

医学文献关键词标引中的组配标引探析

金永勤

温州医学院期刊社, 325035, 浙江温州

摘要 探讨文献主题标引中组配标引的概念及利弊、组配标引的理论基础、组配标引的2种形式及规则,并对组配标引作了必要的补充。

关键词 医学期刊;期刊文献;组配标引;关键词

Analyze the combination indexing of medical journal articles//

JIN Yongqin

Abstract This article analyzes the combination indexing of medical journal articles and its advantages and disadvantages. This article also describes its theory foundation and two styles and rules. Furthermore, this article makes essential descriptions of combination indexing.

Key words medical journal; journal article; combination indexing; key words

Author's address Periodicals Office of Wenzhou Medical College, 325035, Wenzhou, Zhejiang, China

关键词作为文献检索的标志,其选择正确与否直接关系到文献检索的效率,影响到科技文献的有效传播和利用^[1],也就是实际上要求以主题词为准。组配标引是主题标引的一种基本规则,是加强文献主题表达的专指性、揭示主题与主题间逻辑关系和多向成簇的多途径检索手段^[2]。大多数医学期刊已利用《医学主题词表》(MeSH)、《中医药学主题词表》上规范化的词汇进行标引,以提高期刊的标引水平。科技期刊论文标引来自图书情报的文献标引,那么就应该按文献标引的规则进行,特别是在进行论文主题分析和关键词的选取上,应遵循其基本的方法和步骤^[3]。在科技信息迅猛发展的今天,全世界每天有十几万篇科技论文发表,学术界已约定利用主题概念词去检索最新发表的论文^[4];因此,医学期刊编辑应当予以足够的重视,如何能进一步地提高医学期刊的标引水平和质量,扩大其在国内外的影响,是医学期刊编辑面临的重要课题之一。

1 组配标引的概念及利弊

文献的主题内容和某些具有检索标志记录下来的文献处理过程,即为标引。用主题词作为检索的标引,为叙词(正式主题词)标引。现在所谓的对医学文献关键词标引,实际上就是叙词标引和自由词标引。组配标引是主题标引的一个重要特点,只是极少数的文献在标引时是不需要组配的。标引工作实质上是:一个是选择叙词,另一个是组配。在标引实践中,为了表

达某个复杂概念,而词表中又没有形式相应的恰当的词,标引人员运用2个或多个有形式逻辑关系的主题词,按一定的规则复合起来,表达一个更有专指性的主题概念,这就是组配标引。

影响组配标引的因素较多,组配又成了标引中最容易产生误差的环节;因此,人们只好把组配方面的标引规则定得更为严格一些。

组配有对文献标引带来专指性强、反映内容全面、可进行多途径检索等好处,但组配本身也不是没有弊病的。在实际标引中,由于种种原因会造成组配误差,如重叠组配、无意义组配等^[5]。例如:灾害/利用——矛盾;冠状动脉/血液供给——重叠。

2 组配标引的理论基础

2.1 与组配有关的逻辑关系

1)逻辑与(写成 A AND B, A、B 是 MeSH 叙词)。同时满足 A、B 这 2 个条件,对一篇文献资料来讲,就是该文献既满足 A 的条件又要满足 B 的条件^{[5]53}。如《柴蒿三黄汤联合环丙氧酸治疗伤寒疗效观察》可采用“逻辑与”,即“柴蒿三黄汤 AND 环丙氧酸”。

2)逻辑或(写成 A or B)。满足 A、B 这 2 个条件之一即可。对一篇文献来讲,就是满足 A 或 B 的条件,或者 2 个条件均满足即可^{[5]54}。如“甲状腺疾病与雌激素受体与孕激素受体”表示成“甲状腺疾病 AND (受体,雌激素 or 受体,孕激素)”。

3)逻辑非(写成 A NOT C)。表示必须满足条件 A,但要排除满足 C 条件的那部分。对一篇文献来讲,就是满足 A 条件,同时又要不满足 C 的条件^{[5]54}。如关于“结核性心内膜炎,不包括心律失常”的有关文章可采用“逻辑非”,即“结核,心血管 AND 心内膜炎 NOT 心律失常”。

2.2 组配标引实质上主要是概念组配 组配标引主要以形式逻辑为基础,概念组配主要是根据概念交叉关系以及概念限定方法,将复杂概念分解为单元概念进行组配。当一个文献主题有多种组配标志时,根据词表收词情况,应首先选用具有交叉关系的叙词,进行概念交叉组配;只有在词表中找不到相应的词进行交叉组配时,选用概念限定组配。无论是概念相交组配还是概念限定组配,都要选择能确切表达文献主题的

组配标志。

2.3 组配标引的要点

- 1)以概念之间的逻辑关系为基础。
- 2)以概念分解为手段。
- 3)以解释概念的本质为目的。

2.4 组配标引的特点

- 1)表达概念时,接近思维逻辑。
- 2)组配所得的概念意义明确,更适应用户从任意角度进行检索提问。如标引一篇题为《门静脉阻断与再开放时肺循环血流动力学动态变化的实验研究》的文献,就选出了有关概念交叉的叙词“门静脉、肺循环、肺动脉、压力、血液动力学过程”,又因为该文献是论述生理学的有关文献内容,按组配标引规则进行概念限定——“门静脉、肺循环/生理学、肺动脉、压力、血液动力学过程”。如此,借助任一个检索词都能检索到此篇文献。

3 组配标引的2种形式及规则

3.1 主题词和主题词的概念组配

1)用2个或2个以上的主题词作概念组配。例如:《再生障碍性贫血自然杀伤细胞的量及活性研究》,标引为“再生障碍性贫血;自然杀伤细胞;免疫活性”。

2)主题词和特征词的概念组配。例如:《慢性常压缺氧高二氧化碳大鼠及肺的超微结构变化》,标引为“低氧血症;血酸过多;心肌/超微结构;肺/超微结构;大鼠”。

主题词和主题词的组配应该采用概念组配,MeSH对一些复合概念采用直接选用专指性的词组作叙词(即概念预先组配)。选用某些专指性的选词可使主题词标引定型,避免错误组配带来的混乱,减少误检。例如:“葡萄球菌性肺炎”标引为“肺炎;葡萄球菌性感染”是错误的,因为这2个词是上位词,应标引为“肺炎;葡萄球菌”(是专指主题词)。

3.2 不能越级组配 越级是指该叙词在范畴表中的类级位置,在组配时不能标出它的类级概念。如《表皮生长因子治疗慢性萎缩性胃炎》,标引为“胃炎,萎缩性/药物治疗;表皮生长因子/治疗作用”,不标为“胃炎,萎缩性/治疗”,治疗只是药物治疗的上位词。

1)要根据词表规定的范围和组配原则进行组配。例如:尿的成分,不标成“尿/异常”,因为尿是在范畴表中属A12类体液和分泌物,不能和“异常”组配,“/异常”仅限于器官部位的形态异常。

2)在MeSH中许多为经过规范的专指主题词,规定要首选这些词而不用其他的词组配标引。例如:“血液中的葡萄糖”标引为“血糖”(先组),不标为“葡

萄糖/血液”。

3)如能用组配表示的1种概念不要标引为2个或2个以上的主题。例如:“胃的消化生理”标引为“胃/生理学”,而不标为“胃;生理学;消化”。

4)首先进组的主题词。如果组配的结果符合某主题概念时,要标先组的主题词,而不用组配。如一篇论及事故预防的文章,标“事故/预防和控制”,这与先组的主题词“事故预防”概念完全一样。

4 组配标引的必要补充

在具体的标引实践中,当运用组配标引表达主题概念时,往往会遇到有的文献不能用词表中现成的叙词进行组配表达其主题面这种情况,作为对组配标引的必要补充,就要考虑用自由词以及其他方法。

4.1 原则 用自由词来补充表达组配标引无法涉及的主题概念时,首先应遵循以下原则:

- 1)该词在本学科领域出现频率高;
- 2)必须采用规范、标准的成熟名词;
- 3)仅限于表达主题概念的主要内容时使用;
- 4)不宜用太细、太专的词做自由词,宜用常见词,如在眼科领域,“角膜检查镜”可作自由词标引。

4.2 药物或化学物质

1)归入最密切相关的多数类。如《邻溴苯乙酸的合成新途径》中邻溴苯乙酸标准溴化苯类。

2)可用最相关的叙词加类似物和衍生物。如《新试剂反相高效液相色谱分离测定氨基酸》,以N-羧基琥珀酰亚胺-3-吡啶乙酸酯为试剂,MeSH中没有该叙词,就以“吡啶乙酸类/类似物和衍生物”标引。

3)或按药物的药理作用标引。如“炎痛喜康的抗炎镇痛作用”标为“镇痛药;抗炎”。

4.3 概念概括 概念概括是指几个并列的同级词,可以用一个上位词代替。组配要遵循所谓的“三”原则,即凡3个以上的同类概念,应用其共有的上位词概括标引^[2]。

5 提高医学文献组配标引质量的对策

1)加深认识与了解《医学主题词表》(MeSH表)的树状结构(tree structure)表,了解主题词之间的隶属关系,掌握主题词的使用方法及相关注释,对副主题词的含义范畴进行辨识,做到准确、恰当地使用副主题词,不错标、不漏标。同时要掌握主题词标引的基本原则,即专指性、一致性、全面性、准确性、客观性和符合性^[6]。还要注意搜索引擎的运用^[7]。

2)借鉴国内外数据库以提高标引质量。医学期刊编辑在进行标引的过程中,还可以浏览、借鉴国内外