



# 中华人民共和国国家标准

GB/T 25514—2010/ISO/TS 17117:2002

---

## 健康信息学 健康受控词表 结构和高层指标

Health informatics—Controlled health terminology—  
Structure and high-level indicators

(ISO/TS 17117:2002, IDT)

2010-12-01 发布

2011-05-01 实施

中华人民共和国国家质量监督检验检疫总局 发布  
中国国家标准化管理委员会

目 次

前言 ..... III

引言 ..... IV

1 范围 ..... 1

2 规范性引用文件 ..... 1

3 术语和定义 ..... 1

4 概述 ..... 2

4.1 基本情况 ..... 2

4.2 概念的定位 ..... 2

4.3 目标和范围 ..... 3

4.4 映射 ..... 3

4.5 系统性定义 ..... 3

4.6 形式定义 ..... 3

4.7 关系的明确性 ..... 3

4.8 参照术语表 ..... 3

4.9 原子参照术语表 ..... 3

4.10 通用术语表 ..... 3

5 术语模型的结构 ..... 4

5.1 术语结构 ..... 4

5.2 组合术语表 ..... 4

6 维护 ..... 5

6.1 基本情况 ..... 5

7 评估 ..... 6

7.1 基本情况 ..... 6

附录 A (资料性附录) 分类的历史 ..... 10

附录 B (规范性附录) 实现原则 ..... 11

参考文献 ..... 16

## 前 言

本标准使用翻译法等同采用 ISO/TS 17117:2002《健康信息学 健康受控词表 结构和高层指标》(英文版)。

本标准的附录 A 为资料性附录,附录 B 为规范性附录。

本标准由中国标准化研究院提出并归口。

本标准起草单位:中国人民解放军总医院、中国标准化研究院。

本标准主要起草人:尹岭、成福春、董连续、任冠华、陈煌、尹书蕊、刘碧松。

# 引 言

1839 年, William Farr 在关于英国出生、死亡与结婚基本情况的第一个年度报告中指出:“查询系统中的术语与物理科学中的重量与计量一样重要,应当立即制定”。从那时起,广大的科学家们一再强调这个主题。当前,人们已经广泛认识到健康记录系统对受控词表的需求(E-1238, E-1239, E-1384, E-1633, ENV 12017)。受控词表为系统统计提供了方便。这些统计数据能够从各种层次粒度进行统计,因此能够扩展为以问题为导向的临床记录检索,针对财务系统的分类,或者针对给定人群的结果数据。随着术语不断地增加,维护一个大范围的术语成了一个繁重的工作。没有一个良好的结构,大容量的术语不能够提供相应的互操作性,而这是今天日益复杂的电子健康记录程序所要求的。

问题的解决归结于标准。在过去的 10 多年,医学信息学研究者一直在研究受控词表。他们审查了当时存在的术语结构和内容,得出它们为什么不适合特定需求的原因,而且提出了相应的解决方案。在某些地方,解决方案已经在使用,并且又获得了新的经验。现在是到了总结这些经验,发布一个可持续发展的受控医学词表标准集的时候了。

在国际上,ISO/TC 215/WG 3 总结了这些经验,并首次将其制定为技术规范 ISO/TS 17117,以满足健康信息学领域的术语标准化工作需求。

本标准采用了 ISO/TS 17117 的全部技术内容,使我国健康信息学标准化领域在概念表达层面能够与国际惯例保持一致,从而支撑我国健康信息学标准的国际接轨。本标准的制定对我国健康信息学领域术语的规范、集成和使用均有重要指导意义。

# 健康信息学 健康受控词表

## 结构和高层指标

### 1 范围

本标准描述了健康受控词表所需要的和足够的重要理念。  
本标准适用于健康保障信息需要保存和利用的所有领域。

### 2 规范性引用文件

下列文件中的条款通过本标准的引用而成为本标准的条款。凡是注日期的引用文件，其随后所有的修改(不包括勘误的内容)或修订版均不适用于本标准，然而，鼓励根据本标准达成协议的各方研究是否可使用这些文件的最新版本。凡是不注日期的引用文件，其最新版本适用于本标准。

GB/T 18391.3—2009 信息技术 元数据注册系统(MDR) 第3部分:注册系统元模型与基本属性(ISO/IEC 11179-3:2003, IDT)

ISO 704 术语工作 原则与方法

ISO 860 术语工作 概念和术语的一致

ISO 1087-1 术语工作 词汇 第1部分:理论与应用

ISO 1087-2 术语工作 词汇 计算机应用

ISO 12620 关于术语的计算机应用 数据分类

ISO/IEC 2382-4 信息技术 词汇 第4部分:数据的组织

ISO/IEC/TR 9789 信息技术 针对数据交换的数据元素组织与表示指南 编码方法与原则

E-1284 支持电子健康档案的临床术语构建标准指南

E-1712 临床实验室检验与分析物命名表示标准规范

ENV 12264 健康信息学 概念系统的分类结构 语义表达模型

### 3 术语和定义

下列术语和定义适用于本标准。

#### 3.1

**术语表 terminology**

在一个特定的领域,表示系统性概念的术语集。

注:这意味着从发布的目标与范围,人们能够决定这个表达覆盖特定领域的合适程度。

#### 3.2

**健康受控词表 controlled health terminology**

为临床应用而发展的术语集。

注:这意味着在特定被表达领域,在术语的目标和范围内,与实践所产生的粒度一致,为表达编码可比较数据所提供的足够的内容和结构。

#### 3.3

**分类 Classification**

为一个特定领域所做的数据集合在规定层次上的抽象。

注1:能够被用分类系统表达的抽象层次的固定性经常被扩展一致,当分类将被应用于不同的用户群时,如一些目前的账务分类模式。

注2:参见附录A分类的历史。

### 3.4

#### 本体论 ontology

概念的有机组合,通过其可以推导出合理的结论。

示例:限定词的层次是一个限定本体论。

注:通俗地讲,这个术语用于描述为一个特定目的构建的一个层次。

### 3.5

#### 限定词 qualifier

限定词对核心术语作进一步的说明,如肿瘤的分期。

示例:“历史的”或“反复发生的”。

### 3.6

#### 修饰词 modifier

修饰词也是对核心术语的进一步说明,如难以控制的糖尿病。

示例:“临床分期”或“疾病的严重性”。

### 3.7

#### 首选术语 canonical term

一个特定医学概念的首选原子词或首选先组词。

### 3.8

#### 术语 term

与一个或多个概念相关的词语。

## 4 概述

### 4.1 基本情况

术语的基本特征影响其临床应用的实用性与恰当性。参照其他国家和国际标准组织确定的术语理念,制定符合本国实际的术语标准。

本标准明确表明,不仅仅被用来作为临床概念表达,同时也提供术语开发者和创作者以构建有用的和可维护的健康受控词表质量指南。这些原则并不试图整合所有健康领域的丰富词汇。然而,本标准确实规定了最低要求,假如不依照最低要求,术语将会很难维护。目前没有达到这些要求的术语集,可依从本标准,通过某些机制向这个目标前进。分类历史参见附录 A,实现原则见附录 B。

本标准为医学术语的可持续性发展提供理论依据。这些基础工作将为术语不断完善提供一个良好的开端。

### 4.2 概念的定位

术语的基本单元应是概念,它是一个特定含义的体现,而不是一个编码或一个字符串。在一个组织很好的术语中,概念的标识符应且唯一与一个意思相符。如 ISO 860 定义的那样,一个概念表达一个意思。然而,多个词语(语言上的表达),如果它们表达了明确的同一概念,就具有相同的意思。

这意味着无冗余性、无歧义性、无模糊性和内部一致性。

#### 4.2.1 无冗余性

术语应在内部被规范化。在一个术语体系中,不应有超过一个概念标识符具有同一个意思,如 ISO 704 和 E-1284 中规定的那样。这并不是排除同义词,而且要求同义词被明确表达。

#### 4.2.2 无歧义性

概念标识符应只有一个意思。然而,一个录入术语可以指向多个概念。

示例:录入术语 MI 既可以代表心肌梗死,也可以代表二尖瓣关闭不全。

注:有些作者将录入术语表示为接口术语。

#### 4.2.3 无模糊性

概念名称应与环境无关。

示例:“糖尿病”不应有子概念“控制良好的”,然而,可以有子概念名“控制良好的糖尿病”。

注:有些作者将与环境无关称为是环境的装载。

#### 4.2.4 内部一致性

在一个术语体系中,跨平行领域概念间的关系应是统一的。

示例:假如心脏瓣膜结构已经被解剖明确规范,那么与每一个结构相关的诊断也应用相应的关系来进行规范。

#### 4.3 目标和范围

任何一个受控词表在操作术语上应有其明确的目标和范围,因此针对一个特定的目标,它的适合程度能够被评定与评估。在适当的场合,可以在数据库模型中和其他的规范工具中通过范例或“用例”来展现其范围。评估原则,如覆盖范围和全面性,需要用原本意图的用途与范围来评判。

示例:一个术语表对一个心血管的症状、体征和不适来说,可能已经非常全面和详细,但是不适合一个特定的心脏学或一个特定的心胸手术单元。相反地,一个足够详细的心脏学或心胸手术术语表可能完全不适用于心血管的一般实践。

##### 4.3.1 覆盖范围

医疗保健过程的每一个段都应有一个明确的深度覆盖,而不是通过宽广的叶子节点分类将特定的临床概念集中在一起。覆盖的不完全性(与原本的范围与目标比较)应明确地进行说明。

示例:从目前标记的类别“未被分类的”(NEC)中区别特定的诊断经常是非常重要的,或者区分出疾病的严重性如进展缓慢的前列腺癌与广泛的代谢性疾病。

##### 4.3.2 全面性

全面性的程度应被明确说明。在目标与范围中,所有医疗保健处理过程中所涉及到的相关约束和要求,如横跨药学、手术、护理和牙科领域的物理发现、危险因素、功能状态等,都需要进行说明。应用这个原则,是因为决策支持,风险调整、结果研究和有用的指南,需要的不仅仅是临床诊断与处理数据。

示例:目前存在的健康保健研究机构、质量指南机构,以及健康保健财务管理机构的死亡模型等。

#### 4.4 映射

在更多的临床数据交换中,政府和付费者需要处理各种表单及分类模式。因此,基于计算机全面和详细的表达病人数据,应有能力映射到那些分类,如 ICD-10。这需要临床医疗保健的多粒度,如 ISO/IEC/TR 9789 规定的那样。术语能够被映射到其他术语的程度应被明确地说明。

示例:内分泌学医师可能比普通的临床内科医师更关注糖尿病的详细信息,虽然这些专家可能处理同一个病人。

#### 4.5 系统性定义

为了使术语使用者能够准确地理解术语概念的含义,原创作者应让术语使用者明白并且能够获取到这些定义。更重要地,如术语间的内在联系,不同的作者也需要这些定义,来确保与原创作者定义的一致性。

示例:临床概念“高血压”可能被定义为持续增高的血压,需要与简单的“BP>140/85”相区别。

#### 4.6 形式定义

一个组合系统应包含非原子概念和形式规则的形式定义,这样就可以从形式定义推导出小类,如 E-1712 中定义的那样。

#### 4.7 关系的明确性

小类的逻辑关系应被定义。所有的链接/关系/属性的形式行为应被明确定义。假如限定词如“更宽/更窄”使用,它应被明确说明。

示例:主要的层次关系应由小类逻辑关系所例证:B 是 A 的一种类型,意味着所有的 B 都是 A,如糖尿病是一种疾病,所有的糖尿病都是疾病。

#### 4.8 参照术语表

首选术语、结构、关系,系统性的和正式的定义的集合,形成了健康受控词表的核心。

#### 4.9 原子参照术语表

在参照术语表中,仅仅由原子概念和它们的系统性定义构成,而不能由两个或多个概念组合以创建一个组合表达式来表达一个原子的概念。

#### 4.10 通用术语表

日常使用的术语表。当在录入数据时,可映射到一个或多个首选术语。

注:这些也被不同的作者称为“录入术语”或“接口术语”。

5 术语模型的结构

5.1 术语结构

术语结构决定了其实际的和有用的接口(如术语导航、录入、或者检索支持)的难易程度,如 ISO 704,ISO 1087-1 和 ENV 12264 中规定的那样。

5.2 组合术语表

5.2.1 组合性

组合概念从原子和先组式概念产生,具有产生组合表达式的能力。

示例:“结肠癌”由“恶性赘生物”和“大肠”原子组分构成。在一个组合系统中,概念表达式能够被分为原子和组合概念表达式。

组合概念表达式能够被更深入地分为“先组式概念表达”和“后组式概念表达”。在一个组合概念中,组分可以被分成三类:“核心概念”,“限定概念(也叫“状态”)”,和“修饰概念”。

注:一个概念是被语言表达的一种观念,它标识了一个想法。然而,术语“概念”在此处用于概念的表达而不是想法本身。

5.2.1.1 原子概念

在一个特定的术语集中,一个原子概念不是由其他简单概念组合而成。在很多场景,原子概念与某些哲学家称作的“自然种类”类似。该实体不能够被意义上的拆分。从广度的实践角度上说,概念应被分在它们的组分中。这些应成为所有概念的基石。

示例:在 SNOMED-RT,“结肠(colon)”是“大肠(large bowel)”的同义词,而“癌(cancer)”是“恶性赘生物(neoplasm, malignant)”的同义词。因此,术语“结肠癌”是一个非原子术语,其可以被分为“结肠”和“恶性赘生物”。这两个原子术语的任一个都有一个独立概念标识符,就如先组式术语“结肠癌”一样。

5.2.1.2 组合概念

一个组合概念由一个原子组合表达式构成,被语义关系所链接(如角色、属性或者链接)。

5.2.1.2.1 先组式概念

该实体能够被打散成几个原子概念,当这些原子概念重新聚集的时候,不会损失实体信息。这些实体在宿主词汇域中被认为是一种简单概念的表达。理想情况下,在术语系统中,这些概念以及他们等价组合概念应被明确定义。

示例:术语“结肠癌”是非原子的,然而它有一个唯一的标识符,在 SNOMED-RT 中,它代表着一个“单独”的概念。它在术语中与位置“大肠”和诊断“恶性赘生物”具有同样的状态。

5.2.1.2.2 后组式概念

后组式概念是一个组合概念,因其不是先组式,因此应用表达语言的多概念表达式来表达。这是从受控词表中建立的一整套概念去完整表达用户查询的一个尝试。

示例:在 SNOMED-RT 中,概念“细菌性液体流出,左膝”不是一个单独的术语。它表达一个临床概念,有些病人有左膝关节感染。由于它不能被一个简单的概念标识符所表达,为了完整地把握其本义,需要用多概念标识符建立一个表达式,否则将会丢失某些自然文本的信息。

5.2.1.3 原子和先组式概念的类型

在一个术语中,唯一概念表达式能够被分为至少三种类型:核心概念、限定概念和修饰概念。这个区分允许用户接口提供更易读的因此也更有用的概念表达。

5.2.1.3.1 核心概念

这是一个原子的或先组式概念,代表了一个原子概念或者先组式概念或者后组式概念的一个主要概念。

5.2.1.3.2 修饰符或限定符

组合概念的组成部分,改变了核心概念的意思,被认为是修饰符或限定符。

示例:“1a 期”在表达式“大肠癌 1a 期”或“易反弹的、难以控制的”在表达式“易反弹的、难以控制的糖尿病”中,它们就是修订符或限定符。



一般而言,这些概念用一个链接加上一个值(“属性-值-对”)来表达。术语应支持包括短暂的持续和长期的趋势的逻辑结构。属性应是术语元素的一部分,符合一个实际的模型,并可扩展为一个术语。

示例:“癌”,在一个特定的时间具有某些特定的症状,经过一段时间后可能会进展,能够被其分期和组织学再细分。

为了结构化的数据录入和数据二次利用,如聚合统计和检索,属性应体现数据的重要特征。核心概念能够被多种方式精化,包括临床感觉、临时感觉和状态术语,如“反复发作的”。

### 5.2.2 内容的规范化

在组合概念中,规范化是支持和映射替代词和速记词的一个过程。所有的先组式概念应被映射到,或者被所有可能等价的后组式概念所逻辑认知。应有一种机制,为用户创建的新的后组概念确定同义词(在术语集中该创新的概念没有先组式概念的时候)。为了适合个人的、区域的和某些特定的约束,定义明确的等价意义,该功能是非常关键的。并且,作为同义词的术语合作,在一个语言里表达而不仅仅主要用于宿主术语,能够达到一个简单的多语言支持形式。

### 5.2.3 语义的规范化

在组合系统中,表达一个概念可以有多种可能的原子集,可能用多种语义关系进行链接。在这种情况下,术语需要能够识别冗余/同义词。这个系统能够进行规范化的程度(广度)应被明确地说明。

示例:“腹腔镜胆囊切除术”可以用以下两种形式表达:

手术:切除{胆囊},{使用腹腔镜方法}

手术:切除{胆囊},{利用腹腔镜设备}

### 5.2.4 多层次

概念应被各种合理的层次路径所访问,即应允许它们有多语义父体。应保持父体和子女之间数目上的平衡。能够利用直觉的路径,保持自然的导航(检索与分析),找到其感兴趣的概念。即用户不应猜测一个特定的概念究竟放在什么地方。

示例:一个多语义血统中“胃癌”应能够在“赘生物”或者“胃肠道疾病”中找到。

### 5.2.5 视图的一致性

在每一个场景中的多层次系统中的术语,概念应是统一的。前面例子中提到的胃癌在癌症轴上和胃肠道疾病轴上,其意义与结构不应有所差别。当引入了某些不易被察觉的变量时,不一致性可能导致检索和决策支持灾难性的后果,并会误导这个系统中的用户。

### 5.2.6 明确地表达不确定性

如“可能”,“怀疑”,“历史上有过的”或者不同的可能性、不同的诊断列表,都应被支持。“确定的”和“非常不确定”信息对决策支持和数据二次利用有着重要的影响。

类似的,在不完全的症状场景中,临床医师应能够记录与病人表现一致的部分准则(criteria)。这个准则在这里单独列出来,是因为许多目前的术语系统在这方面没有明确地足够地表述。

### 5.2.7 表现形式

在术语中的表现形式应是没有特别意义的。概念标识符的计算机编码不应加上术语主观的限制,如数字的位数、属性,或者组合元素。如果这样做,经常会导致当层次结构更改时,这个编码将不再适用于当前系统。如统一医学语言系统(UMLS)的长的、顺序的、字母数字的标签就是一个很好的例证。

## 6 维护

### 6.1 基本情况

在一段时间内,技术选择能够影响术语进化、改变和保持可用性的能力。

#### 6.1.1 与环境无关的标识符

附加于概念上的唯一编码不应被层次结构或其他的环境所绑定;它们的格式不应带有任何含义。因为健康知识一直处于更新中,健康概念的分类很有可能改变。因为这个原因,附加于概念的编码不应被术语的层次结构所固定,当这个层次结构被重新组织时编码不用更改。改变编码可能引起历史数据

的困惑或错误。

示例：“胃溃疡”现在被认为是一种细菌感染性疾病，但是以前并不这样认为。

注：这个“与环境无关的标识符”与“无语义的标识符”同义。

### 6.1.2 标识符的一致性

当一个概念过时或被取代时，编码不应再使用。当概念改变编码（概念未变，编码改变，增加了编码），经过一段时间后，病人描述的一致性就成为不可能；更糟糕的是，编码的含义改变（编码未变，概念改变）。这种情况不仅仅干扰了历史集合数据的分析，而且对于个人患者的管理是非常危险的，因为他的数据有可能被错误地理解与翻译。

注：这实际上是概念的持久性理念。

### 6.1.3 版本控制

通过一致的版本标识符，更新与修改的部分应能够被查到。对病人记录的使用应携带版本信息。因为解释病人在特定的时间的编码数据是术语的一个基本功能。

示例：AIDS 病人的编码与以前 AIDS 术语的编码就不一致。

当该术语被使用时，术语系统表达应明确该术语当时的状态；版本信息最容易实现这个功能，而且可以被普通的视图所隐藏，如 ISO 12620、ISO 1087-2、GB/T 18391.3 和 ISO/IEC 2382-4 所描述的那样。

#### 6.1.3.1 编辑信息

新的和修订的术语、概念和同义词应有他们进入或者影响系统的时间，而且应指明它们的来源和/或者机构。为了历史检索的目的，一个新录入词的以前表达方式应被记录。

#### 6.1.3.2 已废弃标记

已经被替代的词项应与他们的后续者一起被标记。因为过时的术语数据可能继续存在病人的历史记录中，他们的未来解释和数据聚合统计依赖于那些已经被使用和交叉引用的后继术语。

示例：人类 T-细胞白血病病毒——类型 III(HTLV III)与人类免疫缺陷病毒(HIV)。

### 6.1.4 确认冗余

当冗余发生时，大范围(large-scale)术语的作者们将需要某种机制去验证鉴别冗余。对任何一个术语系统的安全进化来说，这是非常必要的。这意味着概念和语义的规范化，但是特别强调的是，需要提供一些工具和资源来完成这项任务。

### 6.1.5 语言独立

支持多语言表达是术语本身的一个目的。当医疗保健面临着全球的经济和多民族实现环境时，常规的术语维护应与多语言支持结合起来。然而，本质上缺乏机器翻译语言的能力和装备，这个简单化的增加将会扩展全球对术语的理解和使用。

### 6.1.6 响应性

更新的频率或子版本，应是足够短地去适合新的代码或者快速地修改原有版本，理想情况是在几周之内。

## 7 评估

### 7.1 基本情况

当我们尝试理解受控词表的质量时，需要一个评价这个系统的标准准则。所有的评价都应反映和界定出其术语本身的范围与目标。

#### 7.1.1 目标与范围的度量

这是一个重要的维度，应从以下的条目进行目标与范围的度量。

##### 7.1.1.1 临床领域

使用术语的临床领域、病人的疾病领域和/或期望的专业用户是什么？医疗保健术语的哪一个部分

是被期望使用和被谁使用？

#### 7.1.1.2 主要用途

术语的主要用途是什么？

示例：一些用途领域包括：报酬报告、计划管理、流行病学研究、文献索引、基于 Web 的检索、与病人直接相关的临床详细记录、决策支持、决策支持的链接等。

#### 7.1.1.3 使用的一致性和广泛性

至少在最初，部分术语主要是为了一个特定的研究或一个特定的地域。假如试图保持一致性，更新与管理方式是什么，等等。

#### 7.1.1.4 自动推论意图程度

开发者应定义是否支持和支持到哪种程度的自动推论意图。开发者应定义分类是否自动化。开发者应定义是否输入验证的可能性和在什么限制之内。开发者应当定义后组式表达是否能够被接受，如果被接受，什么能够被推论出，什么限制应被附加在上面（即正式的核准是否需要）。

#### 7.1.1.5 转换(映射)到其他术语

什么转换(映射)被支持，为哪一个目的？什么是转换的敏感性和特异性？

示例：为临床用途的变形，比起为文献检索的变形可能需要更少的精确性。

#### 7.1.1.6 用户/开发者的可扩展性

用户和程序开发者是否可以扩展术语？如果可扩展，限制是什么？假如不能，什么机制能够满足新的需求？

#### 7.1.1.7 自然语言

自然语言的输入与输出是否支持(为分析或输入)？到什么精度水平？

#### 7.1.1.8 其他功能

其他的功能是什么？

示例：链接到一个特定的决策支持系统，链接至一个事后市场监督等。

#### 7.1.1.9 目前状态

系统最后将到什么程度，目前的进程？假如术语的不同部分处于不同的完成阶段，如何表明？

### 7.1.2 质量度量，术语工具

#### 7.1.2.1 互联性(映射)

##### 7.1.2.1.1 术语和其他编码系统

术语与其他编码系统或参照术语的映射到达哪种程度？

##### 7.1.2.1.2 术语和术语的扩展

术语适应本地扩展的程度？

##### 7.1.2.1.3 术语和网络

术语服务器是否能够对本地网、广域网的查询进行响应？

#### 7.1.2.2 精度和废止

##### 7.1.2.2.1 术语

用一个标准的较好的方法，对比与一个固定的和国家认可的标准查询集，什么是术语的精度和为映射诊断、手术、展示、解剖、组织等的废止？这应就术语系统本来的目标与范围内评价。

##### 7.1.2.2.2 搜索引擎

在映射过程中是否使用标准的搜索引擎？

#### 7.1.2.3 可用性

##### 7.1.2.3.1 可用性的确认

术语的可用性是否被确认？

##### 7.1.2.3.2 接口考虑

接口考虑如何与术语评价分开？

### 7.1.2.3.3 原型

有效的用户接口是否已经被建立？针对一个有效的用户接口，针对其本来意图的使用，术语是否已经被展示？假如不行，突出的问题是什么？如果用不同的方式，是否有证据说明实际录入的速度、精度、全面性？假如不行，是否有概念性的证据？

### 7.1.2.3.4 应用程序编程接口

是否有计算机接口和系统实现的支持？是否有概念实现的明显证据(在软件上)？如其所说的主要目的能否被展示？是否有实现上的失败？

### 7.1.2.4 可行性

假如术语意图用于电子健康档案，信息存储的选项是什么？是否被展示为可行？

### 7.1.3 其他质量度量

可概括性(适用性)，任何一个研究设计报告(评估报告评价)应可以被评价。

#### 7.1.3.1 健康保障/临床关联性

什么是术语的健康保障/临床关联性？

#### 7.1.3.2 金标准

在评估中什么是金标准？

#### 7.1.3.3 研究人口

假如用人口比率用于比较，人口比率如何得出？

#### 7.1.3.4 特定的目标

特定的目标是否清晰？

#### 7.1.3.5 盲法

研究是否采用合适的盲法？

#### 7.1.3.6 随机

样本集的选择是否随机？或者在某种程度上代表了最终样本？

#### 7.1.3.7 试验场所

##### 7.1.3.7.1 独立性

是否试验场所不同于发展者的场所？

##### 7.1.3.7.2 研究设计的适当性

试验场所是否适合于研究设计(工具、资源，等等)？

##### 7.1.3.7.3 主要的研究组织

以下是主要的研究组织：

- 大学；
- 学术医疗中心；
- 有限公司或公司；
- 医院；
- 政府机构；
- 首诊中心(医疗保健机构)；
- 私人实践；
- 学术组织。

##### 7.1.3.7.4 主要的调查者

主要的调查者是否独立于被评估的术语系统？在这个研究领域中主要的调查者是否有跟踪记录？在进行这项研究中是否有利益上的冲突？

#### 7.1.3.8 项目完成

项目是否在一个合理的时间内完成？

7.1.3.9 样本大小

影响力度:样本大小是否足够,以显示的期望的影响,谁应存在?

统计:谁审查统计方法?

7.1.3.10 个人

7.1.3.10.1 训练水平

研究人员的平均训练水平和经验如何?

7.1.3.10.2 审查者

变异性:什么是内部审查者的变异性?

类型:在这个研究中,审查者的类型是什么(内科医生、护士、其他临床医生、编码者和工程师)?

独立性:这个审查者是否对其他审查者的判断未知(即审查者的独立性)。

附录 A  
(资料性附录)  
分类的历史

自从 17 世纪中叶采用死亡统计单以来,在术语维护的数据方法和原则方面没有太多变化。最广泛被接受的表示病人疾病状态的 ICD-9-CM,就是一个很好的例证。ICD-9-CM 完全依赖于表格结构,具有有限的概念层次,并且没有对同义词的明确机制,值域限制,继承或者语义和非语义的链接。该健康保障分类的维护环境是一个文字处理器,它的分发几乎完全基于纸本。

当前内科医师处理术语(CPT)第一版出现于 1966 年。在美国,CPT 编码系统被医疗保险公司如 Medicare 和几乎所有第三方付费者所使用,包括工人补偿和 Medicaid。作为 Medicare 的 B 部分——内科医师赔付目录,CPT 代码与基于资源的相对量表(Resource Based Relative Value Scale (RBRVS))一起用于决定赔付服务。CPT 代码集是健康保障经济管理通用程序编码系统(Health Care Financing Administration Common Procedure Coding System (HCPCS))的第一层。CPT 代码集,目前是第 4 版,包含数字修订符,注释、指南和一个索引,便于正确利用该编码系统所提供的解释信息。美国医疗学会(AMA)目前正准备发展下一代 CPT(CPT-5)。ICD-10(国际疾病统计分类)和相关的健康问题的维护环境中也具有了相同的限制,而 ICD-10 正在被越来越多的国家采用为诊断编码的国家标准。

在疾病和治疗过程表示方面,重要的进展发生在 1928 年的纽约医学科学院,结果是产生了工业范围内的支持,并且成为疾病与操作的标准术语。这个深刻的技术革新是采用了多轴分类体系。现在一个病理过程(如炎症)能够与一个原子位置(如口咽部分:扁桃腺)组合成为一个诊断(扁桃腺炎)。采用多轴术语编码系统的自然组合,极大地提高了可管理术语的范围和表达能力,并且,提升了为病人编码的粒度水平。

美国病理学院(CAP)通过创建系统化病理术语集(SNOP)及随后的系统化医学术语集(SNOMED),将这件事情更推进一步。在这些系统中,能够被 SNOMED 原子术语合成的术语数字、范围和大小增长到几乎是天文数字。一个比较公认的缺陷,是这种表达缺乏语法规则、组合规则以及概念和语义两者之间的正规化。正规化是一个过程,通过它系统能够知道两个组合的结构是否具有一个相同的意思,实际上也应如此(如术语“大肠癌”等同于“恶性赘生物”和部位“大肠”的组合)。让 SNOMED 成为强壮的医疗保障的参照术语,也是 CAP 成立以来的一直想要实现的一个目的。

其他重要的原创是临床术语 V3(只读编码),由英国的国家健康服务(NHS)和医师协会维护与发布,其术语描述的组合规则非常详细。这个只读编码是一个非常巨大的术语库,现在是第 3 个版本,采用分层体系的结构设计,并且将在全英国范围内使用。统一医学语言系统(UMLS)和其词库由美国国立医学图书馆发起,目的是为目前存在着的多个健康保障术语系统,提供一个医学领域的世界术语。一个值得注意的进展是 CAP 和 NHS 新签订的协议(1999 年 4 月),将 SONMED-RT 和临床术语 V3 合并,产生出一个新的成果,即 SNOMED 临床术语集。

**附 录 B**  
**(规范性附录)**  
**实 现 原 则**

**B. 1 概述**

**B. 1.1 基本情况**

在临床应用中,术语的基本特征影响其实用性与适合性。

**B. 1.2 概念的定位**

什么是概念的定位?这是一个重要的问题。一个标识符能够与多少个含义相关联?

**B. 1.2.1 无冗余性**

在术语系统中,初始概念不应具有冗余性。

**B. 1.2.2 无歧义性**

概念定义是不应有歧义?

**B. 1.2.3 无模糊性**

概念定义应独立于他们的环境。

**B. 1.2.4 内部一致性**

术语之间的关系应用是否一致?当然应一致。

**B. 1.3 目标和范围**

术语的目标是什么?术语的范围是什么?这些应在操作术语中说明(即术语意图实现的功能是什么?)。

**B. 1.3.1 覆盖范围**

术语的覆盖范围是什么?

**B. 1.3.2 全面性**

在一个原本的覆盖范围领域,术语的全面性的程度(用完全百分比来表示)?什么研究能够被引用以支持这个论断?(使用第4章中给定的原则来评估这个参考研究的有效性和概括性)。

**B. 1.4 映射**

是否这个术语集能够映射到其他的分类或术语集?如果可以,哪一个术语集?假如是能够部分映射到某些分类或者其他术语集,到什么程度(用完全百分比表示)?(使用短语4中给定的原则来评估这个参考研究的有效性和概括性)。

**B. 1.5 系统性定义**

是否术语集中每一个特定的概念能够被用户所理解?这个要求应被满足。

**B. 1.6 形式定义**

是否术语支持形式定义?假如这样的话,完整定义支持到哪种程度?(用完全百分比表示)?什么研究能够被引用以支持这个论断?(使用第4章中给定的原则来评估这个参考研究的有效性和概括性)。参照术语支持形式定义是必要的。

**B. 1.7 关系的明确性**

术语是否支持正式的小结?层次体系通过描述逻辑自动产生到什么程度(用术语集中所有包含概念和一个百分比表示)?这是一个期望的特征。

**B. 1.8 参照术语表**

是否这个术语意图被用于参照术语?

### B.1.9 原子参照术语表

是否有一个明确的机制确定参照术语的原子位置？是否先组式术语能够被用于组合表达式？这应是所有参照术语的目标。

### B.1.10 通用术语表

特别地，在通用术语与参照术语之间的关系是怎样？这两个术语如何保持以避免产生数据实例的歧义或冗余？这是所有意图用于临床的参照术语应解决的问题。

## B.2 术语模型的结构

### B.2.1 术语结构

术语结构决定了其实用的与有用的接口的难易程度，如术语导航、录入或者检索能否被支持，如 ISO 704, ISO 1087-1 和 ENV 12264。

#### B.2.1.1 组合术语表

##### B.2.1.1.1 组合性

术语是否支持组合术语表达式？组合表达式怎样产生？如果是在规则指导下产生的，它们将是精致的。假如这样，等价的表达能否被任意的组合表达式鉴定，并且用什么方法？

##### B.2.1.1.1.1 原子概念

术语作者是否明确表示其概念是原子的？

##### B.2.1.1.1.2 组合概念

组合概念作为一个表达式，由原子概念与语义链接（如角色、属性或链接）构成。

先组式概念：术语作者是否明确表明其概念是先组式的？在所有的组合术语中，这种情况都应出现。

后组式概念：术语是否支持后组式表达式的创建？

##### B.2.1.1.1.3 原子和先组式概念的类型

在一个术语中唯一性的概念表达至少能够被分成三个显著的部分：核心概念、修饰符和限定符（其包含了状态概念）。这种区分能够允许用户界面提供更多的可读的因此也更加有效地进行组合概念的表达。

核心概念：是否术语单独地表明了核心概念？在组合术语中这个应当被表明。

修订符与限定符是精化核心概念的词汇。在术语中是否明确表明了修订符与限定符？如果表明，它们如何被使用？在组合术语中应当被表明。

##### B.2.1.1.2 内容的正规化

术语的内容是否被正规化？什么研究能够被引用以支持这个论断（用给定的短语？规则评估有效性和研究引用的概括性。）？在所有的组合术语中，这个应被完成。

##### B.2.1.1.3 语义的正规化

是否术语的语义被正规化？什么研究能够被引用以支持这个论断（用给定的短语？规则评估有效性和研究引用的概括性。）？在组合术语中，表达同一个概念能否用不同的语义？在所有的组合术语中，这个应被完成。

##### B.2.1.1.4 多层次体系

是否多层次体系被支持？在目前的术语版本中他们是否出现？

##### B.2.1.1.5 视图的一致性

术语中视图的一致性是否被维护？在支持多层次体系的术语中，这个应被实现。

##### B.2.1.1.6 明确地表达不确定性

是否术语支持明确的不确定性和不完整症状的输入？这个应是组合术语的一个特性。

##### B.2.1.1.7 表现形式

是否术语的概念标识符表现形式具有某些限制？如果有，是什么限制？这种情况不应发生。



### B.3 维护

#### B.3.1 基本情况

技术选择能够影响术语在一段时间内进化、改变和保持有用性的能力。

##### B.3.1.1 与环境无关的标识符

是否术语支持与环境无关的标识符？应支持。

##### B.3.1.2 标识符的一致性

因为不同的概念引起编码曾经被重用的现象是否发生过？假如是，什么情况下可能发生？这种情况不应出现。

##### B.3.1.3 版本控制

是否术语的编码明确地与版本紧密绑定？应与版本绑定。

###### B.3.1.3.1 编辑信息

当术语被修订，导致更新的时间、来源或者机构信息是否被记录？应被记录。

###### B.3.1.3.2 已废弃标记

在术语的条目中，废弃的标记是否被包含在其中？应包含。

##### B.3.1.4 确认冗余

术语能否识别冗余？应识别冗余。假如这样的话，怎样完成？

##### B.3.1.5 语言独立

术语是否被多语言表达？当然应被多语言表达。假如不是，是否术语有多语言表达的能力？假如这样的话，应当给出一个解释。

##### B.3.1.6 响应性

术语更新的频率是多少？是否少于或等于 12 周？应少于。

### B.4 评估

#### B.4.1 基本情况

当我们创建或理解受控词表的时候，需要一个评估这些系统的标准准则。所有的评估应反映和明确地鉴定被评估术语的目标和范围。这个准则促进了受控词表评估方法的研究。这个准则也有利于指导那些希望实施健康受控词表的个人或组织。

##### B.4.1.1 目标与范围的度量

这是一个重要的维度，应从以下方面来定义和说明目标与范围。

###### B.4.1.1.1 临床领域

术语的临床领域是什么？病人的疾病领域和/或者期望的专业用户？在健康保健术语中的哪一部分打算被使用和被谁使用？

###### B.4.1.1.2 主要用途

术语本来的最主要用途是什么？

例：一些用途领域包括：报酬报告、计划管理、流行病学研究、文献索引、基于 Web 的检索、与病人直接相关的临床详细记录、决策支持、决策支持的链接等。

###### B.4.1.1.3 使用的一致性和广泛性

术语是否意图保持一致性和进化？如果意图保持一致，更新和改变的机制是什么，等等？

###### B.4.1.1.4 自动推论意图程度

术语是否意图支持自动分类？输入的验证是否进行，或者具有什么限制？后组式表达式是否被接受，假如是的话，什么能够被从他们中推论出来，什么限制会附加于上面？

###### B.4.1.1.5 转换(映射)到其他术语

什么转换(映射)被支持和本来用于什么目的(如为文献检索的目的而转换的精度要低于为临床使

用而转换的精度。)？转换的敏感性和专一性是什么？

#### **B.4.1.1.6 用户/开发者的可扩展性**

术语是否意图被使用者或者应用程序开发者扩展？假如可以的话，有什么限制？假如不行，可通过什么样的机制去满足新出现的需求？

#### **B.4.1.1.7 自然语言**

是否支持自然语言的输入与输出(为分析或输入)？到达哪种程度的精度？

#### **B.4.1.1.8 其他功能**

原本还有哪些其他功能？

示例：链接到特定的决策支持系统，链接到后市场监管，等等。

#### **B.4.1.1.9 目前状态**

到什么程度系统将会“完成”或正在进行中？假如术语的不同部分处于不同的完成阶段，应如何表示出来？

### **B.4.1.2 质量度量，术语工具**

#### **B.4.1.2.1 互联性(映射)**

##### **B.4.1.2.1.1 术语和其他编码系统**

术语能够与其他编码系统或参照术语映射到什么程度？

##### **B.4.1.2.1.2 术语和术语的扩张**

术语能够适应本地扩张的程度？

##### **B.4.1.2.1.3 术语和网络**

通过网络的查询(本地网、广域网)，词汇服务器的响应如何？

#### **B.4.1.2.2 精度和废止**

##### **B.4.1.2.2.1 术语**

用一个标准的程序化的方法，对比与一个固定的和国家承认的标准查询集，评价术语为映射、诊断、手术、展示、角度、组织等的精度和废止？这应就术语系统本来的目标与范围内评价。

##### **B.4.1.2.2.2 搜索引擎**

是否一个标准的搜索引擎用于映射过程？

#### **B.4.1.2.3 可用性**

##### **B.4.1.2.3.1 可用性的确认**

术语的可用性是否被确认？

##### **B.4.1.2.3.2 接口考虑**

接口考虑如何与术语评价分开？

##### **B.4.1.2.3.3 原型**

有效的用户接口是否已经被建立？针对一个有效的用户接口，针对其本来意图的使用，术语是否已经被展示？假如不行，突出的问题是什么？如果用不同的方式，是否有证据说明实际录入的速度、精度、全面性？假如不行，是否有概念性的证据？如果用户接口未建立，是否有概念实现的证据？

##### **B.4.1.2.3.4 应用程序编程接口**

是否有计算机接口和系统实现的支持？是否有明显的证据在软件上实现？如其所说的主要目的能否被展示？是否有实现上的失败？

#### **B.4.1.2.4 可行性**

假如术语意图用于电子健康档案，信息存储的选项是什么？是否被展示为可行？

### **B.4.1.3 其他质量度量**

可概括性(适用性)，任何一个研究设计报告(评估报告评价)应可以被评价。

#### **B.4.1.3.1 健康保障/临床关联性**

什么是术语的健康保障/临床关联性？

#### **B.4.1.3.2 金标准**

在评估中什么是金标准?

#### **B.4.1.3.3 研究人口**

假如用人口比率用于比较,人口比率如何得出?

#### **B.4.1.3.4 特定的目标**

特定的目标是否清晰?

#### **B.4.1.3.5 盲法**

研究是否采用合适的盲法?

#### **B.4.1.3.6 随机**

样本集的选择是否随机?或者在某种程度上代表了最终样本?

#### **B.4.1.3.7 试验场所**

##### **B.4.1.3.7.1 独立性**

是否试验场所不同于发展者的场所?

##### **B.4.1.3.7.2 研究设计的适当性**

试验场所是否适合于研究设计(工具、资源,等等)?

##### **B.4.1.3.7.3 主要的研究组织**

以下是主要的研究组织:

- 大学;
- 学术医疗中心;
- 有限公司或公司;
- 医院;
- 政府机构;
- 首诊中心(健康保健机构);
- 私人实践;
- 学术组织。

##### **B.4.1.3.7.4 主要的调查者**

主要的调查者是否独立于被评估的术语系统?在这个研究领域中主要的调查者是否有跟踪记录?  
在进行这项研究中是否有利益上的冲突?

#### **B.4.1.3.8 项目完成**

项目是否在一个合理的时间内完成?

#### **B.4.1.3.9 样本大小**

影响力度:样本大小是否足够,以显示期望的影响,谁应存在?

统计:谁审查统计方法?

#### **B.4.1.3.10 个人**

##### **B.4.1.3.10.1 训练水平**

研究人员的平均训练水平和经验是什么?

##### **B.4.1.3.10.2 审查者**

变异性:什么是内部审查的变异性?

类型:在这个研究中,审查者的类型是什么(内科医生、护士,其他临床医生、编码者和工程师)?

独立性:这个审查者是否对其他审查者的判断未知(即审查者的独立性)。

参 考 文 献

- [1] ISO 12200 Computer applications in terminology—Machine-readable terminology interchange format (MARTIF)—Negotiated interchange
- [2] E-1238 Standard Specification for Transferring Clinical Observations Between Independent Computer Systems
- [3] E-1239 Standard Guide for Description of Reservation/Registration—Admitting, Discharge, Transfer (R-ADT) Systems for Electronic Health Record (EHR) Systems
- [4] E-1384 Standard Guide for Content and Structure of the Electronic Health Record (EHR)
- [5] E-1633 Standard Specification for Coded Values Used in the Electronic Health Record
- [6] ENV 12017 Health Informatics—Medical Informatics Vocabulary
- [7] Masys DR. Of codes and keywords: standards for biomedical nomenclature. *Acad Med.* 1990;65:627-629
- [8] Solbrig HR. Final submission to the CorbaMED Request for Proposals on Lexical Query Services (CorbaLex) OMG. 1998. <http://www.omg.org/cgi-bin/doc?formal/99-3-6.pdf> or <http://www.omg.org/cgi-bin/doc?formal/99-3-1.pdf>
- [9] Cimino JJ. Desiderata for Controlled medical Vocabularies in the Twenty-first Century. *Methods of Information in Medicine.* 1998;37(4/5):394-403
- [10] Cote RA, Rothwell DJ. The Classification-nomenclature Issues in Medicine: A Return to Natural Language. *Medical Informatics.* 1989;14(1):25-41
- [11] Rocha RA, Rocha BH, Huff SM. Automated Translation Between medical Vocabularies using a Framebased Interlingua. *Proceedings of the Annual Symposium of Computer Applications in Medical Care.* 1993:690-694
- [12] Bernauer J, Franz M, Schoop D, Schoop M, Pretschner DP. The Compositional Approach for Representing Medical Concept Systems. *Medinfo 95;8 Pt (1):70-4*
- [13] Campbell KE, Musen MA. Representation of Clinical Data Using SNOMED III and Conceptual Graphs. *Proceedings of the Annual Symposium on Computer Applications in Medical Care* 1992:354-8
- [14] Rossi Mori A, Galeazzi E, Gangemi A, Pisanelli DM, Thornton AM. Semantic Standards for the Representation of Medical Records. *Medical Decision Making* 1991;4(Suppl):S76-80
- [15] Tuttle MS, Olson NE, Campbell KE, Sherertz DD, Nelson SJ, Cole WG. Formal Properties of the Metathesaurus. *Proceedings of the Annual Symposium on Computer Applications in Medical Care* 1994:145-9
- [16] Campbell KE, Cohn SP, Chute CG, Rennels G, Shortliffe EH. Galapagos: Computer-based Support for Evolution of a Convergent Medical Terminology. *JAMIA* 1996;SympSuppl:269-73
- [17] Cimino JJ. Formal Descriptions and Adaptive Mechanisms for Changes in Controlled Medical Vocabularies, *Methods of Information in Medicine* 1996;35(3):211-217
- [18] Elkin PL, Chute CG. ANSI-HISB Code Set Evaluation Criterion Survey, 1998; Minutes ANSI-HISB meeting 4/98
- [19] William Farr; Regarding the Cullenian system of 1785, First Annual Report of the Registrar-General of Births, Deaths, and Marriages in England. London;1839 p. 99
- [20] Evans DA, Cimino JJ, Hersh WR, Huff SM, Bell DS, for the Canon Group. Toward a

Medical-Concept Representation Language. JAMIA 1994,1:207-217

[21] Musen MA, Wieckert KE, Miller ET, Campbell KE, Fagan LM. Development of a controlled medical terminology: knowledge acquisition and knowledge representation. *Methods of Information in Medicine*. 34(1-2):85-95,1995 Mar

[22] Cimino JJ. Desiderata for Controlled Medical Vocabularies in the Twenty-first Century. *Methods of Information in Medicine* 1998, in press

[23] Cote RA, Rothwell DJ. The Classification-nomenclature Issues in Medicine: A Return to Natural Language. *Medical Informatics* 1989;14(1):25-41

[24] Rocha RA, Rocha BH, Huff SM. Automated Translation Between medical Vocabularies using a Framebased Interlingua. *Proceedings of the Annual Symposium of Computer Applications in Medical Care*. 1993:690-694

[25] Campbell KE, Musen MA. Representation of Clinical Data Using SNOMED III and Conceptual Graphs. *Proceedings of the Annual Symposium on Computer Applications in Medical Care* 1992:354-8

[26] Rossi Mori A, Galeazzi E, Gangemi A, Pisanelli DM, Thornton AM. Semantic Standards for the Representation of Medical Records. *Medical Decision Making* 1991;4(Suppl):S76-80

[27] Tuttle MS, Olson NE, Campbell KE, Sherertz DD, Nelson SJ, Cole WG. Formal Properties of the Metathesaurus. *Proceedings of the Annual Symposium on Computer Applications in Medical Care* 1994:145-9

[28] Campbell KE, Cohn SP, Chute CG, Rennels G, Shortliffe EH. Galapagos: Computer-based Support for Evolution of a Convergent Medical Terminology. *Journal of the American Medical Informatics Association* 1996;SympSuppl:269-73

[29] Elkin PL, Chute CG. ANSI-HISB Code Set Evaluation Criterion Survey, 1998; Minutes ANSI-HISB meeting 4/98

---

中 华 人 民 共 和 国  
国 家 标 准  
健康信息学 健康受控词表  
结构和高层指标

GB/T 25514—2010/ISO/TS 17117:2002

\*

中国标准出版社出版发行  
北京复兴门外三里河北街16号  
邮政编码:100045

网址 [www.spc.net.cn](http://www.spc.net.cn)

电话:68523946 68517548

中国标准出版社秦皇岛印刷厂印刷  
各地新华书店经销

\*

开本 880×1230 1/16 印张 1.5 字数 36 千字  
2011 年 4 月第一版 2011 年 4 月第一次印刷

\*

书号: 155066·1-42411 定价 24.00 元

如有印装差错 由本社发行中心调换  
版权专有 侵权必究  
举报电话:(010)68533533



GB/T 25514-2010