach to address directly unmet medical needs of high interest to consumers, health care providers and the pharmaceutical industry.[1] The guidelines and policies of diverse regional or international ethical, regulatory, medical and scientific bodies emphasize the importance of enhancing health care through genetics-based research.[2–8]

Informed consent (IC) is the means by which potential research subjects make a judgment about the contribution that their involvement in the research can make, relative to the risks or benefits to them as individuals. Pharmacogenetic research in drug development involves special considerations and disclosures in the informed consent process. These disclosures are of practical importance in the context of available options and strategies for incorporating genetic objectives into clinical studies and are of ethical importance in terms of the implications of genetic data to be derived from such studies, including the potential risks for genetic discrimination. However, early experience in the design and implementation of such studies by pharmaceutical sponsors indicates a need for consistent practices in the IC process. In the context of multi-national drug trials and registration efforts, IC policies may be inconsistent among individual institutional review boards (IRB) or independent ethics committees (IEC), and in some cases IRB/IEC requirements may be contradictory among different countries. Several approaches to protecting human subjects in clinical genetic research through IC have been proposed. These have merit in particular contexts, but none of these address all issues in all circumstances.[9–13]

The Pharmacogenetics Working Group (PWG) is a voluntary association of pharmaceutical companies involved in clinical drug trials and genotyping whose goal is to advance the understanding and development of pharmacogenetics by addressing non-competitive ethical, regulatory, and legal issues. This group previously defined harmonizing terminology for sample collection in clinical genetic studies.[1] This communication considers the key elements of the IC process[7] of special relevance to the design and approval of pharmacogenetic trials whereby pharmacogenetics is described as the study of DNA sequence variation as it relates to differential drug response. The major goals of this

communication are to assist researchers, IRBs/IECs and regulatory agencies in better understanding issues specific to pharmaceutical company-sponsored pharmacogenetic research so that they can best assure protection of subjects, while at the same time facilitating timely review, approval and implementation of pharmacogenetic trials.

INFORMED CONSENT PROCESS FOR PHARMACOGENETIC INVESTIGATIONS

Informed consent can be defined as both the document used to obtain consent as well as the process utilized to communicate the intended disclosures and to ensure accountabilities for the consequences of obtaining consent. Defining the best approach to obtain consent for pharmacogenetic research is challenging. Because genetic terminology and concepts may be intrinsically difficult to understand, study subjects should be encouraged to ask questions to confirm their understanding of the purpose for participating and what is to be learned from the studies. Signature of the person conducting the IC discussion facilitates accountability for this understandable communication and dialogue. When appropriate to enable clear communication, consent materials should be given to subjects prior to the research visit so there is sufficient time for review. In order to facilitate understanding of a pharmacogenetic study, some sponsors have created educational pamphlets and videos as an adjunct to the consent form to provide additional information to assist the subject in making decisions about participation. Also, special attention should be paid to the document itself, to assure that the language is understandable, and at an appropriate reading level and that the risks, benefits and purpose of the study are clearly explained. It is noteworthy that some background information a subject may have received prior to being asked to participate in a pharmacogenetic trial is likely to have come from media or a historical perspective suggesting (to the subject) that a pharmacogenetic trial involves cloning, reproductive choices or risks for serious monogenic disorders. To avoid these possible misconceptions, the IC consent process should explain not only what pharmacogenetics is, but also what it is not.

KEY ELEMENTS FOR DISCLOSURE AND CONSENT IN PHARMACOGENETIC STUDIES

Purpose(s) and Intent of Pharmacogenetic Studies

As is true for all clinical studies, the researcher must describe, in terms understandable to the study subject, the overall objectives of the study and the subject's role in it. For pharmacogenetic studies, the subject should be provided with some background information about the biologic function of genes (for example, that genes affect physical features and health status) and how such studies may help scientists and clinicians learn more about health, disease, and drug treatments.

The specific purpose of the study should be clearly described including both short-term objectives and potential long-term applications. A description of the disease(s) or clinical conditions of investigative interest during the current trial, as well as those disorders of potential interest for study in the future should be described. Depending on the study design, specific language or broad descriptions may be appropriate. Readily understandable study endpoints should be conveyed, such as 'to identify genetic reasons why certain individuals respond differently to drugs' or 'to identify variations of genes which may cause or modify a disease'.

Within certain countries or for specific IRBs/IECs, explicit identification of the genes (or genetic pathways) to be studied may be required for approval. Such IC requirements have the important disadvantage of limiting the use of DNA or other genetic materials for pharmacogenetic evaluations in a way that will not allow researchers or research sponsors to maximize the value of donated samples in light of future knowledge or hypotheses. Moreover, such requirements cannot be satisfied for studies designed to discover unknown genetic determinants of drug response, ie, where no *a priori* assumption is being made. Since these objectives represent meritorious goals with potentially positive health implications, it may be desirable to limit the imposition of such restrictions.

The IC document should identify all intended uses of the pharmacogenetic information and clinical information to be derived from the study. If the research is sponsored by a commercial entity or has commercial (and intellectual property) implications, this should be clearly described in the IC document because some individuals may choose not to participate in such studies. Plans for archiving the subject's DNA and/or creating immortalized cell lines (which could provide an inexhaustible source of DNA for future studies) should be clearly revealed, and any plans for distribution of the subject's genetic materials to secondary users should be presented, even if such parties are not yet defined. This is of special importance in the context of regulations restricting blood/DNA shipment from certain countries/regions. Some, but not all, existing guidelines or policy statements imply or state that the study subject has a right to decide (prospectively) the future uses of his/her sample.[9,10] Such guidelines imply a requirement for re-contact or re-consent. This option is not always practical or even possible and is dependent on the category and the relative anonymity of the genetic samples collected for a pharmacogenetic study.[1]

Trial Procedures

The IC process for pharmacogenetic trials should clearly describe the procedures involved in collecting and handling samples, and the options available to the patient once a sample has been acquired and genetic information has been derived. These procedures and options will vary depending on the trial design, the research sponsor's internal standard operating procedures (SOP) and the preferences of the researchers and their institutional policies.

Voluntary participation
Enrolment in pharmacogenetic studies requires voluntary participation through IC as for any clinical research protocol.[14] However, pharmacogenetic trials commonly utilize separate IC documents for a drug research protocol and a related pharmacogenetic sampling protocol. The latter is often prepared as an amendment to the main study protocol to enable study subjects to make an informed choice about participating in the pharmacogenetic study independent of their decision to participate in the drug research protocol. A separate IC document for pharmacogenetic sampling must provide sufficient information for the subject to make an informed decision to donate genetic materials based on the merits and risks of the pharmacogenetic objectives and procedures alone. Importantly, this approach allows for efficient patient enrolment in parent drug trials (with optional pharmacogenetic components), which may be compromised if subjects are uncertain whether to volunteer for the pharmacogenetic studies. In studies where the pharmacogenetic aspects cannot be separated from the parent trial (for example, when genotyping is an inclusion criteria), a single consent form is appropriate.

Pharmacogenetic sample collection, storage and distribution
A complete description of sample collection procedures (for example, phlebotomy, volume of blood and buccal swab technique) should be part of the IC process. This should include an indication of which procedures are part of routine clinical care, and which are specifically applicable to the pharmacogenetic research objectives, descriptions of who will be handling the samples, and where and how long the samples will be stored. The subject must be informed, if relevant, that enrolment in pharmacogenetic studies may require a detailed family history of disease or other genetic traits, and that such research will require the collection of materials (for example, blood cells, buccal cells, or other tissues) used as a source for the extraction of DNA or other genetic materials. Information concerning sample storage, sample replication (for example, creation of immortalized cell lines), and/or the distribution to third parties (such as other investigators and commercial entities) for additional collaborative studies should be fully disclosed. Specific plans and timelines for sample destruction or depletion should also be defined, especially in the context of subject protection against informational risks (see the section, 'Confidentiality of Subject Information').

Withdrawal options and timelines
In contrast to standard clinical studies or drug trials, the process of 'withdrawal' of subject participation from pharmacogenetic studies may involve a request by the subject (or others) to destroy genetic materials collected. The possibility of this option is dependent on the category of genetic samples as defined in the study protocol,[1] the type and quantity of genetic material collected or derived, and the time interval during which the genetic materials are maintained by the researcher or research sponsor. The IC process should indicate when sample destruction will not be possible, for example, because of pooling of individual samples in the laboratory, and should also indicate the circumstances under which individual genetic results cannot be retrieved, such as after data pooling or entry into anonymous/anonymized databases. As for drug trials, pharmacogenetic data collected and/or analyzed up to the time of a request by a study subject to withdraw should be maintained by the sponsor, as it is not consistent with good clinical practices to delete individual data from stored data sets.

The IC process should clearly describe the requisite procedures for withdrawal and the time interval or circumstances after which withdrawal may not be possible. In situations where withdrawal is feasible, the sponsor must ensure that logistical capabilities exist to permit the identification and disposal of a pharmacogenetic sample when requested by the subject. An explanation should be provided of the circumstances allowing or preventing the destruction of a genetic sample. For example, it should be explained to subjects that it is possible to destroy an *identified, coded,* or *de-identified/double-coded* sample if requested, but that this option is not possible for *anonymized* or *anonymous* clinical samples (see Table 1).[1] Withdrawal of consent by destroying *identified, coded, or de-identified/double-coded* samples may be possible, but may be limited by the length of time a researcher or research sponsor maintains study records or by the requirement to maintain data for regulatory submissions. In turn, this time interval may be determined (required) by regulatory guidelines in registration studies.[15] For studies in which samples will eventually be *anonymized* (for example, following the completion of a registration study) the IC process should indicate that withdrawal may be possible for only a specified time interval (for example, three months) to allow

Table 1 Outline of the major characteristics of the five categories of sample labeling[a]

Category	*Link between subject identity and genetic data is possible*	*Option for re-contact or re-consent for subsequent studies is possible*
Identified	Yes, directly	Yes
Coded	Indirectly, via code numbers	Yes
De-Identified or Double coded	Very indirectly via two levels of code numbers	Yes
Anonymized	No, key between first and second codes is deleted	No
Anonymous	No	No

[a]Table adapted from Spear *et al.*