

e-PKGene: A knowledge-based research tool for analysing the impact of genetics on drug exposure

Houda Hachad,¹ Casey Lynnette Overby,² Sophie Argon,¹ Catherine K. Yeung,¹ Isabelle Ragueneau-Majlessi^{1*} and René H. Levy¹

¹Department of Pharmaceutics, University of Washington, Seattle, WA 98195, USA

²Department of Medical Education and Biomedical Informatics, University of Washington, Seattle, WA 98195, USA

*Correspondence to: Tel: +1 206 543 4669; Fax: +1 206 543 6131; E-mail: imaj@uw.edu

Date received (in revised form): 20th December 2010

Abstract

e-PKGene (www.pharmacogeneticsinfo.org) is a manually curated knowledge product developed in the Department of Pharmaceutics at the University of Washington, USA. The tool integrates information from the literature, public repositories, reference textbooks, product prescribing labels and clinical review sections of new drug approval packages. The database's easy-to-use web portal offers tools for visualisation, reporting and filtering of information. The database helps scientists to mine pharmacokinetic and pharmacodynamic information for drug-metabolising enzymes and transporters, and provides access to available quantitative information on drug exposure contained in the literature. It allows in-depth analysis of the impact of genetic variants of enzymes and transporters on pharmacokinetic responses to drugs and metabolites. This review gives a brief description of the database organisation, its search functionalities and examples of use.

Keywords: Database, pharmacogenetics, metabolism, transporters, cytochrome P450 enzymes

Introduction

Differences in drug response among patients are common, often leading to challenges in optimising a dosage regimen for an individual patient. Genetic factors have long been known to cause interindividual differences in the pharmacokinetics, efficacy and adverse events of a number of drugs,¹ and drug metabolising enzymes have been shown to be the greatest source of pharmacogenetic (PGx) variability identified to date. It is estimated that over half of the ~170 genes with products affecting drug disposition are polymorphic,² and clinically important polymorphisms have been identified for most major enzymes involved in both phase I and phase II drug metabolism.³ More recently, the polymorphic variability of several transporter proteins, such as the hepatic uptake transporter, organic anion

transporter polypeptide 1B1 (OATP1B1), has been shown to have an impact on the exposure to, and safety of, widely prescribed drugs.⁴ Thus, incorporating the knowledge gained from PGx research to make decisions in drug development and clinical care has the potential to increase the safety and efficacy of drug treatment, and is central to the strategies of personalised medicine.⁵ In spite of current efforts to incorporate the use of PGx information in drug development, clinical practice and in making cost-effective healthcare decisions, however, information uptake remains low. Translational research is required to move PGx discoveries effectively to evidence-based application in these areas. Translational research has been described as having four iterative phases with feedback loops, to allow integration of new knowledge.⁶ Phase 1 (T1) and

Phase 2 (T2) translational research informs the development of clinical interventions and evidence-based guidelines; Phase 3 (T3) research assesses the implementation of guidelines in health practice; and Phase 4 (T4) research evaluates the health outcomes of changes in practice following the implementation of guidelines.⁶ All phases have become data intensive, with studies in PGx discovery increasing rapidly in both number and throughput. For example, in 2009 there were at least 12 PGx genome-wide association studies conducted.⁷ In a review of 100,000 PubMed-listed publications on pharmacogenomics, less than 2 per cent were identified as *original* research manuscripts,⁸ illustrating the difficulty that exists in locating reliable information for translational research pursuits.

e-PKGene (www.pharmacogeneticsinfo.org) is a manually curated knowledge-based product which facilitates easy access to, and search for, quantitative information contained in the PGx literature base. It provides in-depth analysis of the impact of genetic variants of metabolising enzymes and transporters on pharmacokinetics. The tool's focal point is the identification of the genetic variants that are best correlated with drug exposure. e-PKGene is designed directly to support drug development or Phase 1 translational research (T1). This tool also has the capacity to support other phases of translational research, however, including Phase 2 (T2) evidence-based evaluations, which can better predict patient responses by providing clinical recommendations. Details about e-PKGene design and content, examples of use and types of support provided to various types of users are described in the following sections.

Application description

Structure

The application has a typical multi-tier architecture in a Microsoft[®].NET environment. The web part of the database, which is accessed by the user over the internet, is hosted on a Microsoft Windows 2003 server running IIS and version 2.0 of the ASP.NET framework. All data are stored on a Microsoft SQL Server 2005 database. The use of the web facilitates worldwide access, as well as

upgrades and updates. e-PKGene is being developed by scientists from the Department of Pharmaceutics, University of Washington, USA.

e-PKGene allows for a high level of development flexibility through incorporating structured data and standardised representations of PGx knowledge with use of controlled terminologies. The tool uses hierarchical categorisation to characterise data sources and evidence.

Content

e-PKGene integrates information from the literature, public repositories, reference textbooks, product prescribing labels and clinical review sections of new drug approval packages. The current pilot version focuses on drug metabolising enzymes and transporters that are routinely assessed in the context of drug development, and which are of interest to clinical practice and healthcare economics. The core content of e-PKGene is represented by published pharmacokinetics studies, performed in human subjects (healthy volunteers or patients). Each research article (citation) may contain one or more pharmacokinetics studies in which one or more genes have been investigated. A study is defined as a set of assessments (pharmacokinetics, pharmacodynamics and safety) following the administration of a target compound to a well-defined population. The population is usually divided into a 'reference' group and an 'impaired' group. The reference group (the one to which the other groups are compared) consists either of individuals who are 'extensive metabolisers' or carriers of two copies of the wild-type allele of the gene of interest.

Each citation is assigned a unique identifier (Accession Number). Detailed records are manually curated from each citation, highlighting the study design, the population characteristics (ethnicity, health status, gender etc.), and the genotyping or phenotyping methods. For each population subgroup, pharmacokinetic parameters (area under the curve [AUC], clearance or plasma concentration) — as determined by the study — are presented to the user. Pharmacodynamics, clinical outcome and side effects are also reported when provided in the citation,

especially when their relationships with pharmacokinetics were investigated. The datasets made available to end-users are summarised in Table 1.

To allow an analysis of the impact of genetic variations on drug exposure, a comparison of the

pharmacokinetic parameters between impaired groups and the reference group is systematically provided for each citation. All changes in PK parameters (Δ) are calculated as follow:

$$\Delta PK_p = \left[\frac{(PK_{variant} - PK_{ref})}{PK_{ref}} \right] * 100$$

Where $PK_{variant}$ = PK parameter of the variant group, PK_{ref} = PK parameter of the reference group (control).

Table 1. Available datasets in e-PKGene

Curated facts from publications
Natural language descriptions/comments
Quantitative data (eg AUC, C_{max} , half-life, T_{max})
Variant details
Genotype details
Phenotype details
Population details (eg ethnicity, gender)
External links to PubMed
Design and methods details
Design and drug administration
Genotyping method
Phenotyping method
Alleles and genes tested
Pharmacokinetics measurements
Pharmacodynamics measurements/protocol
Side effects
Drug information
Pharmacology and interactions
Metabolising enzymes and transporters involved
Classification (therapeutic uses, therapeutic range)
PubChem record, CAS number
Genetic information
Gene (common name, synonyms)
Genetic variations (DNA sequence variations [SNP], rs numbers)
AUC, area under the curve; C_{max} , maximum (peak) observed drug concentration; T_{max} , time to reach maximum concentration following drug administration; CAS, Chemical Abstracts Service

Gene and allelic variants

e-PKGene defines each variant allele by one single nucleotide polymorphism (SNP) which is unique to the allele. This SNP is identified as the ‘diagnostic SNP’, but it is important to note that the diagnostic SNP may not be responsible for the observed functional changes. The reference SNP number (rs number, as assigned by the Single Nucleotide Polymorphism Database [dbSNP]⁹) allows the identification of the * allele and the SNP as well. The * allele nomenclature has been adopted according to the Human Cytochrome P450 (CYP) Allele Nomenclature Committee.¹⁰ Each SNP or variant allele is described by its nucleotide change, its location within the gene, its impact on the amino acid change in the protein and the mechanisms by which it affects its activity *in vivo* (Figure 1). When an allele is defined by a haplotype formed by a constellation of SNPs, each SNP is described as well. When available, the functional activity *in vivo* of the enzyme or transporter is assigned. For certain genes, such as *CYP2D6*, an ‘activity score’ is also assigned.

Examples of queries and output

The e-PKGene website allows initial searching based on compound (or metabolite), gene or population. Selection of a search category retrieves a list of all available options. For example, a search by gene generates an alphabetical list of all compounds and metabolites with available data. In the following example (Figure 2), a tamoxifen-based search is conducted. The tamoxifen query generates a results screen (Figure 3) which indicates the total

Genetic characteristics for the most common CYP2D6 alleles (diagnostic Single Nucleotide Polymorphism) ^{(1) (2) (4) (5) (6) (7) (8) (9) (10) (11) (12)}

Allele name	diagnostic nucleotide change (SNP)	SNP position	Effect of variation	Amino acid change	rs number	Activity score model-B	Enzyme activity in vivo
CYP2D6*1	none		'wild type' [reference]	N/A	N/A	1	normal
CYP2D6*2	g.2850 C>T (also seen as 2938 C>T)	exon 6	missense	R296C	rs16947	1	normal
CYP2D6*3	g.2549 del A (also seen as 2637 del A)	exon 5	frameshift (nonsense)	R259	rs4986774	0	none
CYP2D6*4	g.1846G>A (also seen as 1934G>A)	intron 3	splicing defect (nonsense)	182ter	rs3892097	0	none
CYP2D6*5	<i>whole gene deletion</i>					0	none
CYP2D6*6	g.1707 del T (also seen as 1795 del T)	exon 3	frameshift (nonsense)	W152X	rs5030655	0	none
CYP2D6*7	g.2935A>C	exon 6	missense	H324P	rs5030867	0	none

Figure 1. Example of a summary of genetic information for the CYP2D6 polymorphic enzyme (partial view). Display from e-PKGene (<http://www.pharmacogeneticsinfo.org>; accessed 18th November 2010).

number of available citations, the genes studied and the examined populations (with the number of citations). Note that tamoxifen has been flagged as a 'prodrug' (circled in red) to remind the user that additional information may be available by searching for the active moiety. Beneath the summary of results table, the citations are listed chronologically, along with their PMID number.

Upon selection of a gene of interest (eg *CYP2D6*), the 'impact of variants' table is generated

(Figure 4), which lists the variants studied and the moieties for which impact has been assigned. Impact is a binary YES/NO classification of genotype/compound pairs which is based on change of exposure between the reference group (wild type) and a variant group. This classification uses a predefined cut-off given by the statistical analysis performed by the authors (statistical significance). The classification is assigned by the database editorial team to provide users with a contextual framework which rapidly

Figure 2. Example of initial search page. Display from e-PKGene (<http://www.pharmacogeneticsinfo.org>; accessed 18th November 2010).

University of Washington
e-PKGene Impact of Genetics on Drug Exposure

Home Search Summaries Logged in as: pgktest@uw.edu Log Out

Search

Tamoxifen - Prodrug

Results (6 citations)

Refine Search By:

Gene	Population
CYP2C9 (1)	Asians (3)
CYP2D6 (6)	Caucasians (1)
CYP3A5 (1)	Chinese (1)
SULT1A1 (2)	Females (6)
	Japanese (2)
	Koreans (1)

Clear Fields

Page: 1

Citations Impact of Variant

Year	Title	PMID
2010	Significant Effect of Polymorphisms in CYP2D6 and ABC2 on Clinical Outcomes of Adjuvant Tamoxifen Therapy for Breast Cancer Patients.	20124171
2008	Association between CYP2D6 *10 genotype and survival of breast cancer patients receiving tamoxifen treatment.	18407954
2008	Effects of CYP2D6 and SULT1A1 genotypes including SULT1A1 gene copy number on tamoxifen metabolism.	17947222
2007	Clinical implications of CYP2D6 genotypes predictive of tamoxifen pharmacokinetics in metastatic breast cancer.	17761971
2006	Quantitative effect of CYP2D6 genotype and inhibitors on tamoxifen metabolism: implication for optimization of breast cancer treatment.	16815318
2005	CYP2D6 genotype, antidepressant use, and tamoxifen metabolism during adjuvant breast cancer treatment.	15632378

Page: 1

Figure 3. Sample search results screen, with gene of interest (*CYP2D6*) highlighted. Display from e-PKGene (<http://www.pharmacogeneticsinfo.org>; accessed 18th November 2010).

highlights variants of interest. Only variants with well-defined genotypes are presented in this screen; phenotypes are not presented on this screen but are captured and presented on the citation displays. Clicking on the ☒ icon retrieves the PMID number of the relevant citation, and clicking on the citation number generates a ‘citation display window’ (Figure 5) for that citation which displays the full details for an individual citation.

The citation listing contains complete reference information, as well as comments entered by the database editorial team when applicable. The full abstract from PubMed can be retrieved by clicking on the PubMed icon (circled in red). The initial ‘effects’ window shows summarised information for all studies contained within each citation (single citations often involve multiple studies if different doses or populations, or multiple enzyme systems are evaluated separately). The database ‘impact assignment’ is shown for each population examined.

More detailed information is retrieved in the ‘impact of pharmacokinetics’ screen for an

individual citation (Figure 6), where specific pharmacokinetic parameters (AUC, clearance or plasma concentration) for the selected compound and any active metabolites are displayed for all the studies. Additional information extracted from the citation is displayed in the ‘full study set’ section.

The ‘full study set’ tab (Figure 7) displays all information that has been captured for an individual study — this information may include study design, phenotyping and/or genotyping methods, alleles tested for (from which the user can infer the alleles that were *not* tested for) and additional population classifications. This information enables the user to determine the validity of the information by evaluating the study design and methods. While not all case reports are suitable for inclusion in e-PKGene database, ‘case report’ will be listed as the study design in this section for those which can be entered into e-PKGene.

Searches can be performed by gene (Figure 8a) or population (Figure 8b), and will retrieve citations relevant to the search criteria.

University of Washington
e-PKGene Impact of Genetics on Drug Exposure

Home Search Summaries Logged in as: pgxtest@uw.edu Log Out

Search

Tamoxifen - Prodrug

Results (6 citations)

Refine Search By:

Gene	Population
CYP2C9 (1)	Asians (3)
CYP2D6 (6)	Caucasians (1)
CYP3A5 (1)	Chinese (1)
SULT1A1 (2)	Females (6)
	Japanese (2)
	Koreans (1)

Clear Fields

Page: 1

Citations **Impact of Variant**

CYP2D6*1/*10		CYP2D6*1/*4		CYP2D6*1/*5		CYP2D6*1/*6		CYP2D6*2/*41		CYP2D6*35/*41		CYP2D6*2/*10		CYP2D6*1/*41		CYP2D6*1/*3		CYP2D6*2/*4		CYP2D6*4/*35		CYP2D6*5/*35		CYP2D6*10/*35	
Yes				No				Yes				No				Yes				No					
endoxifen				View																					
4-hydroxytamoxifen				View				N-desmethyltamoxifen				View				tamoxifen				View					
endoxifen				View				tamoxifen N-oxide				View													
N-desmethyltamoxifen				View																					
endoxifen				View				4-hydroxytamoxifen				View				N-desmethyltamoxifen				View					
								tamoxifen				View													

Figure 4. Sample 'impact of variants' screen (partial view), with highlighted citation of interest. Display from e-PKGene (<http://www.pharmacogeneticsinfo.org>; accessed 18th November 2010).

Allowing searches by compound, gene or population enables the user quickly to focus on the desired parameters. Pharmaceutical researchers may find the information useful in designing studies on human subjects by highlighting populations (ethnic and genotypic) that will require more in-depth examination. Similarly, clinicians may find this platform valuable for identifying drugs that may require dosing adjustment in subjects with a known ethnic background, genotype or phenotype.

A summary section allows the user to access both gene and drug summaries. The gene summaries give a brief description of the gene, as well

as the classification of its alleles and genetic description. The drug summaries show the metabolism of the drug and its metabolic scheme, as well as a table view with the maximum percentage change in AUC and/or clearance for a drug (and its metabolites when available) encountered in particular genotype variants (Figure 9).

Domains of use

In spite of the wealth of information relating variants of metabolising enzymes and transporters to pharmacokinetic parameters, it is rare for a

University of Washington
e-PKGene *Impact of Genetics on Drug Exposure*

PMID 17947222 [Print](#) [PubMed](#)

CITATION
Title : Effects of CYP2D6 and SULT1A1 genotypes including SULT1A1 gene copy number on tamoxifen metabolism.
 Annals of oncology : official journal of the European Society for Medical Oncology / ESMO
 Volume 19, Issue 1, Pages 56-61
 Published 2008
 J Gjerde, M Hauglid, H Breilid, S Lundgren, JE Varhaug, ER Kisanga, G Mellgren, VM Steen, EA Lien
Comments:

Full Study Sets **Effects** **Impact on Pharmacokinetics**

Study 1	tamoxifen 20 mg oral daily dose for at least 80 days			
STUDY 1	86 Patients (reference) CYP2D6*1/*1; Homozygous Extensive Metabolizers; Caucasians; Females	49 Patients CYP2D6*1/*4; CYP2D6*1/*5; CYP2D6*1/*6; CYP2D6*1/*3; Heterozygous Extensive Metabolizers; Caucasians; Females	11 Patients CYP2D6*4/*4; CYP2D6*4/*5; Poor Metabolizers; Caucasians; Females	5 Patients CYP2D6*1/*2X2; Ultrarapid Metabolizers; Caucasians; Females
tamoxifen - Prodrug				
Impact	NO	NO	NO	NO
Study 2	tamoxifen 20 mg oral daily dose for at least 80 days			
STUDY 2	72 Patients (reference) SULT1A1*1/*1; Homozygous wild-type; Caucasians; Females	56 Patients SULT1A1*1/*2; Heterozygous variant; Caucasians; Females	23 Patients SULT1A1*2/*2; Homozygous variant; Caucasians; Females	
tamoxifen - Prodrug				
Impact	NO	NO	NO	NO

Figure 5. Sample 'citation display window' for PMID 17947222. Display from e-PKGene (<http://www.pharmacogeneticsinfo.org>; accessed 18th November 2010).

University of Washington
e-PKGene *Impact of Genetics on Drug Exposure*

PMID 17947222 [Print](#) [PubMed](#)

CITATION
Title : Effects of CYP2D6 and SULT1A1 genotypes including SULT1A1 gene copy number on tamoxifen metabolism.
 Annals of oncology : official journal of the European Society for Medical Oncology / ESMO
 Volume 19, Issue 1, Pages 56-61
 Published 2008
 J Gjerde, M Hauglid, H Breilid, S Lundgren, JE Varhaug, ER Kisanga, G Mellgren, VM Steen, EA Lien
Comments:

Full Study Sets **Effects** **Impact on Pharmacokinetics**

Study 1	tamoxifen 20 mg oral daily dose for at least 80 days			
STUDY 1	86 Patients (reference) CYP2D6*1/*1; Homozygous Extensive Metabolizers; Caucasians; Females	49 Patients CYP2D6*1/*4; CYP2D6*1/*5; CYP2D6*1/*6; CYP2D6*1/*3; Heterozygous Extensive Metabolizers; Caucasians; Females	11 Patients CYP2D6*4/*4; CYP2D6*4/*5; Poor Metabolizers; Caucasians; Females	5 Patients CYP2D6*1/*2X2; Ultrarapid Metabolizers; Caucasians; Females
tamoxifen - Prodrug				
Cmin (ng/mL) (Medians (Range))	90 (34-291)	89 (27-302)	82 (31-149)	97 (58-151)
Δ Cmin %		-1.11	-8.89	7.78
4-hydroxytamoxifen - active metabolite				
Cmin (ng/mL) (Medians (Range))	5.8 (2.2-17.2)	5.7 (1.7-11.3)	5.1 (3-8)	5.9 (5.7-12.7)
Δ Cmin %		-1.72	-12.07*	1.72
endoxifen - active metabolite				
Cmin (ng/mL) (Medians (Range))	52.3 (21.3-184.8)	49.6 (27.3-108.2)	36.7 (30.7-68.6)	46.3 (37.6-141.4)
Δ Cmin %		-5.16	-29.83*	-11.47
N-desmethyltamoxifen				
Cmin (ng/mL) (Medians (Range))	217 (90-596)	241 (156-691)	255 (142-481)	185 (127-369)
Δ Cmin %		11.06	17.51*	-14.75*
tamoxifen N-oxide				
Cmin (ng/mL) (Medians (Range))	8.9 (3.2-29)	9.5 (3.9-36.4)	10.5 (3.5-37.9)	10.5 (2.7-18.6)
Δ Cmin %		6.74	17.98	17.98

Figure 6. Sample 'impact of pharmacokinetics' screen for PMID 17947222. Display from e-PKGene (<http://www.pharmacogeneticsinfo.org>; accessed 18th November 2010).

University of Washington
e-PKGene Impact of Genetics on Drug Exposure

PMID 17947222 Print PubMed

CITATION
Title : Effects of CYP2D6 and SULT1A1 genotypes including SULT1A1 gene copy number on tamoxifen metabolism.
Annals of oncology : official journal of the European Society for Medical Oncology / ESMO
Volume 19, Issue 1, Pages 56-61
Published 2008
J Gjerd, M Hauglid, H Breilid, S Lundgren, JE Varhaug, ER Kisanga, G Mellgren, VM Steen, EA Lien
Comments:

Full Study Sets Effects Impact on Pharmacokinetics

Population
86 Patients (reference)
CYP2D6*1/*1;
Homozygous Extensive
Metabolizers;
Caucasians
Females
Breast cancer

Design and Drug Administration
Drug Administration: tamoxifen 20 mg oral daily dose
for at least 80 days .
Design: Multiple Dosing

Phenotyping Method
Taq-Man
Genotyping Method
PCR-Restriction fragment length polymorphism
(PCR-RFLP)
Long/Extra-Long PCR

Alleles Tested For
CYP2D6*2, CYP2D6*4, CYP2D6*5, CYP2D6*2X2,
CYP2D6*3, CYP2D6*6;

Pharmacokinetics

86 Patients (reference) CYP2D6*1/*1; Homozygous Extensive Metabolizers;	tamoxifen	4-hydroxytamoxifen	endoxifen	N-desmethyltamoxifen	tamoxifen N-oxide	N-desdimethyltamoxifen
Cmin (ng/mL) (Medians (Range))	90 (34-291)	5.8 (2.2-17.2)	52.3 (21.3-184.8)	217 (90-596)	8.9 (3.2-29)	37.7 (11.4-93.4)

Comments
All patients are treated adjuvantly with tamoxifen. CYP2D6 alleles that tested negative for all the variants were classified as CYP2D6*1. Poor Metabolizers: CYP2D6*4/*4 (n=8), CYP2D6*4/*5 (n=3) - Heterozygous EMs: CYP2D6*1/*3 (n=1), CYP2D6*1/*4 (n=43), CYP2D6*1/*5 (n=4), CYP2D6*1/*6 (n=1).

Figure 7. Sample 'full study set' screen for PMID 17947222. Display from e-PKGene (<http://www.pharmacogeneticsinfo.org>; accessed 18th November 2010).

genotype to be a critical determinant in selecting a drug dose.¹¹ While there are several instances of PGx and personalised medical applications, few have undergone the rigorous evaluations required for regulatory approval.¹² Even in cases where there

is adequate knowledge on genotype–drug response–phenotype correlations to appear in drug labels,¹³ PGx language is often for informational purposes only. This is because of what is described as the 'evidence dilemma', since there is often a

A

University of Washington
e-PKGene Impact of Genetics on Drug Exposure

Home Search Summaries

Search

Compound Go

(Example: You may search by a single term like "codeine" or "norfluoxetine")

Gene CYP2D6 Go

CYP2D6 Load list

Population Go

Load list

B

University of Washington
e-PKGene Impact of Genetics on Drug Exposure

Home Search Summaries

Search

Compound Go

(Example: You may search by a single term like "codeine" or "norfluoxetine")

Gene Go

Population Asians Go

Asians Load list

Caucasians

Pooled Asians

Figure 8. Examples of searches by (A) gene or (B) population. Display from e-PKGene (<http://www.pharmacogeneticsinfo.org>; accessed 18th November 2010).

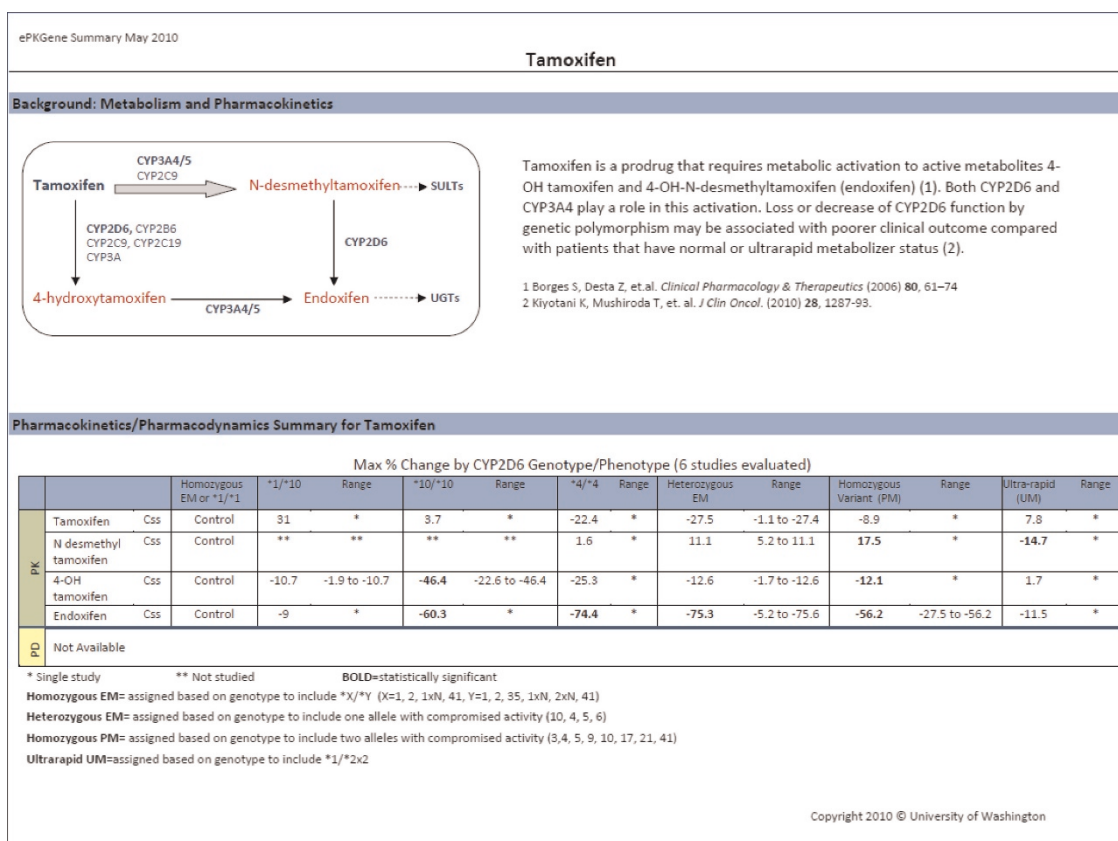


Figure 9. Example of drug summary from e-PKGene pilot version. Display from e-PKGene (<http://www.pharmacogeneticsinfo.org>; accessed 18th November 2010).

lack of sufficient evidence to weigh the benefits and risks of population screening or routine use of PGx applications and decision making in clinical practice.¹⁴ The web-based application e-PKGene emphasises the quantitative analysis of PGx data related to pharmacokinetics, pharmacodynamics and the safety of drugs in various populations. It provides access to information that is important to consider when evaluating genetic factors affecting drug exposure, populations affected and potential clinical relevance. The tool has the capacity to support all phases of translational research, but is of direct use to Phase 1 (T1) research, which seeks to move PGx discoveries into candidate health applications. Additionally, the tool will provide a portal for better understanding the scientific steps involved in applying PGx information and integrating genomics into drug development programmes:

for example, validating drug targets, choosing and validating lead compounds, identifying optimal early patient populations, determining surrogate endpoints (biomarkers) for the design of clinical trials and predicting likely variability in clinical trials. As the field of pharmacogenetics continues to expand and mature, the tool will shape into a large repository that will become invaluable in guiding both research and clinical practice decisions by providing primary literature findings necessary for establishing genotype–phenotype relationships.

References

1. Lesko, L.J. and Woodcock, J. (2004), ‘Translation of pharmacogenomics and pharmacogenetics: A regulatory perspective’, *Nat. Rev. Drug Discov.* Vol. 3, pp. 763–769.
2. Katz, D.A., Murray, B., Bhatena, A. and Sahelijo, L. (2008), ‘Defining drug disposition determinants: A pharmacogenetic-pharmacokinetic strategy’, *Nat. Rev. Drug Discov.* Vol. 7, pp. 293–305.

3. Wilkinson, G.R. (2005), 'Drug metabolism and variability among patients in drug response', *N. Engl. J. Med.* Vol. 352, pp. 2211–2221.
4. Aoki, M., Terada, T., Ogasawara, K., Katsura, T. *et al.* (2009), 'Impact of regulatory polymorphisms in organic anion transporter genes in the human liver', *Pharmacogenet. Genomics* Vol. 19, pp. 647–656.
5. Woodcock, J. and Lesko, L.J. (2009), 'Pharmacogenetics — Tailoring treatment for the outliers', *N. Engl. J. Med.* Vol. 360, pp. 811–813.
6. Khoury, M.J., Gwinn, M., Yoon, P.W., Dowling, N. *et al.* (2007), 'The continuum of translation research in genomic medicine: How can we accelerate the appropriate integration of human genome discoveries into health care and disease prevention?', *Genet. Med.* Vol. 9, pp. 665–674.
7. Daly, A.K. (2010), 'Genome-wide association studies in pharmacogenomics', *Nat. Rev. Genet.* Vol. 11, pp. 241–246.
8. Holmes, M.V., Shah, T., Vickery, C., Smeeth, L. *et al.* (2009), 'Fulfilling the promise of personalized medicine? Systematic review and field synopsis of pharmacogenetic studies', *PLoS One* Vol. 4, pp. e7960.
9. Single Nucleotide Polymorphism Database (dbSNP). <http://www.ncbi.nlm.nih.gov/projects/SNP/> (Accessed 1st November 2010).
10. Human Cytochrome P450 (CYP) Allele Nomenclature Committee. <http://www.cypalleles.ki.se/> (Accessed 1st November 2010).
11. Bhatena, A. and Spear, B.B. (2008), 'Pharmacogenetics: Improving drug and dose selection', *Curr. Opin. Pharmacol.* Vol. 8, pp. 639–646.
12. Deverka, P.A. (2009), 'Pharmacogenomics, evidence, and the role of payers', *Public Health Genomics* Vol. 3, pp. 149–157.
13. US Food and Drug Administration. (2010), Table of valid genomic biomarkers in the context of approved drug labels, 09/21. <http://www.fda.gov/Drugs/ScienceResearch/ResearchAreas/Pharmacogenetics/ucm083378.htm>, US Food and Drug Administration (Accessed 1st November 2010).
14. Khoury, M.J. (2010), 'Dealing with the evidence dilemma in genomics and personalized medicine', *Clin. Pharmacol Ther.* Vol. 87, pp. 635–638.