

# 复合性诊断系统的编制及其信度、效度测试

邹义壮 沈渔邨 舒良 王玉凤 冯锋 王健 刘学俊

**【摘要】** 目的 研制适合国内临床使用的精神疾病计算机辅助诊断系统。方法 在复合性国际诊断交谈表(CIDI)基础上,将绝大部分项目由他评变为自评,简化了检查过程,在诊断中使用了人工神经网络技术,建立了复合性诊断系统(CDS)。在精神分裂症、神经症各 70 例和 70 名正常人中对该系统进行跨效度检验。结果 4 名检查者使用 CDS 的他评一致性为 94.39%;被检查者的自评重测一致性为 92.44%。CDS 跨效度检验与临床诊断的符合率的 Kappa 值为 0.96~0.99;对疑难病例诊断的符合率的 Kappa 值为 0.89。CDS、CIDI-国际疾病分类第 10 版(ICD-10)和 CIDI-美国精神疾病诊断与统计手册第 3 版修订本(DSM-III-R)的诊断结果与临床诊断比较,符合率的 Kappa 值分别为 0.96, 0.76 和 0.72。结论 CDS 具有满意的信度和效度,使用和操作远较 CIDI 方便,培训简单、可以使用光学符号阅读机输入答卷,有希望成为精神科临床广泛使用的辅助诊断工具。

**【关键词】** 精神病学 诊断,计算机辅助

**Composite Diagnostic System and artificial neural network** ZOU Yizhuang\*, SHEN Yucun, SHU Liang, et al. \*Beijing Huilongguan Hospital, Beijing 100096

**[Abstract]** **Objective** To develop a computerized psychiatric assistant diagnostic system —— the Composite Diagnostic System (CDS) which was revised from the Composite International Diagnostic Interview (CIDI), for its efficient application in China. **Methods** Most of the items were changed into self-rating from others-rating and the artificial neural network was used to assist diagnosis in the CDS. The research samples included 70 schizophrenics, 70 neurosis and 70 normal people. **Results** The reliability between 4 raters was 94.39% and the test-retest reliability was 92.44%. The overall correct Kappa between the CDS and the clinical diagnosis was 0.96~0.99 and with the 18 difficult-diagnostic patients, the Kappa was 0.89. The CDS and CIDI and the ICD-10 were parallel compared with clinical diagnosis, the overall Kappa were 0.96, 0.76 and 0.72, respectively. **Conclusions** The CDS is of higher diagnostic validity and more convenient than the CIDI in administration, and it could be widely used by clinicians in China.

**[Key words]** Psychiatry Diagnosis, compute-assistant

复合性国际诊断交谈表<sup>[1,2]</sup>(Composite International Diagnostic Interview, CIDI)在国内现场的测试结果显示,CIDI 具有良好的信度、效度和文化适应性<sup>[3,4]</sup>,但其项目多、检查过程复杂,常使临床工作者感到难以接受,从而降低了实际应用价值。本研究目的是在 CIDI 系列研究的基础上,对该量表的内容和结构进行改造,将大部分项目由他评改为自评,简化检查过程,并应用计算机人工神经网络技术辅助诊断,建立复合性诊断系统(Composite Diagnostic System, CDS),通过临床测试,对其信度和效度进行评价。

## 资料和方法

### 一、资料

全部患者和正常人的资料均于 1994 年 1~6 月采集于北京回龙观医院。

1. 患者组:共 140 例,其中精神分裂症 70 例:男 49 例,平均年龄(34±10)岁,平均受教育(11±3)年;女 21 例,平均年龄(33±10)岁,平均受教育(12±3)年。神经症 70 例:男 45 例,平均年龄(31±10)岁,平均受教育(13±3)年;女 25 例,平均年龄(35±14)岁,平均受教育(12±3)年。

考虑到 CDS 系统将主要应用于门诊或新入院、急性期、症状比较活跃的患者,因此患者组样本较多是新住院的患者,精神分裂症患者多在急性期和症状活跃期,少部分为慢性患者。并要求由 2 名以

作者单位:100096 北京回龙观医院(邹义壮、冯锋、王健、刘学俊);北京医科大学精神卫生研究所(沈渔邨、舒良、王玉凤)

上医生(其中至少 1 名是主治医师)根据中国精神疾病分类方案第 2 版独立进行诊断,诊断一致方可入组。此外,精神分裂症患者的病程要求在半年以上。

2. 正常人组:共 70 名,均为随机抽取的医院职工,其中男 30 名,平均年龄( $35 \pm 11$ )岁,平均受教育( $12 \pm 3$ )年;女 40 名,平均年龄( $34 \pm 9$ )岁,平均受教育( $13 \pm 3$ )年。

为了检验 CDS 系统对临床疑难病例的诊断能力,有意采集了 18 例临幊上诊断困难和不典型的病例,疑难病例的标准是:(1)在治疗过程中曾更改过诊断;(2)症状表现不典型,经科主任或院长查房最后确诊的病例。18 例中神经症 11 例,精神分裂症 7 例。

## 二、方法

1. 量表的编制:以在量表长期研究与应用实践中积累的经验,将自评量表与他评量表相结合,建立自评为主、他评为辅的量表设计思路。CDS 的项目大部分取自 CIDI,并根据我国的实际情况删除和增加了一些项目,问卷共有 556 个项目,所有问题只有“是”和“否”两种回答。CDS 分为自评(507 题)和他评(49 题)两部分。对阳性症状采取患者自评,对难以自评的定向力、记忆力、智力、病前应激、自知力、行为障碍特别是阴性症状和答題行为等进行他评。CDS 的检查过程分以下三部分:(1)查时首先由经过培训的心理技师或医师询问被试者的一般情况,检查定向力、常识、计算力、记忆力、理解力、病前应激、自知力等情况,这一阶段平均需要 5~8 分钟。要求检查者在这一段检查中与被试者建立起较好的合作和相互信任关系。(2)指导被试者自己回答问题。允许检查者在初步了解被试者主要问题的情况下,在自评部分让被试者首先回答与主要问题相关的部分。被试者自己回答 CDS 自评部分的时间一般约需 1 小时,对个别一次不能完成的患者,可分为两次完成。自评阶段允许多个患者同时进行,要求检查者在一旁观察并回答患者的问题。(3)自评部分完成后,检查者要对被试者的行为进行评定,检查光电答卷纸的填写情况,并对患者答題态度是否认真做出判断。答卷通过光学符号阅读机输入计算机。本研究中的统计部分使用社会科学统计软件包(SPSS/PC+)。人工神经网络在北京邮电大学钟义信教授的指导下,由该校的王明会、刘伟权博士协助开发。

2. CDS 的信度:(1)CDS 他评信度:对检查者间的一致性进行测试,4 名检查者轮流担任提问者和

观察者,共对 9 例患者进行了 CDS 检查。(2)CDS 自评信度:CDS 被检查者的自评信度使用了短期重测信度的方法评定,并按以下规定程序操作,以减少对信度评估的干扰:2 次 CDS 的间隔时间控制在 1~5 天;2 次 CDS 在不同日的同一时间进行,如第 1 次 CDS 是在上午 9 点开始,则第 2 次应在隔一天或几天后仍是上午 9 点进行;2 次 CDS 在同一地点(房间)中进行;2 次 CDS 由同一名检查者施测。比较被试者在 2 次 CDS 测查中对各项目的回答,以评价其重测信度。

3. CDS 的效度:使用 Back Propagation(BP)人工神经网络模型<sup>[5]</sup>,全部自评和他评结果共 556 个项目作为网络的输入,网络输出根据诊断类别设置正常、神经症和精神分裂症等 3 类。样本的训练集使用 60 例样本,跨效度(cross validity)检验集有 150 例样本。人工神经网络的隐层神经元节点由试验确定为 40。各节点的原始权值由计算机随机产生,通过训练建立人工神经网络的诊断模型。各节点的原始权值由计算机随机产生。

## 结 果

1. 检查者间信度:他评部分(共 36 题)检查者间的总一致性为 94.39%,其中定向记忆智力部分 97.05%、自知力 100.00%、阳性症状 89.51%、病前应激 88.88% 和阴性症状 80.86%。

2. 自评的重测信度:参加重测信度的被试者共有 91 人,其中正常人 29 名,神经症 28 例,精神分裂症 10 例,其他 24 例。总体重测信度为 92.44%,其中正常人为 98.61%,神经症为 93.10%,精神分裂症为 93.15%。

3. 效度:使用跨效度方法比较 CDS 诊断与临床诊断,CDS 诊断与临床诊断的总符合率为 99.33%, $Kappa = 0.99, P < 0.01$ (表 1)。18 例疑难病例的 CDS 的诊断一致性为 94.44%, $Kappa = 0.89, P < 0.01$ 。

表 1 CDS 诊断与临床诊断比较(例数)

临床诊断	例数	CDS 诊断		
		正常	神经症	精神分裂症
正常	50	50	0	0
神经症	50	0	50	0
精神分裂症	50	0	1	49

4. CDS 与 CIDI 的两套计算机诊断程序比较:以

CDS 和根据国际疾病分类第 10 版 (ICD-10) 标准进行诊断的 CIDI/ICD-10、根据美国精神疾病诊断与统计手册第 3 版修订本 (DSM-III-R) 标准诊断的 CIDI/DSM-III-R<sup>[1]</sup>, 观察对同一组样本的诊断效度。样本共 99 人, 其中正常人 45 名, 神经症 31 例, 精神分裂症 23 例。3 套诊断系统与临床诊断的符合率的 Kappa 值分别为: 0.96、0.76 和 0.72。表 2。

表 2 3 种诊断系统对 99 例样本的诊断效度比较 (%)

项目	CIDI/ICD-10	CIDI/DSM-III-R	CDS
神经症			
敏感性	64.55	64.55	100.00
特异性	100.00	100.00	98.53
符合率	88.89	88.89	98.99
Kappa 值	0.71	0.71	0.98
精神分裂症			
敏感性	82.61	78.26	95.65
特异性	98.68	100.00	100.00
符合率	94.95	94.95	98.99
Kappa 值	0.85	0.85	0.97
总符合率	84.85	83.84	97.98
Kappa 值	0.76	0.72	0.96

## 讨 论

Robins 等<sup>[2]</sup> 报道对 575 例患者进行 CIDI 检查, 评定者间的一致性为 67% ~ 99%。舒良等<sup>[4]</sup> 报道 CIDI 的他评信度为 98.70%。邹义壮等<sup>[3]</sup> 对 CIDI 他评信度的试验结果为 93.50%。本结果显示 CDS 的他评部分的评定者间的一致性信度为 94.39%, 自评部分重测信度为 92.44%, 是令人满意的。

诊断效度的检验以有经验的临床医生 (2 名) 的诊断作为金标准, 对 CDS 进行跨效度检验。此外, 为了排除可能存在的样本典型造成的误差, 试验中还特别使用了疑难病例对 CDS 的效度进行了检验。CDS 的诊断符合率的 Kappa 值为 0.96 ~ 0.99, 对疑难病例的诊断符合率的值 Kappa 为 0.89。应当指出的是, 疑难病例是一种比较特殊的情况, 其实际诊断效度应比对非疑难病例的诊断效度要高一些。

另外一个重要结果是, CDS 在大量使用自评方法的情况下, 诊断效度明显高于 CIDI/ICD-10 和 CIDI/DSM-III-R, 3 套诊断系统与临床诊断符合率的 Kappa 值分别是: 0.96、0.76 和 0.72。提示在诊断量表中使用自评和他评方法的结合是可行的。自评量表不用面对医生、答题阻力小、施测方便, 可以对多人同时进行检查, 特别适合于正常人、神经症和能够

合作的精神分裂症患者使用。但是, 对于不合作的患者, 自评量表的使用受到一定的限制。他评量表则能较好地补充自评量表的不足, 能对症状较重的精神分裂症患者进行评定, 特别是给对自评难以解决的阴性症状的评定提供了手段。CDS 集自评和他评方法于一身, 相互取长补短, 最大限度地发挥了各自的优势。

CDS 与国内其它诊断工具比较: 郑延平等<sup>[6]</sup> 报道, 便携式微电脑诊断系统与临床诊断总符合率的 Kappa 值为 0.62。樊彬等<sup>[7]</sup> 用 PSE/CATEGO 与临床诊断合率的 Kappa 值为 0.22 ~ 0.44; 张明圆<sup>[8]</sup> 使用 DIS/DSM-III 与临床诊断符合率的 Kappa 值为 0.42 ~ 0.56; 邹义壮等<sup>[9]</sup> 使用 MMPI 计算机解释诊断系统与临床诊断符合率的 Kappa 值为 0.77; 使用人工神经网络方法辅助 CIDI 诊断的符合率的 Kappa 值为 0.94<sup>[10]</sup>。而 CDS 与临床诊断跨效度符合率的 Kappa 值为 0.96 ~ 0.99, 对疑难病例诊断符合率的 Kappa 值为 0.89。这显示 CDS 的诊断效度是最高的, 如此高的诊断效度是传统方法难以实现的。

CDS 诊断效度高主要归因于: (1) CDS 量表编制合理, 自评他评结合, 建立在 CIDI 基础之上, 拥有大量能够区分不同疾病的项目。(2) CDS 中使用的人工神经网络具有迄今为止最强的分类诊断能力, 采用多层次的 BP 结构, 网络的容错性好, 善于处理模糊和经验的数据<sup>[5]</sup>, 从而充分利用了 CDS 中的诊断信息, 最大限度地模拟了临床医生的诊断经验。

总之, 在 CDS 的研制过程中, 使用了先进的人工神经网络技术, 对数据在项目水平上进行分类诊断, 充分利用了答卷中的信息, 提高了系统处理模糊数据的能力, 保证了 CDS 系统具有较高的诊断效度。CDS 集自评和他评的优势于一身, 在量表的内容和诊断上更加符合我国国情, 并使用光学符号阅读机输入答卷, 提高了效率, 减少了误差。在 CDS 设计上集中了当代多学科发展的最新成就。本结果显示, CDS 具有令人满意的信度和效度, 培训简单、操作方便、容易推广, 初步达到了本研究预想目的。CDS 系统有希望成为一种准确可靠和方便实用的精神科临床辅助诊断工具。

## 参 考 文 献

- 1 世界卫生组织. 复合性国际诊断访谈检查 (CIDI) 核心本检查者用本、使用者手册和计算机使用手册. 北京医科大学精神卫生研究所编译. 1991.
- 2 Robins LN, Wing J, Wittchen HU, et al. The Composite International

- Diagnostic Interview. Arch Gen Psychiatry, 1988, 45: 1069-1077.
- 3 邹义壮, 沈渔邨, 舒良, 等. 精神疾病诊断的方法学研究——量表诊断与人工神经网络: [学位论文]. 北京: 北京医科大学精神卫生研究所, 1994.
- 4 舒良, 张鸿燕, 汪向东, 等. “复合性国际诊断交谈检查—核心本”现场测试——信效度检验. 中国心理卫生杂志, 1993, 7: 208-210.
- 5 钟义信, 潘新安, 杨义先. 智能理论与技术——人工智能与神经网络. 北京: 邮电出版社, 1993. 406-488.
- 6 郑延平, 许乐意, 黄能达, 等. 便携式微电脑诊断系统在临床中的试用. 中华神经精神科杂志, 1987, 20: 84-89.
- 7 樊彬, 张明园, 王征宇, 等. 精神科的计算机诊断——PSE/CATEGO 的应用. 中国神经精神疾病杂志, 1986, 12: 280-282.
- 8 张明圆. 标准化精神检查和诊断标准的评价. 中国神经精神疾病杂志, 1985, 11: 161-166.
- 9 邹义壮, 赵传铎. MMPI 临床诊断效度的研究. 中国心理卫生杂志, 1992, 6: 211-214.
- 10 Zou Y, Shen YC, Shu L, et al. Artificial neural network to assist psychiatric diagnosis. Br J Psychiatry, 1996, 169: 64-67.

(收稿: 1998-01-12 修回: 1998-07-28)

(本文编辑: 杨小昕)

## ·论著摘要·

### 东北三省移居国外者发生精神疾病情况的调查

张聪沛 李晶 张延承 王琳

通过调查在中国长大、在国外发病的 56 例精神疾病患者, 分析突然面对异国文化所出现的不适应及精神疾病。

1. 对象和方法: 56 例均符合以下条件: (1) 出国前均长期居住在东北三省; (2) 在国外连续居住至少半年以上; (3) 在国外首次发病; (4) 回国后居住不超过 1 年。诊断标准符合国际疾病分类第 10 版精神与行为障碍分类和中国精神疾病分类方案与诊断标准第 2 版。病例来自我院住院患者 31 例、门诊 14 例、作者出国期间线索调查面诊 11 例。56 例中男 19 例, 女 37 例; 年龄 21~47 岁。国外职业: 学生 18 例, 家务 11 例, 杂工 11 例, 职员 9 例, 研究人员 3 例, 经商 2 例, 其他 2 例。出国后平均发病时间 1.7 年。诊断: 精神分裂症 19 例, 应激适应障碍 14 例, 神经症 23 例 (癔病 1 例, 焦虑症 6 例, 抑郁性神经症 9 例, 神经衰弱 7 例)。发病所在国家为: 日本 23 例, 俄罗斯 11 例, 美国 9 例, 澳大利亚 7 例, 加拿大 6 例。在国外的适应调查采用 1991 年日本、中国和美国协作调查中的精神卫生

心理适应问卷。其中有 40 个条目, 分为语言困难、思念亲人、生活不习惯、人际关系困难、学习工作困难、经济问题 6 个项目。对精神分裂症患者除自己填写外, 还要由知情人填写一份进行比较。

2. 结果: 年龄与患病种类间的差异无显著性 ( $\chi^2 = 6.62, P > 0.05$ )。应激适应障碍多在移居后的半年内发病 (11 例, 79%), 而精神分裂症、神经症多在移居半年后发病 (35 例, 83%)。暂时移居者中 (23 例) 以应激适应障碍最多 (11 例, 48%), 永久移居者 (33 例) 以神经症和精神分裂症者为多 (30 例, 91%), 暂时移居者与永久移居者间差异有非常显著性 ( $\chi^2 = 10.92, P < 0.01$ )。主动情愿移居 (多为留学、投亲、赚钱等) 的 19 例中, 以精神分裂症为多 (11 例, 58%); 被动不情愿移居 (主要是随家迁居、陪读等) 的 37 例中, 以神经症为多 (18 例, 49%), 二者差异有显著性 ( $\chi^2 = 7.37, P < 0.05$ )。适应问卷结果, 56 例均存在不同程度的不适应。其中思念亲人和生活不习惯发生率最高, 分别为 57% 和 52%。精神分裂症、神经症和应激适应障碍患者间各种适应问题的发生率差异有显著意义者为思念亲人和经济问题。思念亲人中以神经症发生率最高 (87%), 精神分裂症者

最低 (26%;  $\chi^2 = 16.01, P < 0.01$ ); 经济问题中则以精神分裂症发生率最高 (37%), 应激适应障碍最低 (7%;  $\chi^2 = 7.08, P < 0.05$ )。此外生活不习惯的发生率, 3 种疾病的患者间比较差异虽未达显著性 ( $\chi^2 = 4.93, P > 0.05$ ), 但若以神经症 (最高) 与精神分裂症比较 (最低), 差异有显著性 ( $\chi^2 = 4.71, P < 0.05$ )。

讨论: 近年来随着我国的改革开放, 大批中国人以不同方式移居国外。他们在传统的中国文化背景下接受教育, 而一夜之间又置身于陌生的文化环境中, 遭受各种困难, 又得不到亲友的支持, 出现了一系列适应困难。这些不适应在被动不情愿移居者中表现得更明显, 并在移居的早期会构成应激适应障碍的发病基础; 本研究结果显示应激适应障碍发病多在移居半年内, 且多为不情愿被动移居者, 证实了这一点。长期处在异族文化背景下, 思念亲人、生活不习惯、经济问题等不适应问题, 会成为长期心理压力, 导致精神卫生问题的增多。提示文化差异、心理适应与精神疾病有密切联系。

(收稿: 1998-11-09)

(本文编辑: 刘宏林)

作者单位: 150056 哈尔滨市神经精神医院 (张聪沛、李晶、张延承); 黑龙江省戒毒中心 (王琳)