

# 人用药品注册技术要求国际协调会

## ICH 三方协调指导原则

基因组生物标志物、药物基因组学、遗传药理学、基因组数据和样本

编码分类的定义

**E15**

现行第四阶段版本

2007 年 11 月 1 日

该指导原则由相应的 ICH 专家小组制定，按照 ICH 进程，已递交管理部门讨论。在 ICH 进程第四阶段，最终草案被推荐给欧盟、日本和美国的管理机构采纳。

## E15

### 文件历史

代码*	历史	日期
E15	指导委员会批准进入第二阶段，并发布以公开征询意见。	2006年10月25日

### ICH 进程第四阶段

E15	指导委员会批准进入第四阶段，并推荐给 ICH 三方的管理机构采纳。	2007年11月1日
-----	-----------------------------------	------------

\* 于 2007 年 11 月经 ICH 指导委员会采纳作为新编码系统代码

# 基因组生物标志物、药物基因组学、遗传药理学、基因组数据和样本

## 编码分类的定义

### ICH 三方协调指导原则

在 2007 年 11 月 1 日 ICH 指导委员会会议上进入 ICH 进程第四阶段，本指导原则已被推荐给 ICH 三方的管理当局采纳。

#### 目录

1. 前言.....	4
1.1 本指导原则的目的.....	4
1.2 背景.....	4
1.3 本指导原则的范围.....	4
2. 指导原则.....	4
2.1 基因组生物标志物.....	4
2.1.1 定义.....	4
2.1.2 附加信息.....	5
2.2 药物基因组学和遗传药理学.....	5
2.2.1 定义.....	5
2.2.2 附加信息.....	6
2.3 基因组数据和样本编码分类.....	6
2.3.1 标识类数据和样本.....	6
2.3.2 编码类数据和样本.....	6
2.3.3 匿名类数据和样本.....	7
2.3.4 佚名类数据和样本.....	8
2.3.5 附加信息.....	8
表 1: 基因组数据和样本编码分类总结.....	9

# 基因组生物标志物、药物基因组学、遗传药理学、基因组数据和样本 编码分类的定义

## 1. 前言

### 1.1 本指导原则的目的

为了协调药物管理，确保国际协调会（ICH）所有成员国均采用统一的名词术语定义是非常重要的。定义的统一将推动药物基因组学和遗传药理学整合到全球药物开发和批准程序中去。

### 1.2 背景

药物基因组学和遗传药理学能促进药物的发现、开发和使用。ICH 的每个成员国均已出版了药物基因组学和遗传药理学的特定指导原则或概念性论文，且其他文件的制定也都在进行中。然而，缺乏常用术语的统一定义会造成管理文件和指导原则之间术语使用的冲突或管理机构、伦理委员会和申办公司间对其解释的不一致。

### 1.3 本指导原则的范围

本指导原则包括药物基因组学和遗传药理学中关键术语，即基因组生物标志物、药物基因组学、遗传药理学、基因组数据和样本编码分类的定义。本指导原则不包括基因组生物标志物的验证和确认过程、预期用途的证据以及 ICH 范围内验收标准方面的问题。由于药物基因组学和遗传药理学中不断出现新的科学知识，因此必要时，本指导原则将被审核并做补充完善。

## 2. 指导原则

基因组生物标志物、药物基因组学、遗传药理学、基因组数据和样本编码分类的定义在下面会有详细的描述。基因组生物标志物构成的定义是理解药物基因组学和遗传药理学的关键，因此在本指导原则中被首先介绍。本指导原则还提供有助于理解每种定义所涵盖内容的附加信息。本指导原则中描述的某些原则可能还适用于蛋白组学、代谢物组学及其他相关学科。

### 2.1 基因组生物标志物

#### 2.1.1 定义

基因组生物标志物定义如下：

一种可测量的 DNA 和/或 RNA 特征，是正常生物过程、病理过程和/或对治疗或其他干预的反应的一种指示物。

### 2.1.2 附加信息

1. 基因组生物标志物可以是某种检测指标，如：
  - 基因的表达；
  - 基因的功能；
  - 基因的调控。
2. 基因组生物标志物可以具有一种或多种脱氧核糖核酸（DNA）和/或核糖核酸（RNA）特征。
3. DNA 特征包括，但不限于：
  - 单核苷酸多态性（SNPs）；
  - 短重复序列变异；
  - 单倍体；
  - DNA 修饰，如甲基化；
  - 单个核苷酸缺失或插入；
  - 拷贝数量异；
  - 细胞遗传重组，如易位、重叠、缺失或倒置。
4. RNA 特征包括，但不限于：
  - RNA 序列；
  - RNA 表达水平；
  - RNA 转录后加工，如剪接和编辑；
  - microRNA 水平。
5. 基因组生物标志物的定义并不限于人体样本，还包括病毒和病原体以及动物样本，即基因组生物标志物适用于所有非临床和/或毒理学研究。
6. 基因组生物标志物的定义不包括蛋白质或低分子量代谢物的检测和特征描述。

## 2.2 药物基因组学和遗传药理学

### 2.2.1 定义

#### 2.2.1.1 药物基因组学

药物基因组学 (PGx) 被定义为:

**对药物反应相关的 DNA 和 RNA 特征变化的研究。**

#### 2.2.1.2 遗传药理学

遗传药理学 (PGt) 是药物基因组学 (PGx) 的分支, 被定义为:

**对药物反应相关的 DNA 序列变化的研究。**

### 2.2.2 附加信息

1. 术语“药物 (drug)”被认为与研究用产品 (investigational product)、医药产品 (medicinal product)、药品 (medicine) 和药剂产品 (pharmaceutical product) (包括疫苗和其他生物制品) 为同义词。
2. PGx 和 PGt 适用于药物发现、药物开发和临床实践的领域。
3. 药物反应包括药物吸收和处置过程 (如药代动力学, PK) 和药物效应 (如药效学(PD), 药物有效性和不良反应)。
4. PGx 和 PGt 的定义不包括蛋白质组学和代谢物组学等其他学科。

### 2.3 基因组数据和样本编码分类

PGx 和 PGt 研究依赖于使用生物学样本来获取数据。这些样本及其相关数据代码的定义一致, 将有助于其在新药研究和开发中的应用。

编码通常分为四类: 标识类、编码类、匿名类和佚名类。编码数据或样本可以是单一代码或双重代码。

使用特异性数据和样本编码分类的含义, 应当在 PGx 和 PGt 研究设计中有所考虑。

某些含义会在本章节中着重强调, 在表 1 中进行总结。

#### 2.3.1 标识类数据和样本

标识类数据和样本以个人识别信息, 如姓名或身份号码 (如社会保障卡号或社会保险号) 进行标记。由于样本及其相关数据可直接追溯到受试者本人, 因此能够根据受试者的要求撤除样本或归还个人结果。。使用标识类数据和样本可进行临床监查、受试者随访或添加新的受试者数据。标识类数据和样本可提供相当于日常医疗实践中常规医疗保密的个人隐私保护。通常认为标识类数据和样本不适用于药物开发中的临床试验。

#### 2.3.2 编码类数据和样本

编码类数据和样本以至少一种特定代码进行标记且不含有任何个人识别信息。

### **2.3.2.1 单一编码类数据和样本**

单一编码类数据和样本通常以单一特定代码标记且不含有任何个人识别信息。通过使用单一编码索引有可能将数据或样本追溯到特定个体。通常，编码索引由临床研究者负责保管。由于样本及相关数据能通过编码索引间接追溯到受试者，因此，可根据受试者的要求撤除样本或归还个人结果。使用单一编码数据和样本可进行临床监查、受试者随访或添加新的受试者数据。单一编码是当前临床研究的使用标准，可为受试者的身份识别提供相当于日常医疗实践中常规医疗保密和个人隐私保护的额外防护。。

### **2.3.2.2 双重编码类数据和样本**

双重编码类数据和样本最初以单一特定代码标记且不含有任何个人识别信息。数据和样本随后接受第二代码的再标记，后者通过第二编码索引与首次密码链接。通过两个编码索引可以将数据或样本追溯到受试者本人。通常，临床研究者负责维护首次编码索引，而无权使用第二编码索引。由于样本及相关数据能通过两个编码索引间接追溯到受试者，因此，可根据受试者的要求撤除样本或归还个人结果。但是，可通过添加电子或技术流程来进一步限制将基因型结果追溯到受试者的能力。例如，可使用一种允许添加新的受试者信息，但却阻止基因型数据再连接到受试者个人识别信息的特异性计算机流程。使用双重编码数据和样本可进行临床监查、受试者随访或添加新的受试者信息。使用第二编码比单一编码能提供更多的保密性和个人隐私保护。只有使用两种编码索引，才能将数据或样本结果追溯到受试者的识别信息上。

### **2.3.3 匿名类数据和样本**

匿名类数据和样本最初以单一或双重编码标记，但随后完全删除受试者识别信息与唯一代码间的联系。一旦联系被删除，数据和样本将不可能再通过编码索引追溯到受试者本人。匿名的目的是阻止受试者再被识别。由于匿名类样本及相关数据不能再追溯到受试者本人，因此，不可能根据受试者的要求撤除样本或归还个人结果。使用匿名类数据和样本也无法进行临床监查、受试者随访或添加新的受试者信息。删除连接数据和样本与受试者识别信息间的编码索引，由于能阻

止通过使用编码索引对受试者进行再次识别，故可以为编码数据和样本提供更多的保密性和个人隐私保护。

#### **2.3.4 佚名类数据和样本**

佚名类数据和样本自最初收集起就不进行个人识别标记，也不生成编码索引。因此，不可能将基因组数据和样本追溯到受试者本人。在某些情况下，仅有限的临床数据（如，样本来自于男性、年龄 50~55 岁、胆固醇>240 mg/dL 的糖尿病受试者）可以采用佚名样本。由于佚名样本及相关数据无法追溯到受试者本人，因此不可能根据受试者的要求撤除样本或归还个人结果。使用佚名数据和样本也无法进行临床监查、受试者随访或添加新的受试者信息。

#### **2.3.5 附加信息**

对已获得知情同意的受试者所使用的特异性编码分类不被包括在本指导原则范围内，因此在此处不予讨论。

在研究相关文件（如知情同意书）中应当描述，在何种条件下，出于何种目的（包括将基因组数据归还给受试者），基因组数据可被追溯到受试者本人。

表 1: 基因组数据和样本编码分类总结

样本编码分类		连接受试者个人识别信息和基因组生物标志物数据	追溯到受试者本人（可能的操作，包括应受试者要求撤回样本或归还个人基因组结果）	执行临床监查、受试者随访或添加新数据的能力	受试者保密性和个人隐私保护的程 度
标识类		是（直接） 允许受试者被识别	是	是	相当于普通医疗保密和隐私保护
编码类	单一	是（间接） 允许受试者被识别（通过单一、特异性编码索引）	是	是	临床研究标准
	双重	是（非常间接） 允许受试者被识别（通过 2 个特异性编码索引）	是	是	比单一编码提供更多的保密性和个人隐私保护
匿名类		否 由于编码索引被删除，无法对受试者进行再识别	否	否	由于编码索引被删除，基因组数据和样本无法追溯到受试者
佚名类		否 不收集识别信息、不使用编码索引，无法对受试者进行再识别	否	否	基因组数据和样本不能与受试者联系