

慢性病观察性研究论文统计学应用评价

洪志恒 王萃 曹卫华

【摘要】 目的 评价中华医学会影响因子在0.5以上的核心期刊中慢性病观察性研究论文的统计学应用情况。方法 选择近5年影响因子在0.5以上的中华医学会系列杂志,采用自行设计的调查表,对是否有研究设计类型、样本量估计、变量及赋值描述,以及统计描述、推断及解释等进行调查,分析慢性病观察性研究论文统计学应用情况。结果 共有《中华流行病学杂志》、《中华预防医学杂志》、《中华心血管病杂志》、《中华内分泌代谢杂志》、《中华内科杂志》和《中华肿瘤杂志》6种期刊中的352篇慢性病观察性论著入选。明确说明研究目的、目标对象、抽样对象、对象入选标准和变量定义率的分别为99.43%、98.57%、95.43%、92.86%和96.87%;计量资料 and 分类变量描述正确率分别为90.94%、91.46%;计量资料、分类变量及模型推断正确表达率分别为100%、95.32%和87.19%;研究结论正确回应研究目的的占89.49%。没有明示研究设计类型的占69.60%,需要进一步说明入选对象排除标准而没有说明的占11.14%,仅有5.16%论著提及了样本量估计情况,变量赋值说明率为24.21%。介绍了数据库建立方法的为24.15%,18.75%论著统计推断方法介绍不全。进行率标准化的论著约有1/4使用不当,对统计检验前提条件进行验证和说明的仅占24.12%,有9.94%论著需要做推断分析而未做。结论 目前中华医学会系列杂志发表的慢性病观察性研究论文统计学应用总体良好,但设计类型表述、样本量估计、变量赋值说明不充分;统计方法介绍不够明确;应加强了解统计推断使用的前提条件。

【关键词】 观察性研究; 统计学; 应用评价

Application of statistics on chronic-diseases-relating observational research papers HONG Zhi-heng, WANG Ping, CAO Wei-hua. Key Laboratory of Epidemiology, Ministry of Education; Department of Epidemiology and Biostatistics, School of Public Health, Peking University, Beijing 100191, China

Corresponding author: CAO Wei-hua. Email: caowehua60@163.com

【Abstract】 Objective To study the application of statistics on Chronic-diseases-relating observational research papers which were recently published in the Chinese Medical Association Magazines, with influential index above 0.5. **Methods** Using a self-developed criterion, two investigators individually participated in assessing the application of statistics on Chinese Medical Association Magazines, with influential index above 0.5. Different opinions reached an agreement through discussion. **Results** A total number of 352 papers from 6 magazines, including the *Chinese Journal of Epidemiology*, *Chinese Journal of Oncology*, *Chinese Journal of Preventive Medicine*, *Chinese Journal of Cardiology*, *Chinese Journal of Internal Medicine* and *Chinese Journal of Endocrinology and Metabolism*, were reviewed. The rate of clear statement on the following contents as: research objectives, target audience, sample issues, objective inclusion criteria and variable definitions were 99.43%, 98.57%, 95.43%, 92.86% and 96.87%. The correct rates of description on quantitative and qualitative data were 90.94% and 91.46%, respectively. The rates on correctly expressing the results, on statistical inference methods related to quantitative, qualitative data and modeling were 100%, 95.32% and 87.19%, respectively. 89.49% of the conclusions could directly response to the research objectives. However, 69.60% of the papers did not mention the exact names of the study design, statistically, that the papers were using. 11.14% of the papers were in lack of further statement on the exclusion criteria. Percentage of the papers that could clearly explain the sample size estimation only taking up as 5.16%. Only 24.21% of the papers clearly described the variable value assignment. Regarding the introduction on statistical conduction and on database methods, the rate was only 24.15%. 18.75% of the papers did not express the statistical inference methods sufficiently. A quarter

of the papers did not use 'standardization' appropriately. As for the aspect of statistical inference, the rate of description on statistical testing prerequisite was only 24.12% while 9.94% papers did not even employ the statistical inferential method that should be used. **Conclusion** The main deficiencies on the application of Statistics used in papers related to Chronic-diseases-related observational research were as follows; lack of sample-size determination, variable value assignment description not sufficient, methods on statistics were not introduced clearly or properly, lack of consideration for pre-requisition regarding the use of statistical inferences.

[Key words] Observational study; Statistics; Application evaluation

采用观察性方法研究慢性病的中文论文占70%以上^[1],是许多临床和公共卫生的知识来源。观察性研究中,对象的特征是客观存在的,非观察因素较难控制^[2],需要充分应用统计学的思维实现正确的科研设计和数据的收集、整理、分析及解读,以较少的人力、物力、财力和时间获得准确可靠的结果^[3]。

统计方法有其适用条件,不少研究存在误用统计学的情况,已有文章对医学科研论文的统计应用规范性进行探讨,但大部分仅罗列统计错误^[4-7],也有研究^[8-11]采用专门的调查表,但是未说明评价标准,以致前后研究结果不便比较,且大多针对的是随机对照试验研究^[12-19]。本研究采用横断面研究设计,对中华医学会系列杂志发表的慢性病观察性研究论文的统计学应用情况进行分析。

材料与与方法

1. 调查对象:中华医学会系列核心期刊上发表的慢性病观察性研究论著。

(1)期刊入选标准:①被2011年版北京大学中文核心期刊收录;②中华医学会系列杂志;③综合类、预防医学类或涉及所研究慢性病的内科和专科类杂志;④根据2010年最新数据,影响因子在0.5以上。同时满足上述条件者纳入本研究,其中《中华流行病学杂志》、《中华预防医学杂志》、《中华心血管病杂志》、《中华内分泌代谢杂志》、《中华内科杂志》和《中华肿瘤杂志》6种期刊入选。

(2)论文纳入和排除标准:①发表于2007年1月至2011年12月;②包含中英文摘要、关键词以及字数在3500字以上的论著;③研究目的与慢性病相关,病种包括恶性肿瘤、冠心病、脑卒中、高血压、糖尿病和COPD等;④研究设计类型为观察性研究;⑤排除仅涉及慢性病危险因素(如吸烟、体力活动、超重与肥胖等)和慢性病相关体格指标(如血糖、血脂、血压)的研究,共纳入352篇符合条件的论著。

2. 研究方法:

(1)统计学应用评价指标体系:在提取STROBE^[20](strengthening the reporting of observational studies

in epidemiology)声明的观察性研究论文报告规范中有关统计学应用的指标基础上,结合专家访谈等方式,制定本研究统计学应用评价指标。从研究设计、统计方法介绍、统计描述、统计推断和统计结果及其解释5个方面评价入选论文的统计学应用情况。

(2)调查实施:根据统计学评价指标制定调查表,在完成预调查并对评价指标进行修改的基础上,对所有符合要求的论著,由两名调查员分别填写调查表,意见不同部分进行讨论并取得一致。

3. 统计学分析:采用EpiData 3.1软件建立数据库,实行双录入质量控制。使用SAS 9.1.3统计软件进行数据分析。对需要进一步推断的四格表资料、趋势资料,分别采用 χ^2 检验、Cochran-Armitage趋势Z检验, $P < 0.05$ 为有统计学意义。

结果

1. 论著基本特征:最终纳入352篇慢性病观察性论著。临床类杂志228篇,占64.77%,其中《中华心血管病杂志》67篇(19.03%)、《中华内分泌代谢杂志》57篇(16.19%)、《中华内科杂志》41篇(11.65%)、《中华肿瘤杂志》63篇(17.90%)。预防类杂志124篇,占35.23%,其中《中华流行病学杂志》79篇(22.44%)、《中华预防医学杂志》45篇(12.78%)。从经费来源看,基金项目203篇,占57.67%,非基金项目149篇,占42.33%;杂志、第一作者单位属性、发表年份等见表1。

2. 统计研究设计:

(1)研究设计类型描述:入选的352篇论著中,有350篇(99.43%)明确说明了研究目的。文中明示研究设计类型的占30.40%,其余需由调查员阅读全文后做出判断,预防类杂志(49.19%)高于临床类(20.18%),差异有统计学意义($\chi^2 = 31.88, P < 0.0001$);2007—2011年5年间,文中明示研究设计类型分布呈逐年增加趋势(分别为18.52%、21.05%、33.33%、38.55%和33.75%),经趋势检验有统计学意义($Z = 2.57, P = 0.0103$)。所选择的研究设计符合研究目的占98.58%。各杂志研究类型说明方式、研

究设计类型分布见表 2。

表 1 352 篇观察性论著基本特征

基本特征	杂志类型		合计
	临床类杂志	预防类杂志	
发表年份			
2007 年	36(15.79)	18(14.52)	54(15.34)
2008 年	43(18.86)	14(11.29)	57(16.19)
2009 年	48(21.05)	30(24.19)	78(22.16)
2010 年	45(19.74)	38(30.65)	83(23.58)
2011 年	56(24.56)	24(19.35)	80(22.73)
论文经费来源			
基金	117(51.32)	86(69.35)	203(57.67)
非基金	111(48.68)	38(30.65)	149(42.33)
第一作者单位			
高等院校	13(5.70)	42(33.87)	55(15.63)
科研单位	42(18.42)	46(37.10)	88(25.00)
疾控机构	2(0.88)	13(10.48)	15(4.26)
医疗单位	171(75.00)	23(18.55)	194(55.11)
单位行政级别			
国家级	66(28.95)	32(25.81)	98(27.84)
省级	146(64.04)	85(68.55)	231(65.63)
地市级	16(7.02)	7(5.65)	23(6.53)
合计	228(64.77)	124(35.23)	352(100.00)

注:括号外数据为篇数,括号内数据为构成比(%)

表 2 352 篇观察性论著研究设计类型及说明方式

分类	杂志类型		合计
	临床类杂志	预防类杂志	
研究设计类型			
横断面研究	98(42.98)	65(52.42)	163(46.31)
历史资料分析	11(4.82)	12(9.68)	23(6.53)
病例对照研究	61(26.75)	28(22.58)	89(25.28)
队列研究	28(12.28)	14(11.29)	42(11.93)
病例系列研究	30(13.16)	5(4.03)	35(9.94)
说明方式			
文中明示	46(20.18)	61(49.19)	107(30.40)
调查员判断	182(79.82)	65(52.42)	245(69.60)

(2)研究对象描述:350 篇明确说明研究目的的论著中,有 331 篇(94.57%)选择的研究对象符合研究目的,明确说明目标总体的占 98.57%,抽样总体的占 95.43%,描述对象入选标准占 92.86%。上述 4 项指标在 6 种杂志中情况相近,但描述对象入选标准以 2007 年发表的最低为 86.79%。需要进一步说明入选对象排除标准而没有说明的占 11.14%,以肿瘤杂志(31.75%)和 2007 年发表(22.64%)的论著最低。

(3)样本量估计及抽样方法描述:需要进行样本量估计的 310 篇论著中,仅有 5.16%详述了样本量估计过程,且预防类杂志(7.07%)高于临床类(4.27%),差异无统计学意义($\chi^2=0.17, P=0.2791$);所有论著中说明抽样方法的占 92.33%,临床类杂志

(93.86%)略高于预防类(89.52%),差异无统计学意义($\chi^2=2.13, P=0.1441$)。在说明了抽样方法的 325 篇论著中,随机抽样占 25.23%,预防类杂志(45.05%)明显高于临床类(14.95%),差异有统计学意义($\chi^2=42.85, P<0.0001$)。

(4)研究变量定义与赋值描述:352 篇论著中,需要定义结局变量、暴露变量、分层变量等三种类型变量共计 766 项(次),明确变量定义率的有 96.87%,该指标各类杂志、发表年份一致;变量赋值说明率仅为 24.21%,临床类杂志(3.48%)低于预防类(12.80%),差异无统计学意义($\chi^2=2.34, P=0.1260$)。2007—2011 年赋值说明率分别为 0.00%、4.41%、6.32%、11.63%和 11.63%,趋势检验有统计学意义($Z=9.43, P=0.0019$)。

3. 统计方法介绍:数据库建立方法介绍率为 24.15%、预防类杂志(41.94%)高于临床类(14.47%),差异有统计学意义($\chi^2=32.98, P<0.0001$),见表 3。所用统计分析软件介绍率为 75.85%,临床类杂志(94.30%)高于预防类(81.45%)。统计推断方法介绍完整率 78.69%,预防类、临床类杂志分别为 82.26%、76.75%;统计推断方法不全介绍率 18.75%。统计意义判断标准描述率 55.68%,并呈逐年增加趋势,趋势检验有统计学意义($Z=4.67, P<0.001$)。

表 3 不同类型杂志发表论著统计学应用情况

评价条目	杂志类型		合计
	临床类杂志	预防类杂志	
统计表($n=352$)			
标题恰当	220(98.21)	124(100.00)	344(98.85)
纵横标目合理	223(99.55)	123(99.19)	346(99.43)
标目明确	213(95.09)	121(97.58)	334(95.98)
符合规范	219(97.77)	121(97.58)	340(97.70)
精度统一	218(97.32)	121(99.19)	339(97.41)
连续资料描述情况($n=287$)			
描述集中离散趋势	194(92.82)	71(91.03)	265(92.33)
正确描述	190(90.91)	71(91.03)	261(90.94)
报告可信区间	9(4.31)	1(1.28)	10(3.48)
类型资料描述情况($n=328$)			
采用相对数指标	197(96.10)	120(97.56)	317(96.65)
正确描述	192(93.66)	108(87.80)	300(91.46)
报告可信区间	5(2.44)	7(5.69)	12(3.66)
统计推断($n=718$)			
检验条件验证说明	74(14.86)	33(15.00)	107(14.90)
选择恰当方法	443(88.96)	194(88.18)	637(88.72)
正确使用的方法	429(86.14)	189(85.91)	618(86.07)
统计结果解释($n=352$)			
充分利用信息	209(89.04)	120(93.55)	329(90.63)
讨论中使用统计结果	219(91.67)	123(96.77)	342(93.47)
统计学显著性意义解释恰当	88(38.60)	87(70.16)	175(53.85)

4. 统计描述:

(1)统计表和统计图:348篇论著使用了统计表,标题恰当率98.85%,纵横标目合理率99.43%,标目明确率95.98%,符号规范97.70%,精度统一率97.41%,见表3。2007—2011年统计表使用合格率分别为88.89%、90.91%、93.59%、91.53和91.38%,趋势检验无统计学意义($Z=0.36, P=0.7213$)。90篇论著使用了统计图,肿瘤杂志统计图使用率较高,为50.79%。统计图坐标轴刻度正确率98.89%,选用的统计图类型恰当率为92.22%。

(2)数据特征描述:287篇(92.33%)论著描述了数据的集中和离散趋势,恰当选择计量资料描述指标并正确使用的占90.94%;328篇论著有96.65%选择相对数描述分类变量特征,正确使用的占91.46%。50篇使用率标化的论著,正确使用的占74.00%。

5. 统计推断:在入选352篇论著中,进行推断统计的论著有258篇,占73.30%。在入选352篇论著中,使用统计推断方法共计718篇次。其中计量资料的统计推断(内容包括 t 检验、方差分析、非参数方法)212篇次,分类资料的统计推断(列联表 χ^2 及其校正及衍生检验)264篇次,模型或方程推断统计242次。

(1)计量及分类资料的统计推断:文中对统计检验的前提条件进行验证及说明率仅为24.12%,计量资料推断前提条件验证及说明率(35.85%)高于分类资料(7.95%),但分类资料推断前提条件验证及说明率呈逐年提高。

(2)模型或方程推断统计:有242篇次进行了模型或方程推断统计,方法依序为多重logistic回归分析135篇次(55.78%)、Cox回归分析63篇次(26.03%)、简单线性回归分析32篇次(13.22%)、多重线性相关分析24篇次(9.91%),其他分析方法6篇次(2.48%)。统计检验的前提条件进行验证及说明占4.13%。正确使用统计分析方法92.56%,正确表达统计分析结果87.19%。

此外,需要做推断分析而未进行推断统计的论著有35篇,占9.94%,临床类杂志(11.40%)略高于预防类杂志(7.26%)。

6. 统计结果解释:352篇论著中正确利用所搜集的资料信息解释统计结果占93.47%;讨论中正确应用了统计结果的占97.16%,统计显著性意义解释恰当占53.85%。这3项指标的正确率,预防类杂志均高于临床类杂志。

讨 论

本研究在科学评估指标体系的基础上,选取中华医学系系列杂志影响因子在0.5以上的6种核心期刊发表的慢性病观察性研究论著为调查对象,以了解其统计学应用情况。

在研究设计方面,多数论著能够说明研究目的,采用适当的研究方法和描述所研究的对象,但有69.60%论著在文中没有明示研究设计类型,需要进一步说明入选对象排除标准的论著而没有说明占11.14%。仅有5.16%论著提及了样本量估计情况。明确对结局、暴露和分层(特征)变量定义率为96.87%,但赋值说明率仅为24.21%,在多变量统计推断中,变量的赋值与结果解释息息相关,相同的结果参数在不同的变量赋值方法下可能意味着不同或是完全相反的研究结论。

在统计学方法介绍方面,数据库建立方法介绍率低(24.15%),临床类论著仅为14.47%;有1/4的论著统计学方法说明缺乏完整性和规范性,18.75%论著统计推断方法介绍不全。针对观察性研究论文,在研究设计已经明确的前提下,在统计学方法介绍应描述根据不同的研究目的、数据类型以及分组方式选取的有针对性的统计描述及分析方法,不应仅用“使用 t 检验、 χ^2 分析”概括之。

在统计描述方面,统计描述包括列表描述、图示描述和数字描述;在统计表中,本研究有关标题、标目、符号、精度等单项评价指标评价情况较理想(97.41%~99.43%)。在统计图应用的文献较少(25.57%),恰当选择计量资料描述指标并使用正确的论著占90.94%。正确使用相对数指标对分类资料的论著进行描述占91.46%,但进行率标准化的论著约占1/4标准化率使用不恰当。

在统计推断方面,普遍缺乏对统计检验的前提条件进行验证及说明(24.12%),有少数论著(9.94%)需要做推断分析而没有进行推断统计。统计结果的合理解释与研究结论正确回应了研究目的等指标评价良好。

本研究表明中华医学系系列杂志影响因子在0.5以上的核心期刊上发表的慢性病观察性研究论著大都满足统计学的要求,主要表现在选择的研究类型、研究对象选择恰当,统计图表使用和资料特征描述较规范,这可能与近年来论文统计学应用更受关注和一些杂志在稿约中统计学方法规定与要求明确有关。存在的主要问题集中在缺少样本量估计,

可信区间报告缺失,较少提供变量赋值情况和统计推断使用的前提条件以及统计学显著性意义解读不全等。本研究入选杂志代表中国医学期刊的较高水平,结果的外推性不佳,并且已经刊出的论著经过审阅和修改,统计学使用情况优于一般医学研究。

(本研究得到李天霖教授的精彩指导)

参 考 文 献

- [1] Liu M, Lv J, Li LM, et al. A retrospective analysis of research articles on the major chronic diseases in China, from 1990 to 2009. *Chin J Epidemiol*, 2011, 32(7): 724-726. (in Chinese)
- [2] 刘森,吕筠,李立明,等. 1990-2009年中国主要慢性病研究文献回顾性分析. *中华流行病学杂志*, 2011, 32(7): 724-726.
- [3] Zhan SY. How to report observational epidemiology research-explain of STROBE. *Chin J Evid Based Pediatr*, 2010, 5(3): 223-227. (in Chinese)
- [4] 詹思延. 第三讲: 如何报告观察性流行病学研究——国际报告规范 STROBE 解读. *中国循证儿科杂志*, 2010, 5(3): 223-227.
- [5] Hu LP. Application of the three-model theory in Statistical expression and description. Beijing: People's Military Medical Press, 2008: 27. (in Chinese)
- [6] 胡良平. 统计学三型理论在统计表达与描述中的应用. 北京: 人民军医出版社, 2008: 27.
- [7] Li ZM. Mistakes of Statistical application in medical research articles. *Chin J Postgrad Med*, 2005, 28(3B): 61-62. (in Chinese)
- [8] 李点梅. 医学论文统计学应用中常见问题分析. *医师进修杂志*, 2005, 28(3B): 61-62.
- [9] Cao XT. How to correctly apply statistical methods in medical science articles. *Chin J Hosp Admin*, 2002, 18(10): 639-640. (in Chinese)
- [10] 曹秀堂. 如何在医学科技文献中正确使用统计学方法. *中国医院管理杂志*, 2002, 18(10): 639-640.
- [11] Gou L, Yu J, Yang ZM. Appropriate application of statistical analysis method in medical science and technology articles. *Chin J Rep Rec Surg*, 2005, 19(11): 925-927. (in Chinese)
- [12] 苟莉,俞军,杨志明. 正确应用医学科技论文统计学方法. *中国修复重建外科杂志*, 2005, 19(11): 925-927.
- [13] Wang W. Analysis of statistical mistakes in medical original papers. *Chin J Neurol Neurosurg*, 2004, 4(5): 335-336. (in Chinese)
- [14] 王伟. 医学科研论文中常见的统计学应用错误分析. *中国现代神经疾病杂志*, 2004, 4(5): 335-336.
- [15] Lai J, Li Z, He JP, et al. Comparative study of application of statistical methods in articles of clinical medicine. *Chin J Otorhinolaryngol Skull Base Surgery*, 2010, 16(2): 124-127. (in Chinese)
- [16] 赖娟,李征,何集篇,等. 临床医学杂志论著中统计学方法应用的对比研究. *中国耳鼻咽喉颅底外科杂志*, 2010, 16(2): 124-127.
- [17] Zhang GY. Application analysis of statistics in thesis of medical master graduates. *J Fourth Mil Med Univ*, 2002, 23(16): 1530-1532. (in Chinese)
- [18] 张功员. 医学硕士学位论文中统计学应用情况分析. *第四军医大学学报*, 2002, 23(16): 1530-1532.
- [19] Xing XH, Geng JH, He MR. Analysis of the status quo of statistical methods employed in MS and PhD graduation theses of medical students. *Northwest Medical Education*, 2009, 17(3): 486-488. (in Chinese)
- [20] 邢晓晖,耿景海,何美蓉. 医学研究生学位论文统计学应用状况的调查分析. *西北医学教育*, 2009, 17(3): 486-488.
- [21] Shen J, Gou L, Tang J, et al. Analysis of common on statistical errors in biomedical journals. *Chin J Rep Rec Surg*, 2007, 21(5): 541-543. (in Chinese)
- [22] 沈进,苟莉,汤洁,等. 生物医学期刊的统计学方法应用情况分析. *中国修复重建外科杂志*, 2007, 21(5): 541-543.
- [23] Rcgg C, Cho D, Eastwood S, et al. Improving the quality of reporting of randomized controlled trials. The CONSORT statement. *JAMA*, 1996, 276(8): 637-639.
- [24] Moher D, Schulz KF, Altman DG. The CONSORT statement: revised recommendations for improving the quality of reports of parallel group randomized trials. *BMC Med Res Methodol*, 2001; 1-2.
- [25] DerSimonian R, Charette LJ, McPeck B, et al. Reporting on methods in clinical trials. *N Engl J Med*, 1982, 306(22): 1332-1337.
- [26] Pocock SJ, Hughes MD, Lee RJ. Statistical problems in the reporting of clinical trials—a survey of three medical journals. *N Engl J Med*, 1987, 317(7): 426-432.
- [27] Moher D, Dulberg CS, Wells GA. Statistical power, sample size, and their reporting in randomized controlled trials. *JAMA*, 1994, 272(2): 122-124.
- [28] Ilollis S, Campbell F. What is meant by intention to treat analysis? Survey of published randomised controlled trials. *BMJ*, 1999, 319(7211): 670-674.
- [29] Schulz KF, Chalmers L, Grimes DA, et al. Assessing the quality of randomization from reports of controlled trials published in obstetrics and gynecology journals. *JAMA*, 1994, 272(2): 125-128.
- [30] Ilotopf M, Lewis G, Normand C. Putting trials on trial—the costs and consequences of small trials in depression: a systematic review of methodology. *J Epidemiol Community Health*, 1997, 51(4): 354-358.
- [31] Wang B, Zhan SY. How to report observational epidemiology research-explain of STROBE. *Chin J Epidemiol*, 2006, 27(6): 547-549. (in Chinese)
- [32] 王波,詹思延. 如何撰写高质量的流行病学研究论文 第一讲 观察性流行病学研究报告规范——STROBE 介绍. *中华流行病学杂志*, 2006, 27(6): 547-549.

(收稿日期: 2012-04-17)

(本文编辑: 卢亮平)