

团 体 标 准

T/CIATCM 098—2023

中医药文本数据命名实体信息分类与代码

Named entity information classification and codes of Traditional
Chinese Medicine text data

2023 - 07 - 15 发布

2023- 08 - 01 实施

中国中医药信息学会

发布

目 次

前言.....	II
1 范围.....	1
2 规范性引用文件.....	1
3 术语和定义.....	1
4 分类方法和编码原则.....	1
4.1 分类方法.....	1
4.2 编码规则.....	2
5 中医药文本数据命名实体信息分类与代码.....	2
5.1 表型实体.....	3
5.2 治疗实体.....	4
5.3 基础实体.....	4
5.4 辅助实体.....	6
5.5 风险因素.....	6
5.6 操作方法.....	6
5.7 其他实体.....	7

前 言

本文件按照GB/T 1.1—2020《标准化工作导则 第1部分：标准化文件的机构和起草规则》的规定起草。

请注意本文件的某些内容可能涉及专利。本文件的发布机构不承担识别专利的责任。

本文件由中国中医药信息学会提出并归口。

本文件起草单位：北京交通大学、中国中医科学院中医药数据中心、湖北中医药大学、中国中医科学院广安门医院、湖北省中医院、河南中医药大学、天士力医药集团股份有限公司。

本文件主要起草人：周雪忠、刘保延、常凯、鄢灯莹、夏佳楠、杨扩、肖勇、舒梓心、张润顺、周霞继、田昊宇、李晓东、周亚娜、余海滨、王文佳、魏宇、孙海龙。

中医药文本数据命名实体信息分类与代码

1 范围

本文件规定了中医药临床病历、古籍、文献等文本数据命名实体的信息分类与代码。
本文件适用于中医医疗、科研、教学、信息化建设等领域数据处理、交换与共享。

2 规范性引用文件

下列文件中的内容通过文中的规范性引用而构成本文件必不可少的条款。其中，注明日期的引用文件，仅该日期对应的版本适用于本文件；未注明日期的引用文件，其最新版本（包括所有的修改单）适用于本文件。

GB/T 20000.1—2014 标准化工作指南 第1部分：标准化和相关活动的通用术语

GB/T 20348—2006 中医基础理论术语

GB/T 40665.3—2021 中医四诊操作规范 第3部分：问诊

T/CIATCM 020—2019 中医临床基本症状信息分类与代码

T/CIATCM 058—2019 中医药信息标准编制通则

3 术语和定义

GB/T 20000.1、T/CIATCM 020—2019界定的以及下列术语和定义适用于本文件。

3.1

命名实体 *named entity*

非结构化文本中主要以短语等形式表达的具有特定意义的信息元素（如城市名称、机构名称和人名等），适用于文本挖掘、信息抽取、自动问答等自然语言处理领域。

3.2

信息分类 *information classification*

信息分类是根据信息内容的属性或特征，将信息按一定的原则和方法进行区分和归类，并建立起一定的分类体系和排列顺序。

[来源：T/CIATCM 020—2019，3.1.8]

3.3

信息编码 *information coding*

信息编码是将事物或概念（编码对象）赋予具有一定规律、易于计算机和人识别处理的符号，形成代码元素集合。

[来源：T/CIATCM 020—2019，3.1.9]

4 分类方法和编码原则

4.1 分类方法

本文件中医药文本数据命名实体分类采用线分类法。线分类法是将分类对象（即被划分的事物或概念）按照选定的若干个属性或特征逐次地分成相应的若干个层级的类目，并排列成一个有层次的，逐渐展开的分类体系。在这个分类体系中，被划分的类目称为上位类，划分出的下一级类目称为下位类，由一个类目直接划分出来的下一级各类目，彼此称为同位类。同位类之间存在着并列关系，下位类与上位类类目之间存在着隶属关系。

4.2 编码规则

本文件中医药文本数据命名实体分类采用字母和数字符号混合编码方式，代码结构如图1所示。

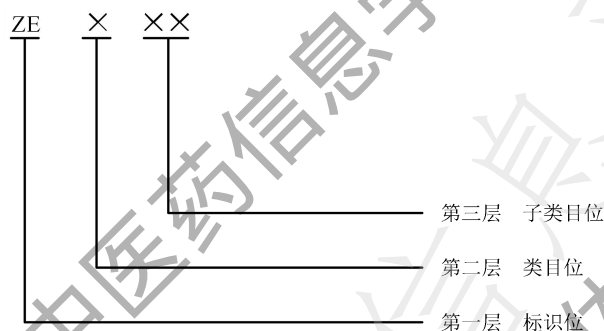


图1 中医药文本数据命名实体分类代码结构图

标识位：以“ZE”作为中医药文本数据命名实体的标识，2位码长。

类目位：根据中医药文本数据命名实体不同类别，分为表型实体类、治疗实体类、基础实体类、辅助实体类、风险因素类、操作方法类和其他实体类，共计7个类目，采用字母编码，1位码长。

子类目位：将中医药文本数据命名实体7个类目进行细分，形成不同类目下的子类目，采用数字编码，2位码长，从“01”开始顺序编码。

5 中医药文本数据命名实体信息分类与代码

中医药文本数据命名实体类目层次结构图及代码表分别如表1、图2所示。

表1 中医药文本数据命名实体类目名称与代码表

类目名称	代码
表型实体	P
治疗实体	T
基础实体	B
辅助实体	A
风险因素	R
操作方法	O
其他实体	Z

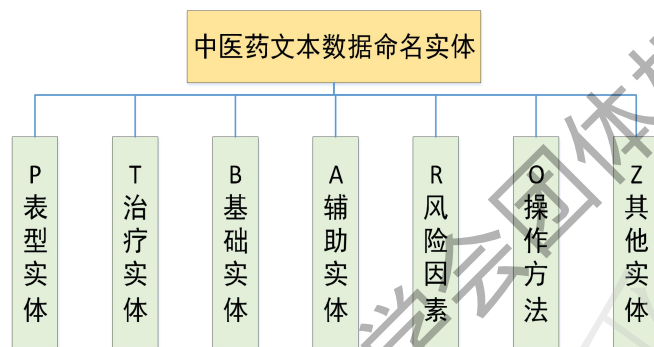


图2 中医药文本数据命名实体类目层次结构图

5.1 表型实体

包括症状、体征、疾病、理化检查等，是医学信息的重要组成部分，是不同医学研究系统之间进行数据集成、交互和信息共享的关键要素。

本文件列举了中医药文本数据中较为常见的表型实体子类目，包括阳性症状、阴性症状、舌象、脉象、证候等。如表2所示。

表2 “表型实体”分类名称与代码表

类目名称	子类目名称	代码	说明
表型实体		P	
	阳性症状	01	指疾病过程中机体内的一系列机能、代谢和形态结构异常变化所引起的患者自身感觉到的不适或痛苦的异常感觉或某些客观病态改变。
	阴性症状	02	指各种生理情况下的临床表现，无异常的临床症状，常用于临床鉴别诊断或描述病情变化时使用，如无发热、无咳嗽等。
	舌象	03	指舌体、舌苔的色泽与形态所构成的形象，中医用以临床诊察病情。
	脉象	04	指脉搏跳动的形象与动态，包括脉的位、数、形、势等。
	证候	05	指中医临床诊疗过程中依据四诊（望、闻、问、切）信息，综合分析病因病机后形成的中医诊断性结果。
	中医疾病	06	指以中医理论为基础，以病因、病机、病理产物、病位、主症、体征为依据命名的疾病。
	西医疾病	07	指在现代医学中，机体在一定的条件下，受病因损害作用后，因自稳调节紊乱而发生的异常生命活动过程。
	生化检查	08	指使用生物或化学的方法来对人进行身体检查的各项检查内容，一般包括多种，如空腹血糖、血钾等，每种检查都有不同的正常范围。
	物理检查	09	指借助物理的方法来对人进行身体检查的各项检查内容。
	曾有诊断	10	指患者既往患病的诊断。
	曾有症状	11	指患者在过去一段时间里出现过的症状。
	生化检查结果	12	指借助生物或化学的方法对人进行身体检查得到的结果。

表 2（续）

类目名称	子类目名称	代码	说明
	物理检查结果	13	指借助物理的方法对人进行身体检查得到的结果。
	转归	14	指病情的转移和发展，如病情的恶化或好转。
	阳性体征	15	指医生运用视觉、听觉、嗅觉及触觉等能客观感知到的各种异常征象。
	阴性体征	16	指各种与异常的阳性体征对应的未出现异常体征描述，常用于临床鉴别诊断或描述病情变化时使用，如麦氏点无压痛。

5.2 治疗实体

包括临床干预或改变特定健康状态所涉及的各类信息元素。

本文件列举了中医药文本数据中较为常见的治疗实体子类目，包括处方、中药、西药、治则、治法等。如表3所示。

表 3 “治疗实体”分类名称与代码表

类目名称	子类目名称	代码	说明
治疗实体		T	
	处方	01	指患者就诊时医生填写的有关药物治疗信息。
	中药	02	指以中医药理论指导采集、炮制、制剂，说明作用机理，指导临床应用的药物，包括植物药、动物药、矿物药及部分化学、生物制品类药物等。
	西药	03	指相对于中药而言，指现代医学用的药物，一般用化学合成方法制成或从天然产物提制而成，包括有机化学药品、无机化学药品和生物制品等。
	中成药	04	指以中药材为原料，在中医药理论指导下，为了预防及治疗疾病的需要，按规定的处方和制剂工艺将其加工制成一定剂型的中药制品。
	治则	05	对临床的具体立法、处方、用药等具有普遍的指导意义，在治疗疾病时必须遵循的基本原则。
	治法	06	以治疗原则为指导，针对不同病证采用的治疗方法与手段。
	穴位	07	指人体脏腑经络气血输注出入的特殊部位，中医可以通过针灸或者推拿、点按、艾灸刺激相应的穴位治疗疾病。
	手术	08	指医生用医疗器械对患者身体进行的切除、缝合等治疗。
	特色疗法	09	指中医特色疗法，包括艾灸、刮痧、熏蒸、拔罐等。
	功效	10	指中药用于防治疾病的作用或疗效的概括性分类。
	曾用药物类别	11	指患者在过去一段时间使用的药物类别。
	曾用药物	12	指患者在过去一段时间使用的药物。
	方名	13	指根据配伍原则，以若干药物配伍组成的药方名称。
	煎服方法	14	指煎药法和服药法。煎药法包括煎药用水、加水量、煎药火候等，还包括先煎、后下、包煎、单煎、烊化、冲服等特殊煎法；服药法包括服药时间、服用方法和药后调护。

表 3 (续)

类目名称	子类目名称	代码	说明
	中药炮制	15	是指根据中医药理论,依照辨证施治用药的要求和药物自身性质,以及调剂、制剂的不同要求所采取的制药技术。
	中药剂量	16	指临床应用药物时的剂量。
	治疗效果	17	指使用药物等干预措施治疗某种疾病后所获得的预期结果或者成效。
	不良反应	18	指按正常用法、用量应用药物预防、诊断或治疗疾病过程中,发生的与治疗目的无关的有害反应。
	方剂剂型	19	指中药方剂组成后,根据病情与药物特点制成一定的形态。
	随症加减	20	是指在主证未变的情况下,随着患者兼有症状的变化,加入或移去某些药物,使之更合乎临床治疗的需要。
	西医疗法	21	指结合利用现代西医科学技术开展的治疗方法,如放射治疗、吸氧等。

5.3 基础实体

中医药文本数据中有关中医药基础知识表达的信息元素。

本文件列举了中医药文本数据中较为常见的基础实体子类目,包括脏腑、经络、病理产物、病因、病机等。如表4所示。

表 4 “基础实体”分类名称与代码表

类目名称	子类目名称	代码	说明
基础实体		B	
	脏腑	01	人体内脏的总称。包括五脏、六腑和奇恒之腑。
	经络	02	经脉和络脉的总称,为人体运行气血、联络脏腑、沟通内外、贯穿上下的径路。
	病理产物	03	指在初始病因作用下,机体气化功能失调所产生并引起新的病理变化的致病因素。包括痰饮、瘀血、结石等。
	病因	04	破坏人体阴阳相对平衡而引起疾病的原因。
	病机	05	疾病发生、发展、变化的机制。包括病因、病性、病位、病势等的变化及其机制。
	体质	06	人体在先天禀赋和后天调养基础上所形成的形态结构和生理功能上相对稳定的特性。
	中药性味	07	指药物的性质和气味,即四气五味。四气指药物的寒、热、温、凉四种特性,又称四性。五味原指药物的辛、甘、酸、苦、咸五种味道,后扩展为体现药物功能归类的标志。
	中药归经	08	归,即归属,是指药物作用的归属;经,即人体的脏腑经络。归经,即药物作用的定位。就是把药物的作用与人体的脏腑经络密切联系起来,以说明药物作用对机体某部分的选择性,从而为临床辨证用药提供依据。

5.4 辅助实体

中医药文本数据中用于描述疾病、症状、体征等实体的信息元素。

本文件列举了中医药文本数据中较为常见的辅助实体子类目，包括发作特征、发作频次、持续时间、部位、性质等。如表5所示。

表5 “辅助实体”分类名称与代码表

类目名称	子类目名称	代码	说明
辅助实体		A	
	发作特征	01	指疾病、症状、体征等出现时所表现出来的外在特征。
	发作频次	02	指某一症状在一定时间内重复出现的次数。
	持续时间	03	指疾病、症状、体征等所持续的时间。
	部位	04	指症状或体征出现在人体某部分的位置。
	性质	05	指某一症状发作时呈现多种情况时，对每一种情况的具体性质的描述，如疼痛的性质可呈现隐痛，刺痛，胀痛等。
	特征	06	指某一症状发作时呈现不同特征的描述，对每一种情况的具体特征的描述，如关节疼痛的游走性等。
	变化情况	07	指症状发生变化或转变的情况。
	加重因素	08	指导致疾病或症状加重或恶化的影响因素。
	缓解因素	09	指导致疾病或症状缓解或好转的影响因素。
	发生因素	10	指导致疾病或症状发生的其他因素，如季节时辰、体态变化、外来刺激、月经胎产、饮食二便、睡眠情志等。
	程度	11	指患者某一疾病或症状进展所达到的状况，如轻度、中度和重度等。
	既往病史	12	指患者就诊以前的病史，特别是涉及心、肺、肝、脾、肾一些重大脏器以及癫痫病史、精神病史等。

5.5 风险因素

中医药文本数据中对不能事先加以控制的因素的描述。

本文件列举了中医药文本数据中较为常见的风险因素实体子类目，包括诱因和高危因素等。如表6所示。

表6 “风险因素”分类名称与代码表

类目名称	子类目名称	代码	说明
风险因素		R	
	诱因	01	指诱导疾病、症状等发生的原因。
	高危因素	02	指影响疾病、症状等发生的高风险因素。

5.6 操作方法

中医药文本数据中有关操作方法表达的信息元素。

本文件列举了中医药文本数据中较为常见的操作方法实体子类目，包括针刺方法、推拿方法和用药方法等。如表7所示。

表7 “操作方法”分类名称与代码表

类目名称	子类目名称	代码	说明
操作方法		0	
	针刺方法	01	指针刺的方法。
	推拿方法	02	指推拿的方法。
	用药方法	03	指用药的方法。

5.7 其他实体

按照分类学要求，设置其他实体子类目，预留未考虑到的实体纳入。