

中医临床医生四诊信息判读及诊断一致性探讨*

□刘国萍** 王忆勤** (上海中医药大学中医四诊信息综合研究实验室 上海 201203)
 赵耐青 (复旦大学公共卫生学院卫生统计与社会医学教研室 上海 200032)
 段艳霞 许朝霞 李福凤 燕海霞
 (上海中医药大学中医四诊信息综合研究实验室 上海 201203)
 董 伟 (上海中医药大学附属龙华医院 上海 200032)

摘要:目的:评价中医临床医生四诊信息判读及诊断的一致性。方法:邀请医院资深临床内科专家,对患者四诊信息进行判读,并给出辨证诊断结果;运用本课题组制作的心系问诊量表采集的心系病例资料,随机抽取25份病例,邀请两位中医学专家分别对其进行辨证诊断;随机抽取病例15份,邀请3位中医学临床医生进行证素辨证诊断,采用Kappa值统计检验,对其进行一致性评价。结果:不同医生,同一患者的现场诊断一致性评价结果:判断舌质淡红的专家有9位,一致性56%;脉象信息判读的一致性中脉沉9位,一致性56.2%;诊断虚证的一致性为89%;同一份病例量表,同一医生先后两次诊断的一致性结果:心气虚证Kap一致性系数为0.69,心阳虚证为0.80,痰浊证为0.78;不同医生诊断一致性结果:心气虚证Kap一致性系数为0.63,心阳虚证为0.72,心阴虚证为0.48。结论:中医临床医生对四诊信息判读及诊断均存在差异性,说明中医临床医生间四诊信息判读及诊断的一致性较差。

关键词:中医学 四诊信息判读 诊断 一致性

doi: 10.3969/j.issn.1674-3849.2010.03.009

在临床科研和实际工作中,医务工作者对临床四诊信息的判读、诊断结果等都可能发生判断的不一致。本研究的目的是评价经验不同的中医临床医

生临床四诊信息判读及诊断的一致性水平。

一、研究方法

病例信息由本课题组制作的心系问诊量表数据库提供,该量表已经通过统计学的信度及效度考核,

收稿日期:2009-11-12

修回日期:2010-05-28

* 上海市卫生局青年科研项目(2008Y128):基于隐结构法的中医问诊诊断模型研究,负责人:刘国萍;上海市高校选拔培养优秀青年教师科研专项基金(syz08003):基于隐结构法的中医问诊(脾胃系)诊断模型研究,负责人:刘国萍;国家自然科学基金青年基金(30901897):基于特征选择的中医问诊信息提取及其辨证推演方法研究,负责人:刘国萍;上海市教育委员会上海市重点学科(第三期)(S30302):中医诊断学重点学科,学科带头人:王忆勤。

** 联系人:刘国萍,助理研究员,主要研究方向:中医四诊客观化及中医证候标准化研究、中医诊断学教学工作,Tel:021-51322286,E-mail:tianshanghuo3838@163.com;王忆勤,教授,主要研究方向:中医四诊客观化及中医证候标准化研究、中医诊断学教学、医疗工作,Tel:021-51322447,E-mail:wangyiqin2380@sina.com。

并已发表(刘国萍,王忆勤,董英,等,中医心系问诊量表的研制及评价.上海中西医结合学报.2009,7(1):1222-1225)。

1. 不同医生,同一患者的现场四诊信息判读及诊断一致性评价

邀请上海中医药大学附属龙华、曙光、岳阳等医院的资深临床内科专家,主要有刘嘉湘、林钟香、何立人、蔡淦、马贵同等共 16 位教授。同时随机选取上海中医药大学附属龙华医院肾内科住院患者(配合者)1 位。按照中医诊察疾病的顺序及方法,请各位专家根据患者情况,对患者四诊信息进行判读,并给出辨证诊断结果。

2. 同一医生先后两次诊断的一致性评价

从采集的心系病例中,剔出病史填写不完整、字迹不清晰及不符合标准者,随机抽取病例 25 份,对 25 份病例复印后,重新编号,从 1~50,采取双盲方法,邀请两位中医学专家对 50 份病例进行辨证诊断。

3. 不同医生,同一份病例的诊断一致性评价

从采集的心系病例中,剔出病史填写不完整、字迹不清晰及不符合标准者,随机抽取病例 15 份,邀请 3 位中医学临床专家,参考课题组提供的辨证诊断标准,给出 15 份病例的辨证诊断结果,并填写诊断依据。

二、统计学方法

应用下列公式进行 Kappa 值检验。95%可信区间用 Bootstrap 方法计算。

3 位医生诊断均为是的比例为 π_{111} , 3 位医生诊断均为否的比例为 π_{000}

第 1 位医生诊断为是的比例为 π_{1++} , 第 1 位医生诊断为否的比例为 π_{0++}

第 2 位医生诊断为是的比例为 π_{+1+} , 第 2 位医生诊断为否的比例为 π_{+0+}

第 3 位医生诊断为是的比例为 π_{++1} , 第 3 位医生诊断为否的比例为 π_{++0}

$$P_0 = \pi_{111} + \pi_{111} \quad \text{一致率}$$

$$P_c = \pi_{1++} + \pi_{+1+} + \pi_{++1} + \pi_{0++} + \pi_{+0+} + \pi_{++0}$$

$$Kap = \frac{P_0 - P_c}{1 - P_c}$$

三、结果

1. 不同医生,同一患者的现场诊断一致性评价结果

(1)舌像信息判读的一致性。

舌质:判断舌质淡红的专家有 9 位,一致性 56%,判读舌红的专家 2 位,舌质淡白的专家 5 位,舌质暗 2 位,一致性分别为 12.5%、31.2%、12.5%。

舌体:齿痕舌 1 位,一致性 0.62%,裂纹舌 1 位,一致性 0.62%,舌体瘀斑 1 位,一致性 0.62%,舌体肿大 2 位,一致性 12.5%。

苔色及苔质:苔黄 11 位,一致性 68.8%;苔白 3 位,一致性 18.8%;苔薄 11 位,一致性 68.8%;苔厚 3 位,一致性 18.8%;苔腻 11 位,一致性 68.8%;苔润 4 位,一致性 25%;苔燥 3 位。一致性 18.8%。

(2)面色信息判读的一致性。

面色:红黄隐隐 2 位,一致性 12.5%;面色淡白 2 位,一致性 12.5%;面色萎黄 5 位,一致性 31.3%;面色黄 2 位,一致性 12.5%;淡黄 3 位,一致性 18.8%;白光白 1 位,一致性 0.62%;青黄 1 位,一致性 0.62%;少泽 11 位,一致性 18.8%;无泽 3 位,一致性 18.8%。

口唇:暗红 9 位,一致性 69%;淡红 2 位,一致性 15.3%;色紫 2 位,一致性 15.3%;唇润 3 位,一致性 42.9%;唇燥 4 位,一致性 57.1%。

(3)脉象信息判读的一致性。

脉沉 9 位,一致性 56.2%;脉结 7 位,一致性 43.8%;脉弦 7 位,一致性 43.8%;脉细 7 位,一致性 43.8%;脉滑 2 位,一致性 12.5%;脉数 1 位,一致性 0.62%;脉无力 3 位,一致性 18.8%。

(4)诊断结果的一致性。

对该患者做出诊断的专家有 9 位,其中虚实诊断中,诊断此患者为虚证者有 8 位,一致性为 89%。病名诊断中,诊断为虚劳者,共有 3 位,一致性 33.3%;诊断为水肿者 2 位,一致性 22.2%;此外,诊断癱闭者 1 位,一致性 11.1%。见表 1。

病性诊断中,诊断为阳虚者 2 位,气虚 1 位,阴阳两虚 1 位专家。兼证诊断中,兼证为血瘀者 2 位,湿浊者 4 位。见表 2。

病位诊断中,其中 6 位专家定位在脾,7 位专家定位在肾,其中脾肾兼有者 5 位。见表 3。

2. 同一份病例量表,同一医生先后两次诊断的一致性(见表 4)

同一份病例量表,同一医生先后两次诊断的一致性结果:心气虚、心阳虚、痰浊、寒凝的辨证诊断一致性相对较高;对于心阴虚、心血虚、血瘀、气滞、心火

亢盛的一致性不高,尤其是阴虚的一致性最差。

3. 不同医生,同一份病例的诊断评价(见表5)

心气虚、心阳虚、心阴虚的诊断一致性的结果相对较好,说明3位专家对心气虚、心阳虚、心阴虚的诊断一致性较高;痰浊、瘀血、气滞诊断的Kap系数分别为0.03、0.14、0.17,95%可信区间为(-0.03,0.12)、(-7.57×10⁻¹⁷,0.33)、(-0.06,0.46),此3项诊断的95%可信区间包含0,因此不能认为3位医生对痰浊、瘀血、气滞的诊断具有一致性。说明这3项诊断的一致性较差。

四、讨 论

1. 发生临床不一致性的环节

临床诊断的不一致,可发生于望、闻、问、切四诊症状及体征的采集过程,以及辨证诊断等诸多环节。

(1) 四诊信息采集的不一致。

由于四诊信息采集的主观性强,并受到外界光线、环境等因素的干扰,导致四诊信息采集的不一致。对于问诊,由于询问方法不一致,导致询问症状一致性降低。据某医院两位高年医师对同一批溃疡病术后者询问病情,两人的意见一致率尚不足2/3。相对于问诊,望、闻、切诊3诊的主观性更强,尤其是脉诊,普遍认为“心中易了,指下难明”。

(2) 四诊信息综合分析的不一致。

经验不同的临床医生,即使在相同的四诊资料的情况下,做出诊断的结果也并非都一致。因其对四诊信息的判断有所侧重;对诊断信息的理解及运用不同,自身持有的诊断标准也存在一定的差异。

2. 产生临床不一致的原因

(1) 观察者的原因。

不同医生对四诊信息判断的不一致,观察者感觉上的差异,如对颜色的判读存在差别;诊断标准不明确,诊断证候分类不清;从事不同疾病的医生偏重从本专业出发考虑诊断。究其原因,与中医学本身的特点及诊断标准的不统一分不开。

①与中医学的传承密切相关 中医学的流传和发扬多以师承为主,即使是现代教育发达的今天,中医学校到处可见也不例外,故形成了许多不同的经验流派。从古至今,存在着多种不同的中医学术流派,如推崇朱丹溪的“养阴派”,推崇张从正的“攻下派”,还有现今流行的“火神派”等等。所谓仁者见仁,智者见智。这是不同医生间的诊断一致性差的主要

表1 病名诊断的一致性

病名诊断	虚证	虚劳	水肿	癃闭
专家人数	8	3	2	1
一致性	89	33.3	22.2	11.1

表2 定性诊断的一致性

证素(定性)诊断	阳虚	气虚	阴阳两虚	血瘀	湿浊
专家人数	2	1	1	2	4
一致性	22.2	11.1	11.1	22.2	44.4

表3 定位诊断的一致性

证素(定位)诊断	脾	肾	脾肾兼有
专家人数	6	7	5
一致性	37.5	43.8	31.2

表4 同一医生先后两次诊断的一致性结果

证型	Kappa 值	95%可信区间
心气虚	0.69	0.42 0.96
心阳虚	0.80	0.52 1.00
心阴虚	0.28	0.06 0.50
心血虚	0.33	0.08 0.58
痰浊	0.78	0.52 1.00
血瘀	0.33	0.08 0.58
气滞	0.54	0.19 0.89
心火亢盛	0.52	0.17 0.86
寒凝	0.72	0.34 1.00

表5 不同医生诊断一致性结果

证型	Kappa 系数	95%可信区间
心气虚	0.63	0.26 0.91
心阳虚	0.72	0.42 1
心阴虚	0.48	0.21 0.77
痰浊	0.03	-0.03 0.12
血瘀	0.14	-7.57×10 ⁻¹⁷ 0.33
气滞	0.17	-0.06 0.46

原因。比如瘀血一证,有的医生认为但见如刺痛、舌紫等1个征象即可诊断;而有的医生则认为要征象表现明显才可诊断瘀血。

②与四诊信息判读与诊断的主观性相关 证候是中医学特有的概念。《中国医学百科全书·中医学》指出“证是综合分析了各种症状和体征,对于疾病处于一定阶段的病因、病位、病变性质,以及邪正双方力量对比各方面情况的病理概括。”辨证是中医学的精髓,体现了中医学理论特色与优势。由于中医传统理论的特点,以及在其指导下的中医药疗

法的作用特点,中医历来从整体角度考察病人的证候变化,是一种复杂的非线性思维,尤其注重病人的自我感受。相对于现代医学主要以实验室及影像学检查等硬指标为主来说,中医学更为重视软指标。所谓软指标主要是指可重复性差、不能准确度量的指标,是通过临床医生的观察采集或患者的自我报告而获得的,无法用仪器或其他手段直接进行精确测量的主观指标^[1]。软指标反映患者主观感受的功能是任何硬指标所无法取代的^[2]。但传统中医证候的确定是建立在望、闻、问、切四诊合参的基础上,由医家思辨而经验性地形成的,不可避免的带有具有主观性、模糊性,并且受情绪和环境等因素的影响。人的感觉又是无形的,“软”亦指其无形,因此,症状软指标信息大多难以准确度量,测量的可重复性也差,测量值受研究样本及环境的影响较大,这是导致四诊信息判读及同一医生前后诊断不一致的主要原因。

③与诊断标准制订的不统一有关 近年来,在国家大力支持和众多学者努力下,政府部门和专业组织相继出台了一系列的证候诊断标准:①教材标准的制订。各诊断学教材及内科学教材等教材中分别制订了相应的证候诊断标准;②国家于1997年发布了《中医临床诊疗术语证候部分》,对证候诊断进行规范;③2001年公布的《中药新药临床指导原则》也成为了临床及科研的重要参考标准;④各内科专业委员会制订的诊断标准。还有对单纯的证候诊断标准,如全国中西医结合虚证与老年病研究专业委员会制定的《中医虚证辨证参考标准》和在某种中医或西医疾病名称下的证候诊断标准,如中国中医药学会内科学会组织中医临床专家讨论拟定了《中风病中医诊断疗效评定标准》等。以上的标准都成为了临床医生及科研人员的参考标准,但这些标准之间并未完全达成共识,也就是现今所说的中医诊断的金标准并未诞生。这就导致临床医生分别结合自己的经验参考相关的标准形成自己的辨证诊断思路。也是现今导致不同医生间诊断不一致的又一主要原因。

(2)被检查者的原因。

患者病情叙述支离破碎,主次不分,或对病情有隐讳;伸舌姿势不对等原因导致舌像判读的不一致;化妆掩盖面色导致面色判读的不一致等原因。

3. 防止临床判读及诊断不一致的方法

中医学的模糊性、多变性、复杂性及隐匿性决定了中医证候诊断必须通过多学科的交叉渗透来完成。所以,在目前证候研究^[3]没有形成“金标准”的情况下,根据系统性、复杂性科学的要求,在中医证候诊断规范化研究中,实行定性与定量相结合、宏观与微观相结合,加强证候概念及专业术语的规范化研究,通过DME、Delphi法较为全面、客观、真实地收集、分析、提炼临床症状信息,充分利用计算机技术、人工智能、信息技术等高新技术挖掘背后的潜在数据规律,构建证候诊断模型,将是证候诊断规范化研究的必然趋势。所以,这是一项非常复杂的工作,只有通过不断的探索,完善现有的研究思路与方法,中医证候诊断的规范化研究才有望取得实质性的突破与进展。

(1)中医诊断的客观化研究。

四诊客观化目的主要在于量化诊断,以提高辨证的准确性。到目前为止,中医师多依据感觉器官,通过望、闻、问、切来获取临床信息,信息收集过程中主观性很大,影响了研究结论的真实性和可靠性。运用相关的计算机数据库技术、数据挖掘技术、图像处理及模式识别等技术方法,进行中医望、闻、问、切四诊信息的规范化和客观化信息采集及提取,可在一定程度上解决四诊信息判读及诊断的主观性。

(2)诊断的标准化研究^[4]。

遵循DME的原则,力求资料的科学性、可靠性、准确性,在文献调研、专家咨询和病例回顾的基础上,开展多中心、大样本、前瞻性的临床研究,对获取的证候资料进行数理统计分析,使研究结论具有普遍性和可重复性,提高证候规范化研究的论证强度。同时也可以按照DME诊断性试验原则选择敏感度、特异度、阳性似然比等相关指标对证候诊断标准的诊断效能进行评价和验证。以此,制订中医不同疾病及证候的统一诊断标准,可在一定程度上减少临床诊断的不一致,有利于不同中医医生间的沟通和交流。

综上所述,目前中医临床医生对四诊信息判读及诊断的一致性不高,中医诊断的准确性差,对疗效的影响值得今后进一步研究。

参考文献

- 1 郑淑美,胡立胜,李友林.中医软指标的筛选方法.中国中医基础医学杂志,2008,14(4):304-305.

- 2 朴海垠,谢雁鸣.中医软指标的测量方法初探.中医药学刊,2006,24(6):1018-1020. 2008,25(6):465-468.
- 3 赵晖,陈家旭.试论证候诊断规范化的研究思路和方法.天津中医药, 4 赵晖,吴崇胜,陈家旭.中医证候诊断标准研究的方法学探讨.上海中医药大学学报,2008,22(4):47-50.

Study on the Diagnosis Agreement of Clinical Doctor of Traditional Chinese Medicine

Liu Guoping¹, Wang Yiqin¹, Zhao Naiqing², Duan Yanxia¹,

Xu Zhaoxia¹, Li Fufeng¹, Yan Haixia¹, Dong Wei³

(1. Diagnostic Information Laboratory of Traditional Chinese Medicine, Shanghai University of Traditional TCM, Shanghai 201203, China;

2. Health Statistics and Social Medical Department, Public Health College, Fudan University, Shanghai 200032, China;

3. Longhua Hospital Affiliated to Shanghai University of Traditional Chinese Medicine, Shanghai 200032, China)

Abstract: This work aimed to assess the concordance of information collected by the TCM four diagnostic methods and diagnosis. 16 TCM experts were invited from Longhua Hospital, Shuguang Hospital, and Yueyang Hospital of Shanghai University of Traditional Chinese Medicine, and randomly selected 1 patient. They recorded patient information by the TCM four diagnostic methods and diagnosis. 25 cases were randomly selected, and 2 copies were made from the numerical data collected through the Inquiry Scale for Diagnosis of Heart System Syndromes developed by our research team. The copies were disorganized and renumbered from 1 to 50. In order to evaluate the concordance of successive diagnosis by the same clinician, 2 TCM experts were invited to differentiate the syndrome and diagnose for the 50 cases by the double-blind method. 15 cases were randomly selected. Three TCM clinicians were invited to make a diagnosis based on Zangfu-organs syndrome differentiation. The Kappa test was adopted. The results of experts judgement for the same patient showed that 9 experts judged red tongue, with the concordance being 56%, 9 experts judged deep pulse, with the concordance being 56.2%, and the concordance of deficiency syndrome was 89%. The results of successive diagnosis by the same clinician showed that the Kap concordance coefficient for the syndrome of deficiency of heart qi was 0.69, that for the syndrome of deficiency of heart yang was 0.80, and that for the syndrome of turbid phlegm was 0.78. The results of diagnosis by different clinicians showed that the Kap concordance coefficient for the syndrome of deficiency of heart qi was 0.63, that for the syndrome of deficiency of heart yang was 0.72, and that for the syndrome of deficiency of heart yin was 0.48. Taken together, diagnosis by the same clinician and different clinicians varied, indicating a poor concordance of the TCM four diagnostic methods and diagnosis.

Keywords: Traditional Chinese medicine, Information from the TCM four diagnostic methods, Diagnosis, Concordance

(责任编辑:李沙沙,责任译审:张立崑)