INTERNATIONAL CONFERENCE ON HARMONISATION OF TECHNICAL REQUIREMENTS FOR REGISTRATION OF PHARMACEUTICALS FOR HUMAN USE

DRAFT CONSENSUS GUIDELINE

TERMINOLOGY IN PHARMACOGENOMICS
E15

Current *Step 2* version

dated 25 October 2006

At Step 2 of the ICH Process, a consensus draft text or guideline, agreed by the appropriate ICH Expert Working Group, is transmitted by the ICH Steering Committee to the regulatory authorities of the three ICH regions (the European Union, Japan and the USA) for internal and external consultation, according to national or regional procedures.

E15
Document History

Current *Step 2* version

First Codification	History	Date	New Codification **November 2005**
E15	Approval by the Steering Committee under *Step 2* and release for public consultation.	25 October 2006	E15

TERMINOLOGY IN PHARMACOGENOMICS

Draft ICH Consensus Guideline

Released for Consultation on 25 October 2006, at *Step 2* of the ICH Process

TABLE OF CONTENTS

1. INTRODUCTION 1

1.1 Objective of the Guideline 1

1.2 Background 1

1.3 Scope of the Guideline 1

2. GUIDELINE 1

2.1 Genomic Biomarker 2

2.1.1 Definition 2

2.1.2 Additional Information 2

2.2 Pharmacogenomics and Pharmacogenetics 3

2.2.1 Definitions 3

2.2.1.1 Pharmacogenomics 3

2.2.1.2 Pharmacogenetics 3

2.2.2 Additional Information 3

2.3 Categories for Genomic Data and Samples Coding 3

2.3.1 Identified Data and Samples 3

2.3.2 Coded Data and Samples 4

2.3.2.1 Single Coded Data and Samples 4

2.3.2.2 Double-Coded Data and Samples 4

2.3.3 Anonymised Data and Samples 4

2.3.4 Anonymous Data and Samples 5

2.3.5 Additional Information 5

Table 1: Summary of Genomic Data and Sample Coding Categories 6

TERMINOLOGY IN PHARMACOGENOMICS

1. INTRODUCTION

1.1 Objective of the Guideline

In order to develop harmonised approaches to drug regulation it is important to ensure that consistent definitions of terminology are being applied across all constituents of the International Conference on Harmonisation (ICH). Agreement on definitions will facilitate the integration of the discipline of pharmacogenomics and pharmacogenetics into global drug development and approval processes.

1.2 Background

Pharmacogenomics and pharmacogenetics have the potential to improve the discovery, development and use of medicines. Each of the ICH regions has published pharmacogenomic and pharmacogenetic specific guidances, or concept papers and is in the process of developing others. However, the lack of consistently applied definitions to commonly used terminology raises the potential for conflicting use of terms in regulatory documentation and guidances or inconsistent interpretation by regulatory authorities, ethics committees and sponsor companies.

1.3 Scope of the Guideline

This guideline contains definitions of key terms in the discipline of pharmacogenomics and pharmacogenetics, namely genomic biomarkers, pharmacogenomics, pharmacogenetics and genomic data and sample coding categories. Validation and qualification processes for genomic biomarkers, evidence for their intended use and acceptance criteria across ICH regions are outside of the scope of this guideline. As new scientific knowledge in the discipline of pharmacogenomics and pharmacogenetics emerges, the current guidance will be reviewed and expanded if appropriate.

2. GUIDELINE

Definitions of a genomic biomarker, pharmacogenomics, pharmacogenetics, and genomic data and sample coding categories are detailed below. The definition of what constitutes a genomic biomarker is key to understanding the definitions of pharmacogenomics and pharmacogenetics and is therefore introduced in this guideline first. Additional information useful to the understanding of aspects covered by each of the definitions is also provided. Some of the principles described in this guideline might be applicable to proteomics, metabonomics and other related disciplines.

2.1 Genomic Biomarker

2.1.1 Definition

A genomic biomarker is defined as:

A measurable DNA or RNA characteristic that is an indicator of normal biologic processes, pathogenic processes, and/or response to therapeutic or other intervention.

2.1.2 Additional Information

1. The definition for a genomic biomarker is not limited to human samples;
2. A genomic biomarker could, for example, reflect:
    - The expression of a gene;
    - The function of a gene;
    - The regulation of a gene.
3. A genomic biomarker can consist of one or more deoxyribonucleic acid (DNA) or ribonucleic acid (RNA) characteristics;
4. The definition for a genomic biomarker does not include the measurement and characterisation of proteins or low molecular weight metabolites;
5. DNA characteristics include, but are not limited to:
    - Single nucleotide polymorphisms (SNPs);
    - Variability of short sequence repeats;
    - DNA modification, e.g., methylation;
    - Insertions;
    - Deletions;
    - Copy number variation;
    - Cytogenetic rearrangements, e.g., translocations, duplications, deletions or inversions.
6. RNA characteristics include, but are not limited to:
    - RNA sequence;
    - RNA expression levels;
    - RNA processing, e.g., splicing and editing;
    - MicroRNA levels.

2.2 Pharmacogenomics and Pharmacogenetics

2.2.1 Definitions

2.2.1.1 Pharmacogenomics

Pharmacogenomics (PGx) is defined as:

The investigation of variations of DNA and RNA characteristics as related to drug response.

2.2.1.2 Pharmacogenetics

Pharmacogenetics (PGt) is a subset of pharmacogenomics and is defined as:

The influence of variations in DNA sequence on drug response.

2.2.2 Additional Information

1. PGx and PGt are applicable to activities such as drug discovery, drug development, and clinical practice;
2. Drug response includes drug disposition (pharmacokinetics, PK) and drug effect (pharmacodynamics, PD);
3. The term drug should be considered synonymous with investigational (medicinal) product, medicinal product and pharmaceutical (including vaccines and other biological products);
4. The definition of PGx and PGt does not include other disciplines such as proteomics and metabonomics.

2.3 Categories for Genomic Data and Samples Coding

PGx and PGt research depends on the use of samples to generate data. A harmonised definition for the coding of these samples and their associated data will facilitate use in research and development of new medicines. There are four general categories of coding: identified, coded, anonymised and anonymous. Coded data or samples can be single or double coded. The implications of using a specific data and sample coding category should be considered in the design of PGx and PGt research studies. Some of these implications are highlighted in this section and summarised in Table 1.

2.3.1 Identified Data and Samples

Identified data and samples are labelled with personal identifiers such as name or identification numbers (e.g., social security or national insurance number). As the samples and associated data are directly traceable back to the subject, it is possible to undertake actions such as sample withdrawal or the return of individual results. The use of identified data and samples allows for clinical monitoring, subject follow-up and the addition of new data. Identified data and samples coding offers privacy protection similar to general healthcare confidentiality in everyday medical practice. Identified data and sample coding are generally not considered appropriate for purposes of clinical trials in drug development.

2.3.2 Coded Data and Samples

Coded data and samples are labelled with at least one unique code and do not carry any personal identifiers.

2.3.2.1 Single Coded Data and Samples

Single coded data and samples are labelled with a single code and do not carry any personal identifiers. It is possible to trace the data or samples back to the individual with the use of a single coding key. In general, the clinical investigator is responsible for maintaining the coding key. As the samples and associated data are indirectly traceable back to the subject via the coding key, it is possible to undertake actions such as sample withdrawal, or the return of individual results. The use of single coded data and samples allows for clinical monitoring, subject follow-up, or the addition of new data. Single coding is the current standard used in clinical research and offers additional safeguards to the subject's identifiers compared to general healthcare confidentiality and privacy protection in everyday medical practice.

2.3.2.2 Double-Coded Data and Samples

Double coded data and samples are initially labelled with a single code and do not carry any personal identifiers. The data and samples are then relabelled with a second code, which is linked to the first code via a second coding key. It is possible to trace the data or samples back to the individual by the use of both coding keys. In general, the clinical investigator is responsible for maintaining the first coding key and does not have access to the second coding key. As the samples and associated data are indirectly traceable back to the subject via the use of both coding keys, it is possible to undertake actions such as sample withdrawal, or the return of individual results. The use of double coded data and samples allows for clinical monitoring, subject follow-up or the addition of new data. The use of the second code provides additional confidentiality and privacy protection for subjects over that of single coded. Access to both coding keys is needed to link any data or samples back to a subject identifier.

2.3.3 Anonymised Data and Samples

Anonymised data and samples are initially single or double coded but the link between the subjects identifiers and the unique code(s) are subsequently deleted. Once the link has been deleted it is no longer possible to trace the data and samples back to the individual through the coding key(s). Anonymisation is intended to prevent subject re-identification. As anonymised samples and associated data are not traceable back to the subject it is not possible to undertake actions such as sample withdrawal, or the return of individual results. The use of anonymised data and samples does not allow for clinical monitoring, subject follow-up or the addition of new data. The deletion of the coding key(s) linking the data and samples to the subject's identifiers provides additional confidentiality and privacy protection over coded data and samples, as it prevents subject re-identification through the use of the coding key(s).

2.3.4 Anonymous Data and Samples

Anonymous data and samples are never labelled with personal identifiers and therefore there is no potential to trace back genomic data and samples to an individual subject. In some instances only limited clinical data is associated with anonymous samples. As anonymous samples and associated data are not traceable back to the subject, it is not possible to undertake actions such as sample withdrawal, or the return of individual results. The use of anonymous data and samples does not allow for clinical monitoring, subject follow-up, or the addition of new data.

2.3.5 Additional Information

The use of a specific coding category in relation to obtaining informed consent from subjects is not within the focus of this guideline and is not being addressed herein. The conditions under which the genomic data can be linked back to the subject's personal identifiers for any purpose should, however, be described in the research related documents, e.g., informed consent document.

Table 1: Summary of Genomic Data and Sample Coding Categories

Sample Coding Category		**Link Between Subject's Personal Identifiers and Genomic Data**	**Traceability Back to the Subject (Actions possible, e.g., sample withdrawal or return of individual genomic results)**	**Ability to Perform Clinical Monitoring, Subject Follow-up, or Addition of New Data**	**Extent of Subject's Confidentiality and Privacy Protection**
Identified		Yes (direct)	Possible	Possible	Similar to general healthcare confidentiality and privacy
Coded	***Single***	Yes (indirect via coding key)	Possible	Possible	Standard for clinical research
	Double	Yes (indirect via the two coding keys)	Possible	Possible	Added privacy and confidentiality protection over single code
Anonymised		No as coding key(s) have been deleted	Not possible as coding key(s) have been deleted	Not possible as coding key(s) have been deleted	Genomic data and samples not linked to subject as coding key(s) have been deleted
Anonymous		No	Not possible	Not possible	Genomic data and samples never linked to subject