

0421763

生物医学文献 主题标引

INDEXING MANUAL OF BIO MEDICAL
LITERATURE SUBJECT HEADINGS

• 主编 李守谅



湖南科学技术出版社

生物医学文献 主题标引

主编 李守谅

编者 萧晓旦 王秀峡

李晓萍 王 唯

[湘]新登字 004 号

生物医学文献主题标引

主编:李守谅

责任编辑:王一方

*

湖南科学技术出版社出版发行

(长沙市展览馆路 3 号 410005)

湖南省新华书店经销

湖南省科技馆青苹果数据中心排版

长沙银都教育印刷厂印刷

*

1992 年 2 月第 1 版第 1 次印刷

开本:850×1168 毫米 1/32 印张:16.25 字数:408 千字 印数:1-2000

ISBN7-5357-0958-3

R·208 定价:8.80 元

地科 88-24

前 言

《生物医学文献主题标引》一书,是由湖南医科大学、白求恩医科大学、同济医科大学及中国医科大学医学图书情报系协议共同编写,湖南医科大学图书情报系主编,供四校五年制医学情报专业学生共同使用的专业课程教材。

本书参照中国医学科学院及中国中医研究院建立中国生物医学分析检索系统的标引规范,将传统中国医学与现代医学的标引规则融为一体,以保证学员通过学习本教材,准确掌握中国生物医学文献主题标引规范。

中国生物医学文献主题标引规范,以国际上较广泛参照的美国国立医学图书馆(NLM)编制的《医学主题词表》(MeSH)及《标引手册》(《INDEXING MANUAL》)为主要参考依据,并将国际医学界日益重视的中国传统医学文献的主题词收入,以利于中国与国外医学情报交流,节省逐年维修词表及标引规则所需花费的大量人力、物力与时间,并保证中国医学文献主题标引能及时反映生物医学新进展,它符合“洋为中用”“推陈出新”的方针。有鉴于此,除了传统中国医学部分参照中医研究院制订的词表,使用汉字标引外,所有现代医学主题词均附有英文对

照,同时将副主题词、重要的标引术语以及《医学主题词注释字顺表》(MeSHAAL)中注释略语,均用中英文对译,并依英文字顺排列副主题词,这样做的目的,是考虑当前读者大都具有和正在努力掌握一定的英语基础,而我国医学情报界不可能将所有 MEDLARS 修订部分及时翻译。因此,培训学员向直接阅读英文医学标引规范及主题词过渡,是图书情报界及检索用户使用现代化检索医学文献的迫切需要。而使用汉字系统对用户来说,则是已经掌握或稍经简短的学习即可轻而易举地掌握。因此本书主要目的之一是:为学员直接查找逐年更新的《医学主题词表》或注释(英文版),提供初步训练,并使学员接触和熟悉部分英语专业词汇,为将来直接用英文标引我国生物医学文献,实现国际资源共享打下一定的基础。至于引用的现代医学的标引实例,大都沿用美国国立医学图书馆《标引手册》所列题例,这是因为标引教学,贵在引导学员了解标引入门的种种原理、方法,而标引人员总是遇到新鲜的事物,只要掌握原理,自能收举一反三之效。

本书还力图帮助学员了解和掌握中国生物医学文献主题标引规范的基本原理,熟悉各项具体标引规则及其依据;对某些标引通则的例外规定,诸如一般规则中应避免重叠组配,却又允许属于泌尿生殖生理范畴(G8)的妊娠(PREGNANCY)与/生理学组

配等等也作了必要的解释。本书还就《医学主题词注释字顺表》及树状结构表的结构与功能,及有关工具书的使用以实例作了详细介绍,这是标引人员及检索者查全、查新、查准的重要手段,也是标引入门必经之路。因此,本书还通过对反映新概念、新进展诸如分子生物学、免疫学的文献的某些难理解的部分篇名作出注释,以启发学员在理解生物医学文献的中心主题基础上准确标引。本教材的上述尝试性的编辑设想,是经过多年标引教学实践所证明的:它能使复杂的标引规则简明化,容易被理解消化,从而启发学员独立思考的能力。由于本书是从文献主题标引的逻辑原理阐述其标引规则和词表的功用,因此,它能使学员易于适应今后词表及标引规则逐年的动态性维护所带来的变化。

此外,本书对于各类生物医学数据库标引人员的培训;图书馆生物医学图书资料主题编目;生物医学期刊杂志编辑部编制关键词、主题索引及医学文献检索用户,均具有参考价值。

由于本课程具有高度实践性,因此本教材在阐述问题时,均援引具体实例,并附有适量习题供学员练习之用。

全书内容反映了1991年《美国生物医学文献分析检索系统》的最新修订规范,但也详述了1990年以前固有的属分参照等,以便于历史性回溯性检索。

总之,本书的上述编辑设想,是为适应当前我国标引人员还是处于实践经验少,而又要担负全科标引重任的需要和现状而确立的。这与美国已具有大量熟练且分专科标引人员的现状迥然不同。

本书的出版,承蒙湖南省新闻出版局,美国国立医学图书馆国际联络部主任 Richard Hsieh 博士、湖南省图书馆馆长常书智副研究员,湖南医科大学图书情报系主任柳晓春副教授、湖南医科大学唐五一医师等在国外资料提供、学术评价论证及出版、排印等方面给予热情支持与关注,此外,还要感谢湖南省科技馆青苹果数据中心戴向阳、谢萍、杨柳、何国荣在电脑录入排版过程中的辛勤劳动,谨志衷心的谢意,萧晓旦同志在原稿整理与篇章改写上,做了不少工作,并志。

李 宁 谅

1991 年 8 月

于湖南医科大学

目 录

第一章 绪论	(1)
一、 医学文献主题标引的发展与时代背景	(2)
二、 计算机在医学标引中的运用	(4)
三、 中国生物医学文献标引规范的设计依据	(6)
四、 中国生物医学文献分析检索系统的建设与展望	(9)
第二章 医学主题词注释字顺表及树形结构	(13)
第一节 医学主题词注释字顺表	(13)
第二节 医学主题词树形结构	(35)
第三章 医学文献标引辅助工具书及参考书	(39)
第一节 技术札记——MEDLARS 标引指导	(40)
第二节 技术札记补篇	(41)
第三节 医学主题词——补充化学记录	(46)
第四节 医学主题词轮排表	(48)
第五节 医学主题词词义注释	(51)
第四章 医学文献主题词标引原则与步骤	(54)
第一节 标引的一般原则	(54)
第二节 主题标引的步骤与方法	(58)
第五章 特征词	(88)
第一节 特征词的含义及其作用	(88)
第二节 特征词的标引规则	(92)
第六章 文献类型	(118)
第七章 组配	(127)
第八章 副主题词	(143)
第一节 概论	(143)
第二节 副主题词的组配范畴	(144)

第三节	副主题词的树形结构及功能	(151)
第四节	副主题词标引原则	(154)
第九章	主题词标引原则	(252)
第一节	A类 解剖学主题词标引原则	(252)
第二节	B类 有机体主题词标引原则	(271)
第三节	C类 疾病主题词标引原则	(284)
第四节	C ₁ 类 肿瘤主题词标引原则	(326)
第五节	D类 化学物质和药品主题词标引原则	(337)
第六节	E类 技术类主题词标引原则	(379)
第七节	F类 心理和精神病主题词标引原则	(397)
第八节	G类 生物学科主题词标引原则	(401)
第九节	H类 自然科学主题词标引原则	(411)
第十节	I类 社会科学主题词标引原则	(415)
第十一节	J类 工艺、工业、农业、食品类主题词标引原则	(424)
第十二节	K类 人文科学主题词标引原则	(429)
第十三节	L类 信息科学主题词标引原则	(433)
第十四节	M类 特殊职称主题词标引原则	(438)
第十五节	N类 保健医疗主题词标引原则	(446)
第十六节	Z类 地理名称主题词标引原则	(452)
第十章	中医药文献主题标引规范	(458)
第一节	词表及参考书	(458)
第二节	中医药副主题词	(459)
第三节	各类中医药文献的标引	(464)
练习		(485)

第一章 绪 论

顾名思义,标引是设立标识以引导人们在复杂广泛的事物中,准确地寻找达到目的的便捷途径。诸如:铁路与公路的路标,江海的航标与灯塔,城市道路的标牌与门牌号码……。而浩如烟海的图书资料,就更需要使用科学的、便利的标识(tag),引导读者尽快地、全面而准确地查找到所需的文献资料,而文献标引(INDEXING)就是基于这种需要而产生的一门学科。

在情报学界,各国对标引有着不尽相同的定义,我国国家标准局GB3860—83《文献主题标引规则》对标引一词的定义是:“标引一是对文献进行主题分析,从自然语言转换成规范化的检索语言的过程。对文献给予分类号标识的过程,称为分类标引;给予主题词标引的过程,称为主题标引。”我们认为这一定义简单明确而具体,既符合我国图书情报检索的实际需要,也与当前国际情报检索语言的发展趋势一致。它体现了国内外分类法与主题法检索标识并用的现实情况。

目前,世界上使用的检索标识都包括:外部特征与内容特征。外部特征标识包括:著者、文献号及其它标识;内容特征主要包括两类标识,即:分类号标识和主题词标识。主题词标识又有四种不同的类型。其中自然标识为关键词,人为标识的有:单元词、标题词及叙词。

本书所阐述的是我国使用的,也是国际医学图书情报界认为最优化的、附有树形结构分类的叙词法(descriptors method)标引系统。这个系统,吸取了美国国立医学图书馆《医学文献分析与检索系统》(MEDLARS)编制的标引规范(IDEXING MANUAL)的优点;依国际《医学主题词表》的体例,按英汉对照拼音字顺排列主题词。主题词包括由中国中医研究院制定的《中医药主题词表》及美国国立医学图书馆(NLM)编制的《医学主题词表》(MeSH)组成;而《中医药主题词

表》虽依汉语拼音字顺排列,但附英文译名,且在 T 类下位设置与 MeSH 表兼容的树形结构,并增加了仅适用于中医的十个副主题词,必要时可允许在以上两种词表外,增加自由关键词,从而形成既可与国际联机数据库兼容或相互变通,又具有中国传统医学特点的中国生物医学文献分析检索系统。

一、医学文献主题法标引发展史略与时代背景

我国早在两千年前,就已使用分类法作为标引方法。西汉年间(公元前 32—公元 23 年间),刘向、刘歆相继整理六艺群书,刘歆的《七略》,就是我国最早的图书目录分类著作,他将当时图书分为七大类,其中医学书籍编入方技略。至隋代(公元 6 世纪末),将图书分为经、史、学、集四大类,医书归入学部医家类;经典著作如《黄帝内经》,也编入官史目录。至唐代,由于科学技术的不断发展,图书资料迅速增多,单纯分类法已不能满足检索的需要,促使标引的结构与方法不断完善。公元 7 世纪,颜真卿所著《韵海镜源》,是以声韵为字顺查找事物的检索目录工具,该书“依四声切韵增出 14,761 字,引诸家著录,并将四部中两字以上成句者,依句尾编入本韵”,该书是标题法的鼻祖。明代(公元 1403—1408 年)编纂的一部带有主题法性质的检索工具,世界名著《永乐大典》,及清代乾隆(公元 1783 年)编纂的《四库全书总目提要》即以《韵海镜源》为蓝本的。

在古代西方,公元前一世纪,塞涅卡(Lucius Annaeus Seneca 就记述过一个哲学家索引表,据 H. B. Whatley 说,西塞罗(Marcus Tullius Cicero)描述那时的索引只是夹在书籍中的羊皮纸签。直到 1611 年在《大不列颠史》中,才附有“本书主要事件字顺索引”,由上可见,中国在标引的设计及实践方面,在世界上居领先地位。而颜真卿所著《韵海镜源》,特别是《永乐大典》,其包罗主题的宽度与深度,及所附著录及引文之完整,均远非西方可比。然而,由于汉字书写方法及音韵均极为复杂,《永乐大典》虽设有四声 76 韵,仍难以确切地排列所有的汉字,而且这种声韵法,查找比较复杂,又无上、下位等级分类,难以

适应现代科学发展与国际国内文献资料交流的需要。

在西方,英国 17~18 世纪的产业革命,推动现代科学技术的迅猛发展。自然科学学科不断分化,交叉学科、边缘学科相继涌现,西方图书资料沿用的分类法尽管作了不少改革,也难以适应查找专指性与交叉性文献的需要,而医学是奠立于广泛自然科学基础上的学科,从而也随着自然科学文献标引而革新。于是,继英国克里斯塔多罗(Crestadoro)于 1854 年提出标题字顺标识系统后不久,华盛顿军事医学图书馆的 John Shao Billings 又提出更完善的主题词标引体系。他自 1865 年开始,用标题法标引军事医学图书馆藏的医学书刊,成为现代医学文献主题标引的创始人。他这一工作受到美国情报界的高度重视,自 1880~1975 年,美国图书情报人员依英文字顺,逐年出版《索引编目》(INDEX CATALOGUE)(1880 年从 A 编到 BERLIMSKY,至 1975 年全部完成 26 个字母顺序的全部馆藏医学图书编目,网罗图书达 305,000 册)。随后,于 1979 年,每年由 A—Z,依字顺出版当年收到的医学书刊,编制出版了《医学索引》(INDEX MEDICUS),该工具书在标题下,依字顺编列作者名,在作者名下又依字顺罗列其著作题录。

《医学索引》使用的标题法是主题法的早期类型,它是抽出了文献标题而形成的、一种手工式和先组式的主题检索方法。所谓标题(SUBJECT HEADINGS),是采用事物自然语言的“名”,经过统一规范化后,用作在检索工具中概括和表达图书资料内容的主题标目或标识。步入 20 世纪后,科学技术更加迅猛发展,出现了书刊数量及文献数量的急剧增长,其增长程度从下列数据得到反映。

各国各科期刊出版增长情况(约数):

年代	期刊种数
本世纪 50 年代	20,000
本世纪 60 年代	40,000
本世纪 70 年代	80,000

本世纪 80 年代 200,000

美国《化学文摘》收录 100 万份文献所需时间:

年限 积累 100 万份文献日期

1907~1938 32 年

1938~1956 18 年

1956~1963 8 年

1963~1968 4 年 8 个月

1968~1971 3 年 4 个月

1971 年以后 2 年

不仅期刊与文献数量迅速增长,且文献内容及涉及面也日趋广泛;学科的横向发展与相互渗透,使得主题概念更加复杂,有不少主题词,往往跨越几个学科领域。使用标题法标引文献所形成的检索工具,无论从文献的多面性、传递速度来说,都难以满足用户的要求。恰恰在这一时刻,电子计算机进入了图书资料检索体系,从而促使主题标引进行了一次重大变革,适应了时代的需要。

二、计算机在医学文献标引变革中的运用

本世纪 60 年代以来,电子计算机开始运用于图书情报检索系统,标引使用的情报检索语言也随之变革。本世纪 50 年代初期,陶布(M B Taube)等提出了一种主题检索方法,叫单元词法(uniterms method)。单元词是指从图书文献的主题内容中,抽出最基本的词素,即在字面上不能再分解的单词作为一个描述主题的概念单元。例如山、水、结核、肝炎、疝气、瘰管……,都是不能或不易再拆开的词,否则在专业上就不具有独立意义了。尽管单元词法在以后的计算机检索应用中,由于词间次序或语义关系缺乏控制而产生许多虚假联系,从而被新的词法所代替,但它所具有的灵活性、多元性等特点,为电子计算机检索开辟了一条新的道路。

1960 年,美国国立医学图书馆使用电脑穿孔、编序及柯达(KODAK)光电系统编制卡片,每分钟达 230 张,大大提高了编目效率;并

且开始为电子计算机检索设计叙词词表及标引规范。1961年，~~他们~~推出了《医学主题词表》(MEDICAL SUBJECT HEADINGS, 简称 MeSH)及《标引手册》(INDEXING MANUAL), 用来编制《~~医学索引~~》。1966年, 为了建立联机检索数据库, 他们又对《医学主题词表》作了详细的注释; 同时使用 1975 年研制的副主题词(Subheadings)或称限定词(Qualifiers)系统; 另行出版了《医学主题词注释字顺表》(Medical Subject Headings Annotated Alphabetic List), 简称 (MeSHAAL), 及《医学主题词树形结构》或称范畴分类表 (Medical Subject Headings—Tree structures); 另外, 还编制了一系列标引工具书, 供标引, 编目及联机检索人员使用。上述一系列标引工具, 均逐年或在必要时进行维护性修订, 这是一个及时反映医学进展, 适用于机械与手工检索的完整的叙词标引系统, 这个美国人工标引系统的专用名称叫《医学文献分析与检索系统》(Medical Literature Analysis and retrieval System), 简称《MEDLARS》。该系统于 1966 年建立数据库文档, 并且于 1971 年开始联机检索服务。

自开展电子计算机与手工检索并行运作以来, 世界上使用不同的标引系统, 大体上分为两种。一种是上述的叙词法 (Descriptors Method); 另一种是关键词法 (Keywords method), 例如美国《生物学文摘》(BA)及《化学文摘》(CA)均使用关键词法。关键词法的特点在于: 使用计算机扫描出文献中的关键词, 并给予文献编号; 而美国《生物学文摘》还在关键词前后, 各抽出一段前后缀语, 作为该关键词的标引说明, 称之为“题内关键词索引”。这种标引方法的优点是: 不用人工标引, 节约大批人力, 加快文献标引与传递速度。但这种方法仍存在一些问題, 有待进一步解决, 一是文献中抽提出的关键词是自然语言, 它有时只是几种同义词中的一种; 例如一篇有关肾结石的文献, 如作者使用 (Renal stone) 则抽出的是 RENAL STONE; 如果使用的是 Kidney stone, 则抽出的关键词是 KIDNEY STONE; 阿司匹林与水杨酸钠也是这样。因此, 在成族文献检索时, 同一类文献被分散在不同词

形的同义词款目下,从而降低了查全率。为了解决这一问题,必须在软件设计中,建立所有同义词互通程序,并加大关键词库的容量;二是电脑没有任何语义与词法限定的智能分析,不同前后缀语的关键词词序不明确,支持性不够;使用题内关键词索引,虽能部分解决上述问题,但也由于文献中文字写作结构因作者而异,一种关键词由于前后缀不同,可能被抽出几次,但表达出的只是同一种概念,使标引累赘;另一方面,有一些重要叙述性概念,则因计算机词库难以周全,在抽提时,一些款目难免被遗漏;三是《题内关键词索引》,无论在机械或手工检索中,贮存量或所占印刷篇幅都很大,噪音多,给用户增加不少不必要的负担,因此,包括我国、日本、欧洲、美洲,多数倾向使用人工标引的叙词法体系。

三、中国生物医学文献标引规范的设计依据

中国生物医学文献标引系统,系由中国医学科学院及中国中医研究院联合设计建立起来的。该系统汲取了美国“医学文献分析与检索系统”及《汉语主题词表》的优点,是中西医并蓄、汉英字顺对照、综合了标题词、单元词及关键词诸法精髓,适应现代科技进步,既利于手工,更适用于计算机传输,并能最高地达到查全率(Recall factors)及查准率(Relevant factors)的先进标引系统。

— (一)特点及优越性

1. 本系统的叙词(discriptors),汲取了分类法的范畴及等级性,将主题词纳入一个或多个不同范畴的树形结构,以利于限制副主题词组配范围,使之逻辑性严密,便利单面或多面成族。这种树形结构,分级不受限制,可以伸缩自如,加之其每级编码可达三位数字(千进位),从而为学科的发展及主题词的扩充,留有充分的维修与扩展的余地。

2. 该叙词系统,既具有单元词的组配功能,在计算机检索中,可作为后组式检索语言,以弥补主题词表的缺陷;或单纯作为参词,以便尽可能专指性地表达文献内容。除了单元词外,又设有大量文献中

出现频率较高的复合主题词,从而加速了族性检索的专指范围。

3. 本系统的叙词表,具有严格的受控性或规范性。表现为:

(1)对主题词的数量及规范作了严格的限制。主题词的增删,是按照文献中出现的频率及专家的鉴定而确定的。现代医学主题词的数量限制在 16,000 左右。这样有限的主题词及其相互组配,加上与 80 个副主题词的方面组配,基本上可以表达现代生物医学领域文献中一般性或专指性概念;除上述现代《医学主题词》外,本系统还另编有《中医药主题词表》,含主题词 4,800 个左右,及附加的十个副主题词,可以标引出汉语中医的基本主题概念。由于主题词数量的严格控制,使标引人员容易通过实践,熟悉有限的词表定义及标引规范,从而提高标引效率,缩短输入周期;这对联机检索人员的咨询及制定检索策略,也同样简便而易于掌握与运用。

(2)健全的参照系统,本系统分三个方面:一是“见”(see)“代”(X)关系,即把一个规范名词代替多个同义词,该同义词仍以小号楷体排入主题字顺,注明见××……。这就避免了同一族文献,分散标引在不同形式的同义词下,从而提高查全率,避免关键词法标引在这方面的缺陷。二是属(see under)分(xu)参照,这是为揭示主要和次要叙词概念之间的语义关系的参照措施。这种参照,揭示了叙词的等级位置,明确了印刷本与机器检索的界限。虽然 1991 年起这种属分关系已经取消,但对 1990 年以前的文献回溯仍具有作用。三是相关参照,这不仅扩大了标入口及检索出口,而且列出的相关参照主题词,有助于近义主题词的辨义,有利于精确运用标引词,以提高检索命中率。

(3)范畴分类表限定了方面组配:本标引系统将主题词分属于 A,B,C,D……N,T 及 Z16 个范畴类别。一般来说,每个范畴或范畴亚类中的主题词,只能与某个或某些特定的副主题词组配,其中个别例外则使用标引注释加以规定。这种范畴分类,不仅使主题词在族性、等级上精确划分,而且可以通过范畴的控制消除矛盾,重叠或无

效概念组配,保证标引的准确与一致。

(4)本标引系统有一系列辅助参考工具书(详见第三章),将化学、肿瘤、真菌等学科浩繁的名称,归纳于有限的医学主题词下,以保证准确地向上靠位;而像《补充化学记录》(Supplementary Chemical Records),则可将两万余种药物或化学物质的专门名词,直接标引输入计算机检索数据库,又可归纳入手工检索本(IM)的约七千个上位主题词。

由上可见,本系统设置的种种控制与条件,其目的就在于保证语义与逻辑精确、简便,从而提高人工标引效率,最大限度地提高查全率与查准率。

3. 每年对本标引系统,进行一次动态性维护,包括主题词、副主题词的增删或定义的变更;某些参照关系或树形结构的改变;标引注释包括组配范围的改变……。这就要求对医学主题词表、树形结构、各种辅助工具书相应的机控程序等……都每年或定期修订一次,从而使本标引系统能及时适应生物医学领域的新进展,包括研究重点和文献表达方式的变化。

4. 本标引系统的叙词能直观地表达事物,加上利用美国国立医学图书馆每年修订的《医学主题词释义》(MeSH Scope Notes)对每个主题词,均给予规范定义,使词义明确,标引误差少。

中国生物医学文献标引系统,虽然在标引规范及现代生物医学主题词等方面,吸收了美国国立医学图书馆 MEDLARS 的优点,但也充分反映了我国实际情况,除了增加《中医药主题词》及副主题词外,对中医标引规则,标引单也作了修改;特别是地理主题词,MeSH 表中,对标引美国地名可达到州及大城市一级,但对其他各国,特别是发展中国家,仅列国名,这就难以反映我国流行病或人种学等方面的地区特征。例如中国鼻咽癌的发生,限于南方几个省;1988 年上海市甲型肝炎流行,显然不能笼统地标为中国。因此,我国文献标引的地理名词必然要增加首都北京及其他直辖市、名省、自治州及其省会或