



中华人民共和国卫生行业标准

WS 370—2012

卫生信息基本数据集编制规范

Specification for drafting of health information basic dataset

2012-03-15 发布

2012-09-01 实施

中华人民共和国卫生部 发布

前 言

本标准由卫生部卫生信息标准专业委员会提出。

本标准起草单位：卫生部统计信息中心、中国疾病预防控制中心妇幼保健中心、解放军总医院。

本标准主要起草人：孟群、胡建平、汤学军、张彤、刘丽华、王存库、魏桐、童心。

卫生信息基本数据集编制规范

1 范围

本标准规定了卫生信息数据集的内容结构、数据集元数据、数据元属性、数据元索引表示方法。
本标准适用于指导卫生信息相关数据集的编制与使用。

2 规范性引用文件

下列文件对于本文件的应用是必不可少的。凡是注日期的引用文件,仅注日期的版本适用于本文件。凡是不注日期的引用文件,其最新版本(包括所有的修改单)适用于本文件。

WS/T 303 卫生信息数据元标准化规则

WS/T 305 卫生信息数据集元数据规范

WS/T 306 卫生信息数据集分类与编码规则

WS 363.1 卫生信息数据元目录 第1部分:总则

WS 364.1 卫生信息数据元值域代码 第1部分:总则

3 术语和定义

WS/T 303、WS/T 305、WS/T 306、WS 363.1、WS 364.1中界定的以及下列术语和定义适用于本文件。

3.1

基本数据集 basic dataset

在特定主题下,为了满足业务信息系统规范化建设和领域内部以及领域间数据交换与共享需求,设计归纳的各个子系统(或者功能模块)所包含的最小数据元素的集合。

4 内容结构

卫生信息数据集至少应由以下内容构成:

封面
前言
名称
1 范围
2 规范性引用文件
3 数据集元数据
4 数据元目录
4.1 数据元公用属性
4.2 数据元专用属性

5 数据集元数据

5.1 数据集元数据选取

根据 WS/T 305 中规定的 7 个元数据子集,本标准选取了核心元数据中的 8 个“必选(M)”元数据元素,分为 2 个元数据子集,对卫生信息基本数据集进行描述。

5.2 数据集描述格式

卫生信息基本数据集采用表 1 的格式进行描述。

表 1 数据集元数据

元数据子集	元数据项	元数据值
标识信息子集	数据集名称	
	数据集标识符	
	数据集发布方—单位名称	
	关键词	
	数据集语种	
	数据集分类—类目名称	
内容信息子集	数据集摘要	
	数据集特征数据元	

5.3 数据集标识符编码规则

数据集标识符采用字母数字混合码,结构为:

数据集类目编码(DCC)_版本标识符(VI)

其中:

- a) 版本标识符(VI): 结构由 4 部分组成,为“V”+“m..m”+“.”+“n..n”。其中,“m..m”和“n..n”为阿拉伯数字构成,在数学上应是具有意义的正整数。“m..m”表示主版本号,“n..n”表示次版本号。如果数据元更新前后可以进行有效的数据交换,则更新后主版本号不变,次版本号等于当前次版本号加 1;如果数据元更新前后无法进行有效的数据交换,则更新后主版本号等于当前主版本号加 1,次版本号归 0。
- b) 数据集类目编码(DCC):即数据集分类编码。采用长度 9 位的字母数字混合码。按业务领域代码、一级类目代码、二级类目代码、顺序号从左向右顺序排列。结构见图 1。

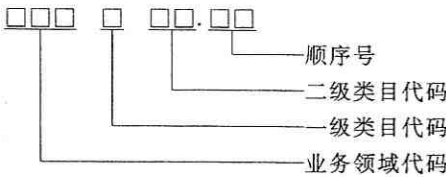


图 1 数据集类目编码(DCC)结构

图 1 中:

业务领域代码:用 3 位大写英文字母表示。卫生信息领域统一用 HDS 表示。

一级类目代码:用 1 位大写英文字母表示,从 A 开始顺序编码。

二级类目代码:用 2 位数字表示,数字大小无含义。如有两个以上类目(含两个),从 01 开始顺序编码。如只有一个类目,编为 00。

顺序号:用 2 位数字表示,代表某二级类目下的数据集序号,数字大小无含义;从 01 开始顺序编码。顺序号与二级类目代码之间加“.”区分。

数据集分类代码详见附录 A。

6 数据元属性

6.1 数据元属性选取

根据 WS/T 303 中规定的五类 22 项数据元基本属性中,本标准选取了五类 14 项数据元基本属性(见表 2),对卫生信息基本数据集的数据元进行描述。

表 2 数据元属性列表

序 号	属 性 种 类	数据元属性名称	约 束
1	标识类	内部标识符	必选
2		数据元标识符	必选
3		数据元名称	必选
4		版本	必选
5		注册机构	必选
6		相关环境	必选
7	定义类	定义	必选
8	关系类	分类模式	必选
9	表示类	数据类型	必选
10		表示格式	必选
11		数据元允许值	必选
12	管理类	主管机构	必选
13		注册状态	必选
14		提交机构	必选

6.2 数据元描述格式

卫生信息基本数据集的数据元描述格式采用摘要式。

在同一个数据集中,版本、注册机构、相关环境、分类模式、主管机构、注册状态、提交机构等 7 个数据元属性的值若完全相同(可称为数据元公用属性),则采用表 3 的描述格式集中描述,在数据元摘要式目录中可不再分述。

表 3 数据元公用属性描述格式

属性种类	数据元属性名称	属性值
标识类	版本	
	注册机构	
	相关环境	
关系类	分类模式	
管理类	主管机构	
	注册状态	
	提交机构	

在同一个数据集中,内部标识符、数据元标识符、数据元名称、定义、数据元值的数据类型、表示格式、数据元允许值等 7 个数据元属性的值均不相同(可称为数据元专用属性),采用表 4 的摘要式目录描述格式进行描述。

表 4 数据元专用属性描述格式

内部标识符	数据元标识符 (DE)	数据元名称	定义	数据元值的 数据类型	表示 格式	数据元允许值

6.3 数据元属性描述规则

6.3.1 内部标识符

内部标识符指数据元在某特定数据集中的唯一标识代码。采用长度 13 位的字母数字混合码,含小数点 2 位。结构见图 2。

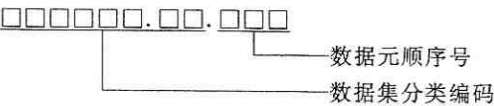


图 2 内部标识符结构

图 2 中:

数据集分类编码(DCC):用 8 位字母数字混合码表示,编码规则见 5.3。

数据元顺序号:用 3 位数字表示,代表数据元在某特定数据集中的序号,从 001 开始顺序编码。数据元顺序号与数据集分类编码之间加“.”区分。

6.3.2 其他属性

WS 363.1、WS 364.1 中对数据元属性已作出的规范适用于本文件。

7 数据元索引表示方法

按数据元首字的汉语拼音顺序提供数据元索引。

附 录 A
(规范性附录)

卫生信息基本数据集分类代码

卫生信息基本数据集分类代码见表 A.1。

表 A.1 卫生信息基本数据集分类代码

一级类目名称	一级类目代码	二级类目名称	二级类目代码
基本信息	A	—	00
卫生服务	B	儿童保健	01
		妇女保健	02
		疾病控制	03
		疾病管理	04
		医疗服务	05
卫生管理	C	—	00
卫生综合	D	—	00