

T/SZGIA

/

-

1

前	II
引	III
1 围	1
2 性引 文件	1
3 术 和定义	1
4	5
5 数据格式属性与描 则	5
6 数据格式 明 方式	7
7 归档 录属性及描 则	8
8 数据元属性与描 则	10
9 数据元值域 方	14

《基因检测产品数据标准》包括 标准和 定量检测产品数据标准，如：

—— 1 分： 标准；

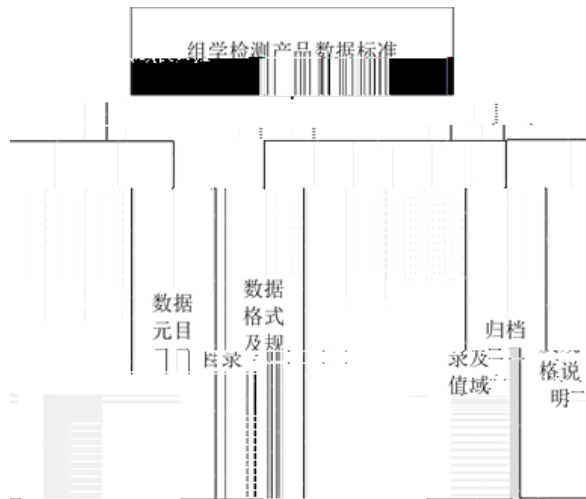
—— 2 分：孕妇外周 儿 DNA产前检测元数据 录；

本标准按 GB/T 1.1-2009 出 则 。

本 分 单位： 圳华大基因 技有 公司、 圳华大 命 学 、 圳华大临床检 中心、 圳华大基因 份公司、 圳基因产学 、北京 技 份有 公司、广州医 大学 属 三医 、 圳 奥康晨 技有 公司、 基因 技（ 圳）有 公司、 圳 技有 公司。

本 分主 人：吕春杰、刘小 、唐 、李 、 奇、李倩一、吴昊、李 强、吴俊、 大伟、 、 敏、 晨晴、杨旭、 、李 、 新华、 志博。

学数据 数据 型可分为 构化数据和 构化数据。其中 构化数据， 数据格式、数据格式 格 明、归档 录描 ； 构化数据， 数据元、值域来描 。



数据格式及 则 明 了数据格式 内容 构、属性与描 则、数据格式 制 则。

归档 录 了归档 录 内容 构、属性与描 则、格式和 引 制 则。

数据元 化定义了数据元 方 及描 属性，描 属性包括 态、来 、基 数据 、基 数据元标 、数据元中文名 、定义、必 性、信息保护、数据元 数据 型、 格式、单位代 、数据元允 值 。值域代 了数据元值域 方 、代 格式和 、代 命名与标 。

本分定了基因学产品数据属性与描述、数据元引与数据元值域方、代格
式与、代命名与标、数据格式内容构与制、数据格式格明、归档
录构内容。

本文件于基因学产品数据标准制。

下列文件对于本文件应是必不可少。凡是日期引文件,仅日期本于本文件。
凡是不日期引文件,其最新本(包括所有修改单)于本文件。

2312 信启、交换字字基本

/ 740 数据元和交换格式信息交换日期和时

/ 10113 分与术

/ 172 5 国易单位代

/ 1 3 1.1 信息技术元数据册() 1 分: 框架

/ 1 3 1.信息技术元数据册() 3 分: 册元型和基本属性

/ 1 4 .1 子政务数据元 1 分: 和

/ 303 卫信息数据元标准化则

/ 305 卫信息数据元数据

/ 306 卫信息数据分与则

363.1 卫信息数据元录 1 分: 总则

364.1 卫信息数据元值域代 1 分: 总则

/ 6 7.1 交信息基数据元 1 分: 总则

数据元录中为数据元分与无关唯一标。

: 定义于国家卫业标准 363.1中3.1。

在数据元录中数据元属性值均同属性。如本标准中册机构。

： 定义 于国家卫 业标准 363.1中3.2。

在数据元 录中数据元属性值不 同 属性。

： 定义 于国家卫 业标准 363.1中3.3。

允 值 合。

： 定义 于国家标准 / 1 3 1.1中 3.75。

具有某 共同属性（或 征） 事 （或 念） 合。

： 定义 于国家标准 / 10113中 2.1.1。

按 定 属性（或 征）区分分 对 ，并将具有某 共同属性（或 征） 分 对 合在一
。

： 定义 于国家卫 业标准 364.1中3.3。

将分 对 按 定 干属性（或 征） 地分为 干层 ， 个层 又分为 干 ，不同层
之 构成 属关 。 分 方 为 分 。

： 定义 于国家卫 业标准 364.1中3.4。

定分 对 干属性（或 征），将分 对 按 属性（或 征）划分成一 ，
一 构成一个 。再按一定 序将各个 平 排列。使 时根据 将有关 中 应
按 指定排列 序 在一 ，形成一个新 复合 。 分 方 为 分 。

： 定义 于国家卫 业标准 364.1中3.5。

定事 （或 念） 一个或一 字 。 些字 可以是 拉伯数字、拉丁字 或便于 子
机和人 别与处 其他 号。

： 定义 于国家标准 / 10113中 2.2.5。

事 （或 念） 予代 。

： 定义 于国家卫 业标准 364.1中3.7。

一个完整代 成方式和 度 合 。

： 定义 于国家卫 业标准 364.1中3. 。

从某一个方 （如含义、 构、 度、 成 ）来 代 某 性。

如：从含义上可分为有含义代 和无含义代 ；从 构上可分为层 和 序 ；从 度上可分
为 代 和不 代 ；从 成上可分为数字代 和字 代 。

： 定义 于国家卫 业标准 364.1中3. 。

对 对 只 标 作 ， 无任何其他 加含义 代 。

： 定义 于国家卫 业标准 364.1中3.10。

对 对 标 作 外， 具有其他 定含义 代 。

： 定义 于国家卫 业标准 364.1中3.11。

拉伯数字（0 ）构成 代 。

： 型 代 仅仅是以 拉伯数字 形式 ，但不是数值型，不可 接 于 。

： 定义 于国家卫 业标准 364.1中3.12。

字 构成 代 。

： 其中所 字 常为 文字 （、 因与1、0 似， 常不使 ）。

： 定义 于国家卫 业标准 364.1中3.13。

字和拉伯数字合构成代码。

：定义于国家标准 364.1中3.14。

以对属性为排列顺序成有层关系代码。

：定义于国家标准 364.1中3.15。

按拉伯数字或字母顺序来对代码。亦。

1：通常情况下，顺序是，代之不出，断。但在情况下，可（式）。

2：定义于国家标准 364.1中3.16。

根据对属性（或征）同或似，将对分为干。再将序分为应干列（也为），并分别各对。在同一列内对对，并扩展。样制代列为列序。

：定义于国家标准 364.1中3.17。

在同一个代码体中，所有对代码度。

：定义于国家标准 364.1中3.1。

在一个完整代码体中，代总度不完全同。

：定义于国家标准 364.1中3.1。

在定境中，可唯一性地标与之关事一系列字，可做来别定对数据元值。

：定义于国家标准 364.1中3.20。

归档录构中为归档录分与无关唯一标志。

在归档录构中数据录属性值均同属性。如本标准中册机构。

在归档录构中数据录属性值不同属性。

数据元 () ;
 数据标 () ;
 核核 () ; A
 主机构 () ; A A
 册机构 () ;
 册机构标 () ;
 提交机构 () ;
 本标 () 。

数据格式属性参 / 303, 一定 5 14 属性, 并按性度分为两: 数据元公属性和数据元专属性。数据元公属性包括7, 数据元专属性包括7, 1。

1

序号	属性	数据元属性名	束	备
1	标	数据格式标	必	专属性
2		数据格式名	必	专属性
4		本	必	共属性
5		册机构	必	共属性
6		关境	必	共属性
7		定义	围	必
	关	分式	必	共属性
11		数据格式允值	必	专属性
12		主机构	必	共属性
13		册态	必	共属性
14		提交机构	必	共属性

数据元()标 字 数字 合 , 包含数据标 ()和 本标 ()两 构。

例1: 1

) 按 分 和 号 合 方式, 字 数字 合 。按 数据元对应 主 分 代 、 大 代 、小 代 、 序 、 加 从左向右 序排列。其中:

主 分 代 : 2位大写 文字 。代 一定为 。

大 代 : 2位数字 , 数字大小无含义。

小 代 : 2位数字 , 数字大小无含义; 无小 时则小 代 为00。小 与大 代 之 加 . 区分。

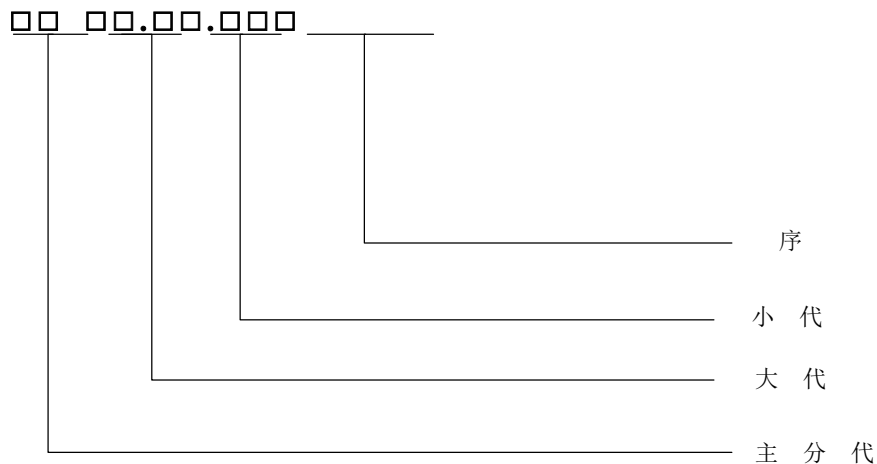
序 : 3位数字 , 代 某一小 下 数据元序号, 数字大小无含义; 从001开始 序 。 序 与小 代 之 加 . 区分。

) 构 4 分 成, 为 + .. + . + .. 。其中 .. 和 .. 为两位 拉伯数字构成, 在数学上应是具有意义 整数。 .. 主 本号, .. 本号。

例2: 1.2 主 本为 一 , 本为 二 。

如果数据元更新前后可以 有效 数据交换, 则更新后主 本号不变, 本号 于当前 本号加1; 如果数据元更新前后无 有效 数据交换, 则更新后主 本号 于当前主 本号加1, 本号归0。

数据标 构 图2。



2

数据元 中文名 应当是唯一 , 并且以字 、 字、数字式 字 串形式 。

数据元 命名应使 一定 构和 术 。

完整 数据元名 对 术 + 性 术 + 术 +(定 术) 。

其中:

个数据元 有一个且仅有一个对 术 。在 学数据元 录中 对 术 为 本人 , 则可 情 。

一个数据元 有一个且仅有一个 性 术 。 性 术 是任何一个数据元名 所必
成分，在数据元 念可以完整、准家无 义 情况下，其他术 可以 情 。

一个数据元 有一个且仅有一个 术 。当 术 与 性 术 有 复或 分
复时，可从名 中将冗余 删 。 术 2。

定 术 专业 域 定。 定 术 是可 。

2

	含义
名	一个对 一个 或
代	替代某一 定信息 一个有内在 则 字 串（字 、数字、 号）
明	描 对 信息 一

) 构 4 分 成, 为 $+ .. + . + ..$ 。其中 .. 和 .. 为两位 拉伯数字构成, 在数学上应是具有意义 整数。 .. 主 本号, .. 本号。

例2: 1.2 主 本为 一 , 本为 二 。

如果数据元更新前后可以 有效 数据交换, 则更新后主 本号不变, 本号 于当前 本号加1; 如果数据元更新前后无 有效 数据交换, 则更新后主 本号 于当前主 本号加1, 本号归0。

数据标 构 图3。

数据归档 录 () 标 字 数字 合 , 包含数据标 () 和 本标 () 两 构。

例1: 1

) 按 分 和 号 合 方式, 字 数字 合 。按 数据元对应 主 分 代 、 大 代 、小 代 、 序 、 加 从左向右 序排列。其中:

主 分 代 : 2位大写 文字 。代 一定为 。

大 代 : 2位数字 , 数字大小无含义。

小 代 : 2位数字 , 数字大小无含义; 无小 时则小 代 为00。小 与大 代 之 加 . 区分。

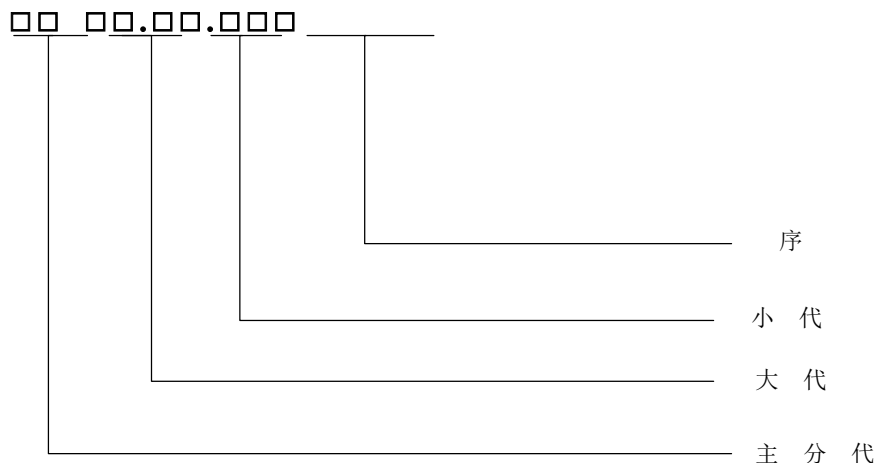
序 : 3位数字 , 代 某一小 下 数据元序号, 数字大小无含义; 从001开始 序 。 序 与小 代 之 加 . 区分。

) 构 4 分 成, 为 + .. + . + .. 。其中 .. 和 .. 为两位 拉伯数字构成, 在 数学上应是具有意义 整数。 .. 主 本号, .. 本号。

例2: 1.2 主 本为 一 , 本为 二 。

如果数据元更新前后可以 有效 数据交换, 则更新后主 本号不变, 本号 于当前 本 号加1; 如果数据元更新前后无 有效 数据交换, 则更新后主 本号 于当前主 本号加1, 本号归0。

数据标 构 图4。



4

数据元 中文名 应当是唯一 , 并且以字 、 字、数字式 字 串形式 。

数据元 命名应使 一定 构和 术 。

完整 数据元名 对 术 + 性 术 + 术 + (定 术) 。

其中:

个数据元 有一个且仅有一个对 术 。在 学数据元 录中 对 术 为 本人 , 则可 情 。

个数据元 有一个且仅有一个 性 术 。性 术 是任何一个数据元名 所必 成 分, 在数据元 念可以完整、准家 无 义 情况下, 其他术 可以 情 。

一个数据元 有一个且仅有一个 术 。当 术 与 性 术 有 复 或 分 复 时，可从名 中将冗余 删 。 术 4。
定 术 专业 域 定。定 术 是可 。

4

	含义
名	一个对 一个 或
代	替代某一 定信息 一个有内在 则 字 串（字 、数字、 号）
明	描 对 信息 一 文字
	以 币为 单位 数 ， 常与 币 型有关
数	币单位数 ， 常与 单位有关。 单位参 录 .1， 定构成十 倍数和分数单位 头 录 .2
日期	以公元 年方式 年、月、日 合
时	以24小时制 时方式 天中 小时、分、 合
日期时	完整时 格式，即 15， 格式
分	具有 同 单位 两个值之 分数形式
	一个 或 与另一个 A 或
标志	又 指 ， 两个且只有两个 ^A 明条件 值，如：是/否、 ^A 有/无
时	两个时 时 度

数据元属性 参 / 303， 一 定 5 14 属性，并按 性 度分为两 ：数据元公 属性和数据元专 属性。数据元公 属性包括7 ，数据元专 属性包括7 ， 5。

5

序号	属性	数据元属性名	束	备
1	标	数据元标	必	专 属性
2		数据元名	必	专 属性
3		信息保护	可	专 属性
4		本	必	共 属性
5		册机构	必	共 属性
6		关 境	必	共 属性
7	定义	定义	必	专 属性
	关	分 式	必	共 属性

		数据元值 数据类型	必	专 属性
10		格式	必	专 属性
11		数据元允 值	必	专 属性
12		主 机构	必	共 属性
13		册 态	必	共 属性
14		提交机构	必	共 属性

数据元()标 字 数字 合 , 包含数据标 ()和 本标 ()两 构。

例1: 1

) 按 分 和 号 合 方式, 字 数字 合 。按 数据元对应 主 分 代、大 代、小 代、序、加 从左向右 序排列。其中:

主 分 代: 2位大写 文字 。代 一定为 。

大 代: 2位数字 , 数字大小无含义。

小 代: 2位数字 , 数字大小无含义; 无小 时则小 代 为00。小 与大 代 之 加 . 区分。

序: 3位数字 , 代 某一小 下 数据元序号, 数字大小无含义; 从001开始 序 。 序 与小 代 之 加 . 区分。

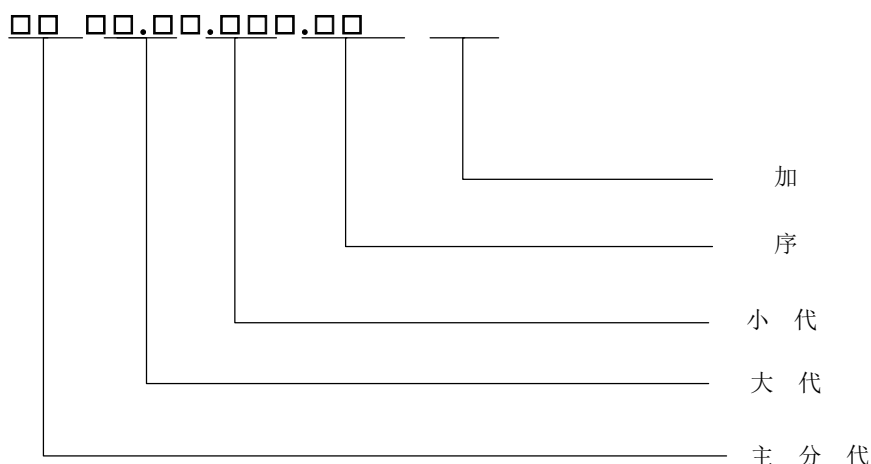
加: 2位数字 , 代 一 数据元 关 ; 从01开始 序 , 加 与 序号之 加 . 区分。无 关 数据元其 加 为 00 。

) 构 4 分 成, 为 + .. + . + .. 。其中 .. 和 .. 为两位 拉伯数字构成, 在数学上应是具有意义 整数。 .. 主 本号, .. 本号。

例2: 1.2 主 本为一 , 本为 二 。

如果数据元更新前后可以 有效 数据交换, 则更新后主 本号不变, 本号 于当前 本号加1; 如果数据元更新前后无 有效 数据交换, 则更新后主 本号 于当前主 本号加1, 本号归0。

数据标 构 图5。



数据元中文名应当是唯一，并且以字、字、数字式字符串形式。

数据元命名应使一定结构和术。

完整数据元名对术 + 性术 + 术 + (定术)。

其中：

一个数据元有一个且仅有一个对术。在学数据元录中对术为本人，则可情。

一个数据元有一个且仅有一个性术。性术是任何一个数据元名所必成分，在数据元念可以完整、准家无义情况下，其他术可以情。

一个数据元有一个且仅有一个术。当术与性术有复或分复时，可从名中将冗余删。术6。

定术专业域定。定术是可。

6

	含义
名	一个对一个或
代	替代某一定信息一个有内在则字符串(字、数字、号)
明	描对信息一文字
	以币为单位数，常与币型有关
数	币单位数，常与单位有关。单位参录.1，定构成十倍数和分数单位头录.2
日期	以公元年方式年、月、日合
时	以24小时制时方式天中小时、分、合
日期时	完整时格式，即15，格式
分	具有同单位两个值之分数形式
	一个或与另一个或
标志	又指，两个且只有两个明条件值，如：是/否、有/无
时	两个时时度

本文件中数据元定义以字、字、数字式字符串形式。

数据元值数据类型描则3。本文件将字符串型()分为三形式，1不可枚举，且以字描形式；2枚举型，且列举值不3个；3代形式。

7

数据类型		描述
字符串型 ()		字符形式值型。可包含字符 (,)、数字字 (2312)
布尔型 ()		又型, 0 (1) 或 1 () 形式值型
数值型 ()		0 到 数字形式值型
日期型 ()		/ 740 中 定 格式值型
日期时型 ()		/ 740 中 定 格式值型。(字 作为时 标志 , 明日 开始。)
时型 ()		/ 740 中 定 格式值型
二进制 ()		上无 其他数据类型, 如图像、 、 二进制文件格式

格式 和 。

字	含义
	字 字
	数字字
	字 或 (和) 数字字
	格式 , 其中, 年份, 月份, 日期
6	格式 , 其中 小时, 分 ,
5	1 格式 , 字 作为时 标志 , 明日 时 开始; 其余字 与上同

别	方
固定度	在数据类型 后 接 出字 度数 , 如 4
可变度	1) 可变度不 定义 最大字 数 在数据类型 后加 .. 后 出数据元最大字 数 , 如 ..10 2) 可变度在定义 最小和最大字 数之 在数据类型 后 出最小字 度数后加 .. 后再 出最大字 数, 如 4..20
有千字度	按固定度或可变度 定 出 字 度数后加 后, 再 出最大数, 如 ..40 3
有小数位	按固定度或可变度 定 出字 度数后, 在 , 后 出小数位数, 字 度数包含整数位数、小数 位数和小数位数, 如 6,2

应 例:

例1: 字符串型

10 固定为10个字 (当于5个字) 度 字 。

..10 可变 度, 最大为10个字 度 字 。

4..10 可变 度, 最小为4个最大为10个字 度 字 。

..20 3 可变 度, 最多3 , 最大 度为20个字 度 字 。

例2: 数字型

4 固定 度为4位 数字。

..4 最大 度为4位 数字。

6,2最大 度为6位 十 制小数格式 (包括小数) , 小数 后保 2位数字。

例3: 日期时 型

格式 (位定) 年月日。

15 格式 (15位定) 年月日时分 。时分 之前加大写字 。

如2010年1月5日 时10分 为20100105 0 100 。

本文件数据元值域有两 型:

)可枚举值域: 允 值列 定 值域, 个允 值 值和值含义均应成对 。其中:

可 值 少 (如3个或以下), 在 数据元允 值 属性中 接列举。

可 值 多 (如3个以上), 在 数据元允 值 属性中写出值域代 名 。如代 属引 标准 , 则 明标准号。

)不可枚举值域: 描 定 值域, 在 数据元允 值 属性中 准家描 值域 允 值。

数据元值域代 构 守以下 :

)代 构 代 标 作 , 免承 多 信息, 以保 构 。

)代 构 合信息处 基本方 , 尽 与 内、外 关标准 构协 一 。

)代 构家保代 加、删 和修改不 坏代 构。

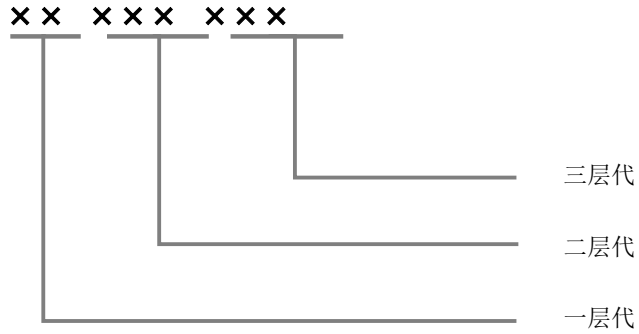
)代 应 便于使 号。

数据元值域代 构描 守以下 :

)应明家描 所 代 、代 构以及 方 。

)层 代 构 可 意图 , 如图6所 。

)当代 构复杂时, 可 例 明。



6

数据元值域代 度 守以下 :

-)在保 前提下,代 度应尽 。
-)尽可 使 代 , 不宜使 不 代 。
-)代 应 对 发展 。

公式

代 度 公式:

$$\sum_{i=1}^n \log_{a_i} Q_i \quad (1)$$

式中:

代 总 度;
层 层数;
层(或) ;

a_i 层(或) 代 字 字 个数;

Q_i 层(或) 对 总数 。

当为 序 或 列 序 时, 1则以上公式变为:

$$\log_a Q \quad (2)$$

代 型及形式应 合下列 :

-)代 字 可 择使 数字型代 、字 型代 、字 数字型代 ;
-)代 字 应 家无、易 易 。应 免使 容易 和 字 。在一个标准中, 、形似字 应 免同时出, 如字 与数字 1;
-)代 最好全 数字或全 字 。字 数字 合 形式一 在 位 使 , 不宜在 机位 使 ;

-) 数字型代 时, 如果有收容 时其代 末位数字为 代 ;
-) 序 时, 代 一 。例如: 001 , 不 1 ; 层 时, 同层 代 ;
-) 在同一个标准中, 代 书写形式 一 , 包括字 大、小写, 代 字体字号。

数据元值域分 与代 (或代) 应以 格 形式列出。依据 / 303 , 格 代 栏 (代 指 值, 可 为 值)、 对 名 栏 (在代 中可 值含义)、 明栏 成, 并可根据实 当增减栏 。

当 格 单时, 为了减少 幅, 可以在一 中并排或两列以上 格。

数据元值域代 书写 如下:

)代 (值) 栏: 代 一 在代 栏内左 格书写; 当代 层 多时, 代 栏可按层 再划分;

)名 (值含义) 栏: 对 名 在名 栏内左 格书写, 个 对 名 占一 。当对 名 时, 可延 下一 , 延 分 与上一 对 。 分 时, 一层 对 名 左 格书写, 二层 一个字, 依 推;

) 明栏: 明 内容在 明栏左 格排。

代 应具备在 定 域 景上 得权威 可 名 。代 名 应准 家 反映代 作为数据元 属性之一 征, 不应放大或 小其使 围。

代 名 应 , 传 明 家 义, 体 代 本 内容。

代 应在 定使 域内具有唯一 标 , 来 别 数据元值域 体 。 学数据元值域代 标 根据 归 家 定。 构为:

+7位数字, 成总 度为11位 字 数字 合 , 包括2个分 号 . 。按 别代 、 序号从左向右 序排列。 构如图7所 。

