

药品不良反应术语系统构建探析^{*}

林 鑫 郭进京 任慧玲

(中国医学科学院/北京协和医学院医学信息研究所 北京 100020)

〔摘要〕 介绍国际常用药品不良反应术语集 WHOART 和 MedDRA 的概况、层级结构等,对二者进行比较分析,同时从层级结构、语义关系、涵盖内容、对外服务平台几方面来探讨我国药品不良反应术语系统构建。

〔关键词〕 WHO 不良反应术语集;监管活动医学词典;药品监测;药品不良反应;术语服务

〔中图分类号〕 R-056 〔文献标识码〕 A 〔DOI〕 10.3969/j.issn.1673-6036.2019.06.013

Analysis of the Building of Adverse Drug Reaction Term System LIN Xin, GUO Jinjing, REN Huiling, Institute of Medical Information, Chinese Academy of Medical Sciences/Peking Union Medical College, Beijing 100020, China

〔Abstract〕 The paper introduces the general situations and hierarchies of WHOART and MedDRA, which are two international Adverse Drug Reaction (ADR) term sets for commonly used drugs, as well as compares and analyzes these two term sets. It also discusses the building of Chinese ADR term system in terms of hierarchy, semantic relation, contents and service platforms.

〔Keywords〕 WHOART; MedDRA; drug monitoring; Adverse Drug Reaction (ADR); terminology service

1 引言

药品作为疾病治疗的重要手段,有时也会因严重的药品不良反应(Adverse Drug Reaction, ADR)给患者带来损害,对其上市后的不良反应进行监测是药品安全的重要保证。在监测过程中,药品不良反应/事件报告中需要填写患者出现的不良反应/事件名称、过程描述等信息。但在实际工作中,受当地文化、习惯及语言的影响,不同地区的报告者

针对相同类型的不良反应/事件可能会使用多种不同方式来报告^[1],专业术语使用存在不规范的现象。而随着药品市场的全球化,ADR 监测也趋向于国际化,其他国家的不良反应监测信息也是我国 ADR 监测工作的重要内容之一^[2],但与此同时语种问题在一定程度上造成监测报告中术语的多样性。因此将 ADR 术语进行标准化将有助于药品上市后的相关不良反应监测数据互换与共享,能够在很大程度上保证描述用语的一致性^[1]。

国际上有关 ADR 标准表述包括由人用药物注册技术要求国际协调理事会(The International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use, ICH)编制的监管活动医学词典(Medical Dictionary for Regulatory Activities, MedDRA)和由世界卫生组织发布的世界卫生组织不良反应术语集(World Health Organization Adverse Reaction Terminology, WHOART),二者虽在我国已进行应用,但仍存在一定问题:WHO 官方

〔修回日期〕 2019-05-06

〔作者简介〕 林鑫,硕士研究生;通讯作者:任慧玲,研究馆员,硕士生导师。

〔基金项目〕 中国医学科学院医学与健康科技创新工程“中文临床医学术语系统构建研究”(项目编号:2017-I2M-3-014);国家自然科学基金项目“基于自然语言处理的内分泌常用药物不良反应监测数据库的构建”(项目编号:NSFC91846106)。

并未推出 WHOART 的中文版,国内广泛使用的中文版本由国家药品不良反应监测中心翻译,但更新速度缓慢,同时 WHOART 自 2015 年第 1 季度最后一次发布后已不再更新;MedDRA 由于术语较多,编码十分复杂,对其进行查询和检索需借助计算机完成,在推广时容易遭到阻碍,同时国内的营利机构使用 MedDRA 每年需缴纳高昂费用,这也使其在我国很难广泛应用。因此构建权威、规范且能够适应我国 ADR 监测现状的中文药品不良反应术语系统能在很大程度上加强我国药事管理不同组织和部门之间的数据互换和共享^[3],对推动药品临床试验和药品上市后的不良反应监管、促进对 ADR 的深度挖掘利用具有重要意义。基于上述背景,本研究以中国医学科学院创新工程项目“中文临床医学术语系统构建研究”为基础,对 MedDRA 和 WHOART 这两个目前广泛应用的药品不良反应术语集进行调研,总结其收录术语与结构特点,在此基础上对我国药品不良反应术语系统的创建进行探讨。

2 国际常用 ADR 术语集——WHOART

2.1 词表概况

WHOART 是精度较高的用于编码药品治疗过程中不良反应信息的术语集,由世界卫生组织(World Health Organization, WHO)于 1969 年首次发布^[4],涵盖不良反应报告中所需的大多数医学术语。WHOART 结构简单,涵盖术语量较少,具有足够的灵活性,可以行列表的形式打印出来,目前在 WHO 各成员国和全球的监管机构及制药公司应用广泛。WHOART 最大特点是将术语集中的一些术语打上星号表示为重点术语,其作为不良反应术语的子集,可用来指明或指示严重疾病状态,被认为是特别重要且需要重点随访跟进的,如过敏性休克等。这一特点能够为及早识别新的严重不良反应提供便利^[5]。早在 2003 年国家食品药品监督管理局药品评价中心暨国家药品不良反应监测中心便组织有关专家遵循实用性和准确性原则对 2000 年 12 月发布的 WHOART 中的首选术语(Preferred Terms, PT)和系统器官分类(System Organ Classes, SOC)进行翻译,包含 ADR 术语 1 548 条。2013 年国家药

品不良反应监测中心组织专家对其进行更新,增加新翻译的高位语(High Level Terms, HLT)和收录术语(Included Terms, IT),ADR 术语量总计 6 053 条^[6-7],于 2016 年出版。

2.2 层级结构

WHOART 具有 4 级层级结构,分别为 SOC、HLT、PT 以及 IT,WHOART 以 SOC 为最高级进行术语的分组,其具体结构为 SOC → HLT → PT → IT。其中 PT 是用来描述 ADR 的主要术语,IT 是 PT 的近义描述和同义词汇,能够辅助将报告中的不良反应术语正确编码成相应的 PT。WHOART 中 1 个 PT 最多可分配至 3 个不同的 SOC,且其分配到的 SOC 是固定的,不会随报告的改变而更改。WHOART 以“记录号(4 位)+序列号(3 位)”的方式来编码 PT 和 IT,每个 PT 都被赋予 1 个记录号,依据 PT 的引入顺序进行排列,IT 具有与其相对应的 PT 相同的记录号,但 PT 被赋予固定的序列号 001,IT 具有更高的序列号^[8]。SOC 和 HLT 则是利用 4 位数字编码,HLT 通过高位术语链接(HL LINK)将其相对应的所有相关 PT 进行链接,其编号选用所有隶属其的 PT 中记录号最小者^[9]。依据 WHOART 中各术语编码可以看出 PT 和 IT 的隶属关系,能够很容易分辨出该术语所属的层级。

3 国际常用 ADR 术语集——MedDRA

3.1 词表概况

MedDRA 术语集是内容丰富详细且包含大量术语的国际医学术语集,适用于所有医疗产品注册以及上市前后的文档与安全监察,以及毒理外的针对人类使用的所有药品开发阶段,其覆盖的产品包括药品、生物制品、疫苗和药物器械综合产品,包含体征、症状、疾病、诊断、适应症、各类检查名称和定性结果、用药错误与产品质量、各种手术及医疗操作、病史/社会史/家族史这 9 类术语^[10]。其编制目的是为药事管理的整个过程提供经过医学确认的医学术语集。自 2001 年起 MedDRA 每半年(每年 3 月和 9 月)更新 1 次,2009 年 9 月发布的 12.1 版本中增加中文版本。截至 2019 年 1 月最新版本是

2018 年 9 月发布的 21.1 版, 包含英文、日文、中文、捷克文、荷兰文等 11 种语言版本。

3.2 层级结构

与 WHOART 相比, MedDRA 的层级结构更为细致, 包含 5 个层级, 分别为 SOC、高位组语 (High Level Group Term, HLGT)、HLT、PT 以及低位语 (Lowest Level Term, LLT)。具体结构及 21.1 版本中各级包含的术语量为: SOC (27 个) → HLGT (337 个) → HLT (1 737 个) → PT (23 389 个) → IT (79 507 个)。MedDRA 中的每个术语均具有 1 个唯一的非表达性 8 位数字编码, 编码从 10000001 起按字母顺序分配且本身不包含任何信息^[10]。与 WHOART 相类似, SOC 也是 MedDRA 层级结构的最高级, 共包含 27 个, 且按照病因学、发病部位和目的进行分组。HLGT 和 HLT 属于概括性术语, 通过解剖、病理、生理、病因或功能等特点与其下位关联^[11]。PT 则是用来对单一医学概念进行独特、明确且无歧义的符合国际标准的表达, 其特异性和描述性强。LLT 作为 MedDRA 的最低层级, 与其对应的上一层级 PT 之间的关系是同义词、异体词、准同义词、子概念和等同 LLT 这 5 类中的任意一种^[10], 具有最高的特异性。MedDRA 从如 ICD-9 等其他术语集中纳入的术语也在该层级出现, 方便历史数据转换且减少数据录入时的主观性。在 MedDRA 中每个 LLT 仅与 1 个 PT 相对应, 但从 PT 层级起每个 PT 可以与多个 HLT 连接, 每个 HLT 可以连接到

多个 HLGT, 而每个 HLGT 又可以与多个 SOC 相对应。MedDRA 中 27 个 SOC 代表互不排斥的平行轴, 这种 1 个 PT 可隶属于多个 SOC 的特性即为多轴性^[12]。但为避免对全部 SOC 进行逐个检索时出现重复计数, 每个 PT 对应的主 SOC 只有 1 个, 其他 SOC 均被称为该 PT 的次 SOC。术语的多轴对应是在 MedDRA 中预先指定的, 可确保在进行数据检索时无论选择哪个 SOC 均能保证检索的全面和一致。

3.3 浏览工具

MedDRA 开发标准 MedDRA 分析查询 (Standardized MedDRA Queries, SMQ), SMQ 作为 MedDRA 独特功能, 能够方便 MedDRA 编码数据检索, 旨在协助识别和检索可能相关的安全报告, 可用于药物警戒和临床研发过程数据分析的第 1 步^[13]。SMQ 中的术语通常是 PT, 其中的 LLT 也必须是与 SMQ 中的 PT 相对应的术语。此外 ICH 还委托 MedDRA 维护和支持服务组织 (Maintenance and Services Organization, MSSO) 开发两类可供用户使用的浏览器^[14], 一类是可安装在本地计算机上的桌面浏览器, 一类可通过互联网访问。借助桌面浏览器, 用户可对 MedDRA 支持的语言版本中 MedDRA 整个层次结构及相应术语进行浏览与检索, 另外用户还可查看该术语可用的 SMQ 及其包含的内容和术语范围, 以及检索历史等。以先天性免疫性缺陷为例, 其在 MedDRA 中文版整个层级结构中的位置及在桌面浏览器中的检索结果, 见图 1。

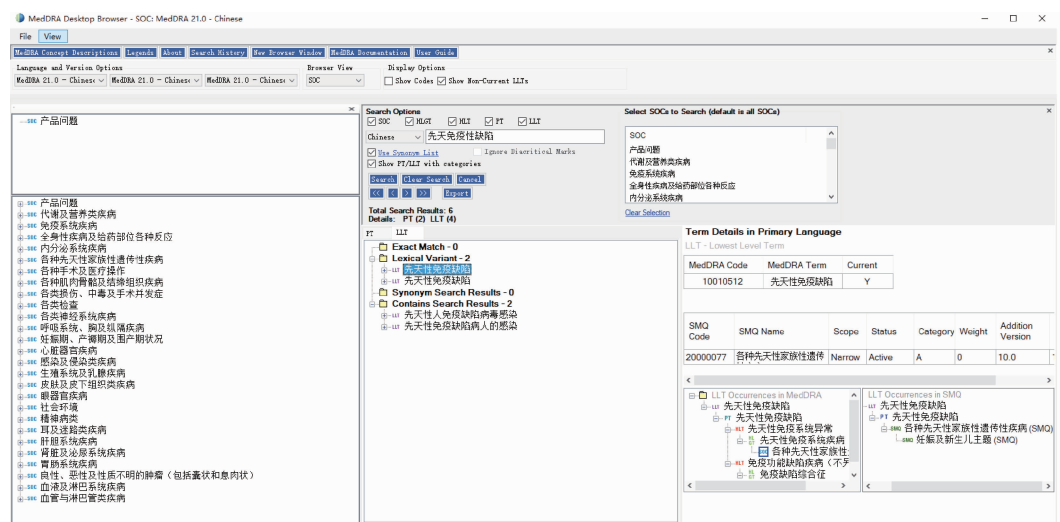


图 1 MedDRA 桌面浏览器

4 MedDRA 与 WHOART 比较分析（表 1）

表 1 MedDRA 与 WHOART 对比

项目	MedDRA	WHOART
术语数量	具有极大的数据量和极宽的范围	首选术语和总术语数量均较少
层级结构	5 级术语：SOC、HLGT、HLT、PT、LLT	4 级术语：SOC、HLT、PT、IT
术语编码	术语的编码本身不包含任何信息	依据编码能够很容易分辨出术语所属层级
应用优势	（1）收录术语范围极宽，大量术语便于编码人员准确编码；（2）能够给药品涉及的合并疾病与病史进行编码，不必借助其他系统；（3）SMQ 便于对数据进行检索，不遗漏相关术语，有利于分析药物警戒中的药物安全性问题；（4）定期更新	（1）术语量较少，方便打印成纸质版进行参考，使用时可脱离计算机；（2）术语较少，在不良反应数据的挖掘与信号预警方面具有很大优势，可以使信号相对集中，提高信号检出的灵敏性；（3）确定需要特别关注的重点术语，提示需要重点关注的安全性信息，使信号筛选更具针对性
现存问题	（1）数据量较大，对其进行查询与检索都需要借助计算机来完成；（2）术语相对较多，信号挖掘时可能会使信号发散，降低信号检出的灵敏性；（3）版本的频繁更新可能会为不良事件的编码和分析带来困难	（1）主要记录不良反应名称，对于疾病和病史的编码需要参考其他术语集如 ICD 等；（2）中文版本难以找到官方版本，不便于非专业人员使用；（3）官方英文版本已停止更新；（4）缺乏细致的使用指导原则

5 我国药品不良反应术语系统构建设想

5.1 概述

我国主要通过自发报告系统收集来自全国的 ADR 报告以对其进行监测，来源广泛，形式多样，存在着信息不完整、专业术语使用不规范等问题。尽管目前 WHOART 和 MedDRA 已应用于国家药品不良反应监测和药物警戒，但由于二者均是由国外机构编制开发，国内使用的 MedDRA 中文版在语言表达习惯等方面与实际描述可能存在一定差异，WHOART 中文翻译版在进行更新时虽然纳入部分中国本地术语^[7]，但之后 WHOART 中文版并未进行再次更新，已不能满足目前国内医疗卫生信息化需求；同时 WHOART 和 MedDRA 目前在国内的应用并不统一，在一定程度上导致国内医疗信息不能互联互通、信息资源共享程度较低；此外 WHOART 和 MedDRA 并不能完全涵盖中医药 ADR 术语，无法完全应用于中药的不良反应监测。随着国内医疗卫生信息化和健康医疗大数据的应用与发展，构建统一、规范且能在我国进行大规模推广应用的药品不良反应术语系统具有重要意

义。除能够对 ADR 专业术语进行规范，也将有助于实现中文病历文本的语义标识和结构化处理，为有关 ADR 的健康医疗数据挖掘提供知识组织工具^[15]。

5.2 层级结构

为构建适应不同需求的 ADR 术语系统并在该系统不同层级上灵活获取相应信息，本术语系统可充分参照 WHOART 和 MedDRA 层级，构建至少包含 4 级的层级结构。其中第 1 级词汇对应 SOC，第 2 级词汇对应 HLT，第 3 级词汇对应 PT，第 4 级词汇对应 LLT。术语系统 4 个层级中的每个术语都具有 1 个唯一的且可以体现其所在层级的数字编码。

5.3 语义关系

在该术语系统中，1 级 SOC 层仍作为层级结构的最高级，2 级 HLT 层中的词汇对 3 级 PT 层中的词汇进行概括性归类，依据解剖、生理、病理、病因或功能等将相关的 3 级词聚为一类，3 级 PT 层中的词汇是用来描述 ADR 的正式主题词，4 级 LLT 层中的词汇作为 PT 层级的下位类，用来收录相应 PT

的子概念、近义描述和同义词汇等。参照 MedDRA, 该术语系统也具有多轴性, 1 个 PT 可能会隶属于多个 SOC, 但每个 PT 对应的主 SOC 是唯一的, 见图 2。

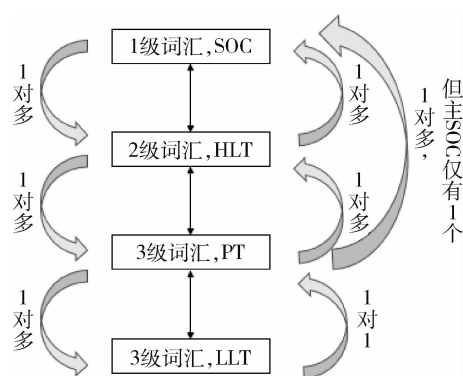


图 2 术语系统语义关系及多轴性

5.4 涵盖内容

相对来说 MedDRA 包含的术语范围比 WHOART 更为广泛, 在术语选择、查询和编码方面更具优势。但 WHOART 在我国已使用多年, 上报时的不良反应/事件名称以 WHOART 中的 ADR 术语作为参考, 报告者和监测人员对其比较熟悉。因此该术语系统在构建时可以 WHOART 中文更新版中收录的 4 级术语为基础, 参照不良反应报告实际情况, 结合 MedDRA 和前期构建的与国家卫生信息化和标准化兼容的中文临床医学术语系统对前 3 个层级的术语进行调整与补充, 归类到相应的上位类下, 对 PT 层级术语的相关同义描述进行整理并补充到相应的 LLT 层级。中文电子病历中包含药物过敏史、病史等多种信息, 是 ADR 记录的重要来源。描述性电子病历作为目前广泛采用的电子病历系统, 能提供更多有关 ADR 的信息, 包括不良反应种类的详细描述等^[16], 此外药品说明书中也包含丰富的不良反应信息。因此术语系统在构建时也要充分收集去隐私化的电子病历和药品说明书, 对其中有关的 ADR 信息进行整理, 收集临床中实际出现的规范与不规范的 ADR 描述及说明书中的 ADR 描述, 将其对术语系统进行补充。这需要国家鼓励国立

医院或医药企业对数据进行积极公开和共享, 尤其是官方数据的公开发布和及时更新等。目前我国尚未有权威、统一的中医药不良反应术语集可用于中医药临床试验和药品上市后不良反应的监管, WHOART 和 MedDRA 也并不能完全适用于中医药的不良反应描述。对于中医药部分 ADR 术语, 与术语系统中已有术语表达完全一致的名词, 如“胸闷”、“心悸”等, 这部分术语可被术语系统涵盖使用; 与术语系统中已有术语内涵一致或相近但表达方式不一致的, 如中医术语“欲寐”和 PT 术语“嗜睡”类似, 可将此类术语进行整理, 将其作为 LLT 对术语系统进行补充; 与术语系统中的术语内涵与表达均不一致, 属于中医特有症状的, 如舌象等, 则在术语系统中新增 1 个 SOC 中医, 组织相关领域人才对术语系统第 2 级及以下的术语类别进行划分, 结合中医临床术语系统、国家标准、中医药词典等对中医类的 ADR 术语进行补充。

5.5 对外服务平台

药品不良反应术语共享服务平台是药品不良反应术语系统对外服务的直接门户, 预期面向我国电子病历、个人健康档案、临床研究等应用需求提供服务。核心内容是围绕 ADR 术语自身维护、检索与浏览以及应用等来进行 ADR 术语整个生命周期的管理^[17]。平台对外将提供 ADR 术语的检索查询、导航浏览、详情展示、下载以及第 3 方调用等服务, 同时在开放共享过程中平台还将制定保障术语共享权益的相关协议。

6 结语

药品不良反应术语集有助于病例报告标准化编码, 从而能够进一步提高对相关数据研究的准确性和利用率。WHOART 和 MedDRA 是国内外广泛使用的药品不良反应术语集, 本文从概况、收录术语与结构特点、相关软件工具等方面对二者进行介绍, 在此基础上对适用于我国的药品不良

反应术语系统构建进行探讨。随着药品不良反应术语系统的实际应用,系统内容以及共享服务平台功能也将不断进行完善,通过实际应用不断发现问题和需求,以尽可能最大程度适应我国药品不良反应术语的特殊性、国际交流以及国内医疗卫生信息化的需求。

参考文献

- 1 吴桂芝,田春华,王丹,等. WHOART 和 MedDRA 在药品不良反应监测中的应用 [J]. 中国药物警戒, 2010, 7 (2): 81-85.
- 2 国家药品不良反应监测中心. 药品不良反应信息通报 (第 15 期) [EB/OL]. [2019-01-15]. <https://wenku.baidu.com/view/662885acaef8941ea76e0586.html>.
- 3 刘晶,谢雁鸣,盖国忠,等. 药品不良反应术语集 WHOART 与 MedDRA 的应用探析 [J]. 中国中药杂志, 2015, 40 (24): 4728-4733.
- 4 马丹华,王玲,孙骏,等. 我国药品生产企业不良反应术语认知及使用现状调查研究 [J]. 药学与临床研究, 2017, 25 (4): 95-98.
- 5 林伟兴,叶小飞,姚洪祥,等. 药品不良反应术语集现状分析 [J]. 中国药物警戒, 2009, 6 (12): 713-716.
- 6 刘佳. 药品不良反应监测近期工作进展及相关要求 [EB/OL]. [2018-12-27]. <http://www.doc88.com/p-4347805293381.html>.
- 7 国家药品不良反应监测中心. 《WHO 药品不良反应术语集 (中文更新版)》定稿及专家咨询会在京召开 [EB/OL]. [2019-01-17]. http://www.cdr-adr.org.cn/xwdt/201302/t20130204_5326.html.
- 8 汤小平. WHOART 和 MedDRA 在药品不良反应监测中的应用 [J]. 中国现代药物应用, 2015, 9 (14): 285-286.
- 9 潘文. 药物不良反应层级系统的构建与分析 [D]. 厦门: 厦门大学, 2011.
- 10 ICH. MedDRA 入门指南. 21. 1 版中文 [EB/OL]. [2018-12-07]. <https://www.meddra.org/how-to-use/support-documentation>.
- 11 李庆娜,陆芳,赵阳,等. 构建基于 MedDRA 术语的中医临床试验编码术语集初探 [J]. 中国药物警戒, 2015, 12 (1): 53-55.
- 12 卜擎燕,熊宁宇,邹建东,等. ICH 国际医学用语词典 (MedDRA): 药事管理的标准医学术语集 [J]. 中国临床药理学与治疗学, 2007, 12 (5): 586-590.
- 13 ICH 工作办公室. 监管活动医学词典 (MedDRA) [EB/OL]. [2019-01-14]. <http://www.cde.org.cn/ich-Web/learnMedDRA/learnMedDRA.jsp>.
- 14 ICH. Browsers [EB/OL]. [2019-01-14]. <https://www.meddra.org/browsers0>.
- 15 任慧玲,郭进京,孙海霞,等. 医学术语标准化研究的思考 [J]. 医学信息学杂志, 2018, 39 (5): 2-7.
- 16 Skentzos S, Shubina M, Plutzky J, et al. Structured vs. Unstructured: factors affecting adverse drug reaction documentation in an EMR repository [J]. Amia Annu Symp Proc, 2011, 2011 (2011): 1270-1279.
- 17 何清行. 语义相似度算法研究与医学术语服务平台的实现 [D]. 长沙: 湖南大学, 2017.

《医学信息学杂志》开通微信公众号

《医学信息学杂志》微信公众号现已开通,作者可通过该平台查阅稿件状态;读者可浏览当期最新内容、过刊等;同时提供国内外最新医学信息研究动态、发展前沿等,搭建编者、作者、读者之间沟通、交流的平台。可在微信添加中找到公众号,输入“医学信息学杂志”进行确认,也可扫描右侧二维码添加,敬请关注!



《医学信息学杂志》编辑部