

· 专家述评 ·

关注眼病人群流行病学研究模式的转变

邹海东

200080 上海交通大学附属第一人民医院 上海市眼病防治中心 上海市眼科医院 上海市眼底病重点实验室

通信作者:邹海东,Email:zouhaidong@263.net

DOI:10.3760/cma.j.issn.2095-0160.2017.12.001

【摘要】 随着慢性病逐渐成为我国人群健康的首要威胁以及大众对眼健康质量要求的迅速提升,眼病人群的流行病学研究模式必须尽快转变,即应开展大规模、多中心和前瞻性的人群队列研究,在具有不同性别、年龄、民族和生活习惯等的人群中广为验证;构建类似广义线性、非线性混合效应等新型分析模型,更好地分析环境外因和个体内因的交互作用;开展系统流行病学和空间流行病学研究,去解构眼病病因的多元化,更准确地推导病因,最终为特定患者的个性化精准预测、诊断和治疗提供循证医学的最佳证据。

【关键词】 流行病学; 人群; 队列研究; 统计; 精准医学; 循证医学

Focusing on the methodology shift in epidemiological study of eye diseases Zou Haidong

Shanghai General Hospital, Shanghai JiaoTong University, Shanghai Eye Diseases Prevention & Treatment Center, Shanghai Eye Hospital, Shanghai Key Laboratory of Ocular Fundus Diseases, Shanghai 200080, China

Corresponding author: Zou Haidong, Email: zouhaidong@263.net

【Abstract】 As chronic diseases gradually become the primary threat to the health of Chinese population and the rapid increase in the need of high eye health quality, the population-based epidemiological study of eye disease in China should be conducted in the following mode: large-scale, multicenter, and prospective cohort studies in different gender, age, ethnicity and living habits population, with new analytical models similar to generalized linear and nonlinear mixed effects to better analyze the interaction between environmental causes and individual internal factors. The systematic epidemiology and spatial epidemiology concepts are suggested to be included in eye research, in order to deconstruct the cause of eye disease diversification. All these facilitate precision prediction, diagnosis and treatment of individual patient.

【Key words】 Epidemiology; Population; Cohort studies; Statistics; Precision medicine; Evidence-based medicine

近30年来,从不断增长的相关领域文献中可以看出,我国越来越多的眼病专业工作者已直接组织或参与了眼病人群流行病学研究,使得研究覆盖面日益拓展,研究区域从中心大城市到边远农村,研究对象从医院患者到社区人群,研究病种从常见盲因到罕见眼病,研究方法从现况调查到随访观察等,所取得的研究成果不仅在我国致盲眼病预防和控制方面做出了突出贡献,同时也为全球一系列眼健康问题提供了科学研究方法和参考资料,赢得了我国独特的学术荣誉。人群流行病学研究的学科地位主要源于其严谨的科学设计、缜密的统计学分析和科学的逻辑推理(Koch's Postulation与Hill's Criteria)。近年来,随着慢性病逐

渐成为我国人群健康的首要威胁以及大众对眼健康质量要求的迅速提升,我国眼病人群流行病学研究在设计、分析和推理方面面临着新的重大挑战。面对这些挑战,只有积极应对,尽快转变人群流行病学研究模式,才能捍卫我国在该领域的学科尊严,并最终更好地服务于广大患者。

1 眼病人群流行病学研究的科学设计向多中心前瞻性队列模型转变

回顾近期我国眼病人群流行病学研究文献,绝大多数的研究设计仍采用单中心、小样本抽样的横断面现况调查。对于解放初期我国首位致盲原因的沙眼等

急性传染病或者某些眼科罕见病而言,横断面调查可以提供有价值的“病因”线索,但对慢性眼病“危险因素”的立论体系而言,这类研究设计只能得到某个基本点,无法构建框架结构。以我国目前不断增加的儿童青少年近视流行病学研究文献为例,许多仍是在 1 个行政区域内某些学校学生中开展的现况调查,其研究结果仅仅是“相关”,如近视度数与受教育程度、家庭成员近视史等相关等,但始终无法确认“因果”关系,即教育、遗传中的这些因素是否确实导致了儿童近视的发生和发展。诚然,这类现况调查研究对人力、物力和财力要求低、容易实施,但大量类似研究得到的结果往往只是同样的点,很少能发现新的有价值的信息。因此,有志于开展眼病人群流行病学的学者应考虑将研究设计转变为多中心、前瞻性队列研究,采取研究资料、图像、标本分享和合作的新模式,并整合多方资源开展大数据下复杂数据库的建设。

多中心、前瞻性队列研究要求仔细完整的准备和长期规范的执行:(1)认真研读国内外文献,明确已有队列的优势和不足,提炼出有我国人群特色的研究目标。(2)大胆探索,将研究指标从传统的个体、环境和生活习惯因素,拓展到社会、经济、个体心理等未知可能。(3)周密考虑实施过程中的各种可能,特别是关于多中心合作和共享、研究长期稳定开展方面的可能问题,从人力、物力和财力上充分保证。(4)充分借助计算机技术、网络技术和生物医学工程技术,高效开展研究信息的采集、过滤和处理。(5)制定标准化实施方案,作为各中心交叉和互补的规范。(6)严格执行,项目负责人担起实施的主体责任,各中心通力合作,抓小抓细,严把质量关。(7)实时分析,提炼结果,客观自我评估和修正,并加强知识产权保护意识^[1]。

2 眼病人群流行病学研究的统计分析向广义线性、非线性混合效应模型转变

与单中心横断面抽样调查相比,多中心动态队列研究所展现出的数据主要存在以下 3 个特点:(1)多层次 在大规模人群研究中往往采取整群抽样的方法,以便于现场的组织实施。但是这种抽样方法会引起数据的多层次特征,存在组内聚集性,即一个整群中的个体并不完全独立。例如同一所学校的学生由于学习环境相似及课程安排相近,往往表现出相似的屈光发育状态,与其他研究中心不同学校同年龄学生的差别可能很大。(2)重复观测 队列数据有别于横断面数据的一大特点在于个体前后多次的观测之间并不独立。在社区人群糖尿病眼病研究队列中,居民每年都

会接受一次详尽的眼科检查,其中例如眼轴、屈光度数等的测量,同一个体的多次观察数据之间存在很强的关联。(3)非线性 自变量和因变量之间的变化关系往往并非线性。譬如本项目组前期队列研究就发现儿童青少年的等效球镜屈光度数和近视患病率等随着年龄的增加存在较明显的拐点^[2]。在糖尿病眼病研究队列中,发现糖尿病视网膜病变的进展概率与患者的糖化血红蛋白水平呈非线性关系^[3]。除了上述特点外,眼病研究中观测结局的表现形式也复杂多样,既存在大量连续型数据(视力、等效球镜、眼轴等),也存在许多离散型数据(发病/未发病、治愈/好转/未愈/加剧等),在进行数据统计分析的过程中,需要经常对因变量进行特定的变换,以获得符合模型前提的形式。

当前,基础的统计学方法(主要指单水平的线性模型)在眼病多中心人群流行病学队列研究中的应用已日益捉襟见肘,往往得不到精准的参数估计,甚至可能对结论造成扭曲。因此,应积极探索非线性混合效应模型在眼科研究中的应用价值。混合效应模型是处理具有多层次结构以及纵向数据的强有力的统计学工具,最常用的混合效应模型是线性混合效应模型,可在 SAS 软件中通过 MIXED 过程实现。但线性混合效应模型仍受到因变量必须为连续型变量,以及因变量与自变量需符合线性变化等限制。为突破该限制,可在线性混合效应模型中融入广义线性的方法,即构造连接函数,例如对二分类的因变量进行 Logit 变换等,形成广义线性混合效应模型,以解决因变量为离散型的情况,并可解决一部分非线性问题。在实际分析过程中,构建该模型可以通过 SAS 软件中的 GLIMMIX 过程实现。此外,在混合效应模型的基础上,结合非线性处理方法即可形成非线性混合效应模型。处理非线性的方法常见的有多项式模型和样条等,并可联合贝叶斯方法进行估计^[4-5]。其中多项式模型一般只能给出数据的整体变化趋势,对局部波动拟合不佳;样条则能较好地拟合局部波动,但不同的样条选择对数据的拟合效果会有差别,需具体考察。联合贝叶斯方法是较新的估计非线性关系的方法,目前应用仍较少。这些近现代统计学模型的算法主要依靠计算机编程实现,需要研究人员耗费大量的时间和精力进行钻研,也受限于统计分析用计算机的性能,造成了近现代统计学方法在眼病人群流行病学研究实际应用中难以推广、转化的困境。因此,眼病工作者应积极与统计学专家合作,努力开发操作简便、估计精确的智能化分析程序,这对优化我国眼病人群流行病学研究的统计分析具有重要的科研和应用价值。

3 眼病人群流行病学研究的逻辑推理向多元化、空间化、精准化转变

与急性眼病不同,慢性眼病人群流行病学病因的逻辑推理属于“多病因论”,包括:(1)多种因素均可引起同一种疾病。多个病因共同存在时,每个病因均可以独立引起疾病。常用的队列研究、病例对照研究都是基于这个推理,从判断单个因素与疾病的因果关联开始,再采用多重线性回归分析,确定每个自变量对因变量的回归系数。(2)疾病是多种因素共同作用的结果。多个病因组合起来才构成疾病的充分病因。这类病因推理,除了各个病因的剂量组合外,还可能存排列先后问题,病因之间的交互作用更加复杂。事实上,大部分慢性眼病的病因属于后一类推理,如仅血糖升高,并不足以使患者罹患糖尿病视网膜病变,只有当同时暴露于易感基因、免疫、炎症等其他多种病因,且各种因素达到一定的量级并交互作用时,才必然罹患视网膜病变。这种多元化的病因推理是传统流行病学设计无法胜任的。

传统的眼病流行病学人群病因研究多基于或大或小的行政区域人口数量,仅着眼于人体本身,而对于人群所处的独特地理和生态环境关注极少。但随着环境对人类影响的日益增加,必须考虑地理分布上的自然环境变化对眼病发生和发展的作用,如研究人群中眼表疾病发生的病因时,特定地理尺度内人群暴露的环境(如雾霾、水污染、土壤污染)、气象、温度、所处的生物群体环境(如昆虫)等均会对病因推理造成或多或少的影响。目前逐步受到重视的是采用地理信息系统、全球定位系统、遥感和空间分析等技术的空间流行病学研究方法,为探讨这类新的病因提供了可能。

2016 年 6 月,我国发布了国家重点研发计划“精准医学研究”的项目安排,这标志着国内精准医学研究大潮已起。国家精准医疗战略专家组负责人詹启敏院士将精准医学定义为:应用现代遗传技术、分子影像技术、生物信息技术,结合患者生活环境和临床数据,实现精准的疾病分类及诊断,制定具有个性化的疾病预防和治疗方案^[6]。群体流行病学研究方法是精准医学研究的共性技术,但面对突如其来的高度个体化的病因推理要求,在海量的生物样本数据及基因、蛋白、转录、代谢等各种组学数据面前,用人群流行病学研究方法推理病因正面临着巨大的挑战。在大数据时

代,流行病学研究的伞式设计、回归树分析、分类树分析以及多种网络模型等正被推介,但仍不能很好地解决复杂的病因推理,滞后于“精准预测”的需要。

综上所述,疾病类型的客观改变和人类认识的主观提升已迫使眼病人群流行病学研究模式必须尽快转变,即应开展大型的、多中心、前瞻性人群队列研究,在不同性别、年龄、民族和生活习惯等人群中广为验证;利用计算机技术,构建类似广义线性、非线性混合效应等新型分析模型,去分析环境外因和个体内因的交互作用;吸纳其他学科的新进展,开展系统流行病学和空间流行病学研究,去解构疾病病因的多元化,更准确地推导病因,最终为特定患者的个性化精准预测、诊断和治疗提供循证医学的最佳证据。

志谢 感谢上海市眼病防治中心 上海市眼科医院陆丽娜、林森林医师给予的宝贵修改意见

参考文献

- [1] 王慧,董尔丹,张作文. 对我国开展人群队列研究的思考和建议[J]. 中华预防医学杂志, 2014, 48(3): 164-166. DOI: 10. 3760/ema. j. issn. 0253-9624. 2014. 03. 002.
- [2] Ma Y, Qu X, Zhu X, et al. Age-specific prevalence of visual impairment and refractive error in children aged 3-10 years in Shanghai, China[J]. Invest Ophthalmol Vis Sci, 2016, 57(14): 6188-6196. DOI: 10. 1167/iov. 16-20243.
- [3] Jin P, Peng J, Zou H, et al. A five-year prospective study of diabetic retinopathy progression in chinese type 2 diabetes patients with “well-controlled” blood glucose[J/OL]. PLoS One, 2015, 10(4): e0123449 [2017-03-26]. <http://journals.plos.org/plosone/article?doi=10.1371/journal.pone.0123449>. DOI: 10. 1371/journal.pone.0123449.
- [4] 王明高,孟生旺. 贝叶斯非线性混合效应模型及其应用研究[J]. 统计与信息论坛, 2016, 31(12): 10-16. DOI: 10. 3969/j. issn. 1007-3116. 2016. 12. 002.
Wang MG, Meng SW. Bayesian nonlinear mixed effects models and their applications[J]. Statistics & Inform Forum, 2016, 31(12): 10-16. DOI: 10. 3969/j. issn. 1007-3116. 2016. 12. 002.
- [5] 李莹,王仲君,赵华玲. 基于贝叶斯推断的 HIV 非线性混合效应联合模型研究[J]. 统计研究, 2012, 29(7): 86-90. DOI: 10. 3969/j. issn. 1002-4565. 2012. 07. 015.
Li Y, Wang ZJ, Zhao HL. Bayesian inference for nonlinear mixed-effects joint model in HIV study[J]. Statistical Res, 2012, 29(7): 86-90. DOI: 10. 3969/j. issn. 1002-4565. 2012. 07. 015.
- [6] 詹启敏. 中国精准医学发展的战略需求和重点任务[J]. 中华神经创伤外科电子杂志, 2015, 1(5): 1-3. DOI: 10. 3877/ema. j. issn. 2095-9141. 2015. 05. 001.

(收稿日期:2017-08-26)

(本文编辑:尹卫靖)