

· 专题交流 ·

确定职业医学国外核心期刊方法探讨

上海市劳动卫生职业病防治研究所 (200003)

杨明山 马美琳 陆志英 郑卫东 韦 星

确定核心期刊是一项极其重要的信息研究工作。1948年英国 Bradford 全面提出了文献分散的布拉德福(下简称布氏)定律。国内这类研究约始于80年代,并于1991年出了一本专著,提出99种专业的国外核心期刊,然无职业医学专科分类。本文试以光盘与国内外专业期刊为引用源,比较并运用五种方法,通过计算机技术处理,确定职业医学国外核心期刊。

1 材料与方

1.1 引文法 本文采用一步引文法

1.1.1 材料来源 国内期刊12种,包括国内本专业的所有一级与多数二级期刊。国外期刊10种,涉及11个国家。所选中外文期刊均为1993年与1994年出版物。引用要求:只分析与职业医学有关的一次文献。

1.1.2 数据获取方法 中文期刊源:以人工方法计算每篇被引文献后的期刊类参考文献,按刊名类加。此数据输入微机,以每种期刊被引量除以总引文量得引文率,按引用频次排序。西文期刊源:方法同中文期刊。两种引文率相加排序,谓总引文法。引文法得序表三张。

1.2 文摘法

1.2.1 材料来源:光盘(下略) NIOSHTIC,为美国职业安全卫生研究所文献数据库;HSELINE为英国职业安全卫生数据库;MEDLINE,内含“医学索引”等文摘;TOXLINEPLUS,含“化学文献”(毒理学)等文摘(库);CISDOC,为国际劳工局的国际职业安全卫生数据库。本研究收集五库中十年(1985~1994)内世界范围内的职业医学文献。

1.2.2 数据处理方法

1.2.2.1 根据 Silver Platter 检索系统,从五库中分别获取每库每刊的文献数。五库数据分别输入80386DX/40兼容机中,用FOXPLUS 2.1软件处理,按刊名累加,得五份原始数据单。

1.2.2.2 把上述五份数据单分别输入IBM486兼容机里,用Foxpro2.5软件处理,计算出每库每刊文摘率,按频次排序,分得序表五张。最终把五表序号总加

并排序,得五库文摘率总和序表一张。文摘法共得序表六张。

1.3 综合法 采用加权平均处理的综合筛选法,内有四个变量:引文率(a)(西文期刊源),引文率(b)(中文期刊源),文摘率(c)与专家推荐率(d)。

1.3.1 数据获取:前三种数据(a、b、c)上已算出。专家推荐率分析,采用问卷调查表法。对象为正教授级职业医学专家。函寄调查表65份,反馈55份,反馈率为84.6%。该数据输入微机,按推荐频次排序。

1.3.2 数据处理 a、b、c、d的权重系数分别为0.3,0.3,0.2与0.2。设每刊四种变量的加权平均值之和为V(A)。再按公式 $V(A) = 0.3a + 0.3b + 0.2c + 0.2d$ 进行计算,最后按V(A)值由大至小排序。综合法得序表一张。

1.4 布拉德福文献分散曲线图像法

1.4.1 以文摘法中五库期刊被摘频次数据为基础,按文献数相同的期刊为同一级,列出各级期刊的顺序累加数据。

1.4.2 制图 设各级期刊数n为X轴,取值 $\log(n)$;各级期刊数的文献收录数R(n)为Y轴,取值 $R(n) \cdot 10^{-3}$ 。把每对数据输入Slidewrite处理文件中,以散点曲线法制图。

1.4.3 观察 典型的布氏分散曲线呈三段式:上升部AC,直线CB,下垂部BD。根据AC与CB的结合C点以确定核心期刊区。五库得布拉德福文献分散曲线图像五张。

1.5 布拉德福定律数学推导式 作者以回归分析法研究五种文摘(库)的文献分布。

1.5.1 英国学者Brooks对布氏定律提出数学推导式,即:期刊核心区相当于 $\ln R(n) = a \cdot n^b$, $1 \leq n < C$ (1),期刊非核心区相当于 $R(n) = k \cdot \ln(n/s)$, $C \leq n \leq N$ (2)。美国学者Worthen,将方程(1)和方程(2)线性化,得 $\ln R(n) = \beta \ln n + \ln a$ (3)与 $R(n) = k \ln n - k \ln s$ (4)。经推导方程(3)与(4)可化成同一形式: $y = bx + a$ 。作线性回归拟合,得拟合模式:

$$R(n) = \begin{cases} \alpha \cdot n^3, & 1 \leq n < C & (5) \\ k \cdot \ln(n/s), & C \leq n \leq N & (6) \end{cases}$$

1. 5. 2 分析 作一元性回归分析, 并根据最小两乘法原理得到核心期刊 C 值。

1. 5. 3 编程 在 IBM-PC 兼容机上, 用 Turbo C 语言编制上述模式的计算机程序, 求 C 值。

2 结果与讨论

2. 1 引文法

用计算机分别计算出被引期刊序表 A (中文期刊源)、B (西文期刊源)、与 C (总引文法) (均略)。其文献总数 (期刊总数) $[R(n), \text{下同}]$ 分别为 3 028 (82)、10 516 (163) 与 13 544 (180)。核心期刊区间一般按累计引文量达 70% 计算, 那么引文法的核心期刊数分别为 82、135 与 107 种 (工作量: 4 人耗时两个月)。一步引文法的引文源是一次文献, 即作者直接记载其科学研究成果的产物。这种方法实践性较强。

但该法引用范围有限, 如总引文法的被引期刊种类数为 472, 只占文摘法总和的 19.8%。少数期刊的自引率过高, 如 Contact Dermatitis 的自引率达 45%, 在总引文法中位居榜首, 显然不符实际, 所得核心期刊数过大, 不符需要。由此可见, 单根据引文法来确定核心期刊实显不妥。

2. 2 文摘法

用计算机分别计算出五库的被引期刊序表 C 至 G (均略), 统计结果见表 1。(工作量: 2 人耗时两周)。

查出五库十年所需文献总和近十万篇, 被摘期刊达 4 656 种, 分别是引文法的 9.86 倍与 6.7 倍, 可弥补引文法之全面性不足。本研究的文摘法偏袒大期刊的现象并不严重, 如日本的 Industrial Health, 年载文献量约 30 篇, 按年载文献量排序约为几百位, 但在五库总和排序表上名列 34 位。工作量小于引文法, 只需 1 人两周。然而, 文摘法不能充分反映职业医学人员的实践活动。毒理学文献较多是动物实验研究, 所以在文献标引中不属职业医学文献, 故毒理学类期刊在文摘法中排序很低。另外, 按 70% 法计算核心期刊, 其数额亦过大, 少则也有 108 种 (NIOSHTIC)。看来, 单独根据文摘法来确定核心期刊似显不足。

2. 3 综合法

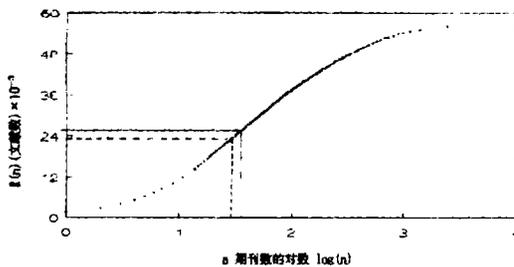
用计算机计算出被摘期刊序表 H。其 $R(N)$ 为 104 304 篇 (4 656 种)。按 70% 引文累计量算出核心期刊为 64 种。工作量同文摘法。

综合法里加入了专家推荐变量, 他们的观点能全面反映职业医学各领域应用国外本专业期刊的最高水

平, 故能纠正上述二法的某些偏倚。在引文法中居首位的 Contact Dermatitis, 现降至第四位; 某些毒理学期刊的序号有所上升, 如在综合法中居 103 位的 Toxicology 上升至 12 位。70% 法的核心期刊区缩小许多 (引文法为 135)。由此可见, 综合法可兼蓄实践性 (引文法) 和全面性 (文摘法) 的优点。

作者认为, 综合法得出的期刊排序表比较全面合理, 可作为本研究最终划定核心期刊区的数据基础。然而, 综合法的核心区数仍显过大, 这可能与职业医学文献离散程度较大有关。为此, 期望运用布氏文献分散曲线与定律予以解决。

2. 4 布拉德福图象法



NIOSHTIC 布拉德福文献分散曲线

用计算机分别计算打印出五个文摘(库)的五张图像。发现 NIOSHTIC 库的图像比较典型, 故本文仅分析 NIOSHTIC 图像 (见图): 曲线斜率较高, 与此库文献总量高 (56 074, 居五库之首) 有关; 曲线三段形态比较明显, 说明职业医学文献分散符合一般呈高中低三个频区的规律; AC 与 CB 的结合点 $[R(n) = 27 000, \log(n) = 1.5]$ 即为核心期刊区的下限。以 $\log(n)$ 查得 $n \approx 40$, 即为 NIOSHTIC 的核心期刊数。其所含引文约为该库文献总数的 40%~50%。

运用布氏曲线与布氏定律要有三个条件: 专业划定清楚, 文献统计充分以及时间界限明确。而本研究的五个库中, MEDLINE 偏重临床学科不全面; TOXLINEPLUS 内含 18 个子库, 收录重复率较高 (约 11%), 数据准确性不强; CISDOC, 收录文献量小 (收文献 17 514 篇, 居五库之末), 收录不充分。从而导致 MEDLINE、TOXLINE 与 CISDOC 三库的图像不典型。

在布拉德福图像法中, C 点很难确定到个数, 误差可达十以上。如图内虚线方格核心区为 32 而实线方格核心区为 42。显然, 布氏图象法难以用作主要的方法。

2. 5 布拉德福定律的数学推导法

经计算,得五库C值等数据(见表1)。取 NIOSHTIC 的C值(42,核心期刊数),算出核心区的引文量为总引文量的48.64%。MEDLINE、TOXLINEPLUS 与 CISDOC 的C值过小,其原因同三库的布氏图象不典型的原因一样:文献收集不全、不足与量

过小。

2. 6 职业医学国外核心期刊的划定

以综合法计算出的期刊序表作为本研究最终划定核心期刊区的数据基础,用 NIOSHTIC 的核心期刊区的百分比48.64%作为划定核心期刊区的分水岭,得职业医学国外核心期刊共25种(见表2)。

表1 五种文摘(库)职业医学文献布拉德福分散现象数学推导式数据表

文献(库)名	期刊总数 <i>n</i>	文献总数 <i>R(n)</i>	平均期刊 被引文献数 <i>R/n</i>	被引文献 频次级数	引文最高 期刊文献数	核心区 期刊数 <i>C</i>	核心区文献与 全部文献比率 <i>R(c)/R(n)</i>
NIOSHTIC	2 379	56 074	23.57	164	1 519	42	48.64
HSELINE	941	18 411	19.57	96	1 251	23	60.72
MEDLINE	2 004	26 321	13.13	103	1 085	25	35.63
TOXLINE PLUS	1 937	28 514	14.72	106	1 360	4	16.31
CISDOC	2 004	17 514	8.74	100	619	5	15.10

表2 国外职业医学核心期刊序表

1. Occup-Environ-Med	(英国,英语)	13. Environ-Res	(美国,英语)
2. J-Occup-Med		14. Environ-Health-Perspect	(美国,英语)
(现名 J-Occup-Environ-Med)	(美国,英语)	15. Am-Rev-Respir-Dis	(美国,英语)
3. Am-Ind-Hyg-Assoc-J	(美国,英语)	16. Appl-Occup-Environ-Hyg	(美国,英语)
4. Contact-Dermatitis	(丹麦,英语)	17. Occup-Med	(英国,英语)
5. Am-J-Ind-Med	(美国,英语)	18. Lancet	(英国,英语)
6. Mutat-Res	(荷兰,英语)	19. Cancer-Res	(美国,英语)
7. Scand-J-Work-Environ-Health	(芬兰,英语)	20. Ann-Occup-Hyg	(美国,英语)
8. Toxicol-Appl-Pharmacol	(美国,英语)	21. J-Toxicol-Environ-Health	(美国,英语)
9. Int-Arch-Occup-Environ-Health	(德国,英语)	22. Am-J-Enidemiol	(美国,英语)
10. Gig-Tr-Prof-Zabol	(俄罗斯,俄语)	23. Arch-Toxicol	(德国,德语)
11. Arch-Environ-Health	(美国,英语)	24. Fund-Appl-Toxicol	(美国,英语)
12. Toxicology	(爱尔兰,英语)	25. Med-Lav	(意大利,意大利语)

3 结论

关于引文法与文摘法的利弊得失及其互补作用,专家已有定论。查阅国内近十几年文献发现,引文法几独统天下。究其根源,约是文摘(库)的文献量浩瀚,因此依靠人工统计是勉为其难的。但是,在计算机开始广泛应用的今天,仍然单独以引文法确定核心期刊者似显不妥。本研究表明,依靠较少的人力与时间,利用光盘信息源与计算机处理,完全可使文摘法简单化,并

最终以较全面合理的综合法确定核心期刊。

无论用何种方法排定期刊序表,最后均要以某百分比来划定核心区。本研究认为,核心百分比的大小与某学科文献的分散规律有内在关系,不可机械套用某百分比,专科文献尤其如此。作者认为,完全可以运用数学推导式利用计算机技术计算出,可避免人为因素。

(参考文献 略)

(收稿:1996-04-04)