

• 方法学研究 •

基于项目反应理论的儿童睡眠问卷项目分析

邵隽¹, 许志飞², 李晓丹¹, 杜江南¹, 王桂香¹, 马军³, 胡佩瑾³, 阎小妍⁴,
张杰¹, 张亚梅¹, 刘原虎¹, 赵靖¹, 郑莉¹, 陈俊⁵, 冯国双⁶, 倪鑫^{1*}

【摘要】 目的 采用项目反应理论对儿童睡眠问卷进行项目分析, 评价儿童睡眠问卷在国内的适用性。**方法** 2015年5—7月采用多阶段分层整群随机抽样方法, 随机抽取了北京市西城、朝阳、昌平、顺义、房山、怀柔 and 门头沟共7个区县的25所幼儿园和中小学学校, 共计11 420例儿童。采用儿童睡眠问卷调查其睡眠情况, 该问卷主要包含与儿童睡眠呼吸障碍相关的4个维度, 即呼吸、睡眠、行为、其他。统计分析采用基于二参数Logistic模型的项目反应理论, 计算各条目的区分度和难度参数, 并绘制项目反应特征曲线。**结果** 大多数条目具有较好的区分度, 但“睡觉偶尔尿床”和“是否超重”的区分度明显低于其他条目(区分度参数估计值分别为0.441和0.124), 且难度参数均较大(分别为4.378和6.172)。测验信息量结果提示删除这两个条目后, 问卷提供的信息量几乎没有变化。**结论** 儿童睡眠问卷在评价儿童睡眠障碍中可以提供较好的初筛效果, 但国内应用可以考虑对其进行简化。

【关键词】 睡眠; 问卷调查; 项目反应理论; Logistic模型

【中图分类号】 R 338.63 **【文献标识码】** A DOI: 10.12114/j.issn.1007-9572.2018.00.129

邵隽, 许志飞, 李晓丹, 等. 基于项目反应理论的儿童睡眠问卷项目分析[J]. 中国全科医学, 2018, 21(23): 2844-2848. [www.chinagp.net]

TAI J, XU Z F, LI X D, et al. The item analysis of Pediatric Sleep Questionnaire based on the item response theory [J]. Chinese General Practice, 2018, 21(23): 2844-2848.

The Item Analysis of Pediatric Sleep Questionnaire Based on the Item Response Theory TAI Jun¹, XU Zhi-fei², LI Xiao-dan¹, DU Jiang-nan¹, WANG Gui-xiang¹, MA Jun³, HU Pei-jin³, YAN Xiao-yan⁴, ZHANG Jie¹, ZHANG Ya-mei¹, LIU Yuan-hu¹, ZHAO Jing¹, ZHENG Li¹, CHEN Jun⁵, FENG Guo-shuang⁶, NI Xin^{1*}

1.Department of Otorhinolaryngology Head and Neck Surgery, Beijing Children's Hospital, Capital Medical University, Beijing 100045, China

2.Department of Respiratory Medicine, Beijing Children's Hospital, Capital Medical University, Beijing 100045, China

3.Institute of Child and Adolescent Health of Peking University, Beijing 100083, China

4.Peking University Clinical Research Institute, Beijing 100083, China

5.Department of Academic Research, Beijing Children's Hospital, Capital Medical University, Beijing 100045, China

6.Clinical Epidemiology and Evidence-Based Medicine Center, Beijing Children's Hospital, Capital Medical University, Beijing 100045, China

*Corresponding author: NI Xin, Chief physician; E-mail: nixin@bch.com.cn

【Abstract】 Objective To evaluate the application of Pediatric Sleep Questionnaire (PSQ) by item analysis using item response theory.**Methods** Data were obtained from the survey of children at the age of 3-14 years from 7 districts in Beijing from May to July 2015. By using multi-stage stratified cluster random sampling method, a total of 11 420 children from 25 kindergartens, primary and middle schools were selected in 7 districts (Xicheng, Chaoyang, Changping, Shunyi, Fangshan, Huairou and Mentougou). All the parents of children were recruited to complete the PSQ, which consists of four dimensions related to sleep disordered breathing in children: breathing, sleep, behavior and others. The item response theory with two-parameter Logistic model was used to conduct item analysis to calculate the discrimination and difficulty of every item and draw item characteristics curve.**Results** The analysis of two-parameter Logistic model showed that most of the items had

基金项目: 国家重点研发计划项目(2017YFC0112502); 北京市科委培育项目(Z161100000116050); 首都卫生发展科研专项项目(2018-1-2091)

1.100045北京市, 首都医科大学附属北京儿童医院耳鼻咽喉头颈外科 2.100045北京市, 首都医科大学附属北京儿童医院呼吸科 3.100083北京市, 北京大学儿童青少年卫生研究所 4.100083北京市, 北京大学临床研究所 5.100045北京市, 首都医科大学附属北京儿童医院科研处 6.100045北京市, 首都医科大学附属北京儿童医院临床流行病学与循证医学中心

*通信作者: 倪鑫, 主任医师; E-mail: nixin@bch.com.cn

a good discrimination, but the items of “Does your child occasionally wet the bed” and “Is your child overweight” had a significant lower discrimination parameter (0.441 and 0.124, respectively) and a higher difficulty parameter (4.378 and 6.172, respectively). The test information function showed almost the same curves of the questionnaire with or without the two items.

Conclusion A simplified version of PSQ is recommended in Chinese children, although a good evaluation effect in the screening of children's sleep disorder is found.

【Key words】 Sleep; Questionnaires; Item response theory; Logistic models

儿童阻塞性睡眠呼吸暂停综合征 (obstructive sleep apnea syndrome, OSAS) 是最常见的睡眠呼吸障碍疾病, 该病可以导致多种并发症的发生^[1], 国外和香港地区报道的患病率约为 5.7% 和 4.8%^[2-3]。目前诊断 OSAS 的金标准是多导睡眠监测 (polysomnography, PSG), 然而 PSG 的检测成本较高, 难以在国内外广泛应用。儿童睡眠问卷 (Pediatric Sleep Questionnaire, PSQ) 是美国儿科学会在 2012 年儿童 OSAS 指南中推荐的筛查问卷^[2], 但该问卷目前在国内尚未推广, 仅有基于小规模人群的评价^[4], 缺乏大规模流行病学调查人群的验证。本研究采用项目反应理论 (item response theory, IRT) 对该问卷在诊断 OSAS 中的效果进行评价, 以提供基于大规模人群的证据。

1 对象与方法

1.1 研究对象 2015 年 5—7 月采用多阶段分层整群随机抽样方法, 按北京市的四大功能区 (首都功能核心区、城市功能拓展区、城市发展新区和生态涵养发展区) 分层, 随机抽取西城区、朝阳区、昌平区、顺义区、房山区、怀柔区和门头沟区共 7 个区, 然后从 7 个区中整群随机抽取 10 所幼儿园、7 所小学和 8 所初中。抽中的学校中, 除小学六年级和初中三年级外, 其余班级均进行调查, 共计 11 420 例儿童。

1.2 调查方法 采用横断面调查方法, 由首都医科大学附属北京儿童医院、北京大学儿童青少年研究所以及抽中 7 个区县的儿童青少年保健所共同合作调查。调查内容均在调查人员的指导下, 由儿童家长或监护人填写。PSQ 是 CHERVIN 等^[5]开发的针对儿童睡眠障碍调查的问卷, 该问卷包含与儿童睡眠呼吸障碍相关的 4 个维度, 即呼吸、睡眠、行为、其他, 共 22 个题项 (见表 1)。选项均为“是”“否”二分类选项, 得分越高提示相关症状越严重。本研究被调查儿童及监护人均知情同意, 且本研究获得首都医科大学附属北京儿童医院伦理委员会批准。

1.3 质量控制 调查前由儿童睡眠学临床专家和流行病学专家对调查员进行规范化培训, 调查以班级家长会的形式进行, 邀请被调查儿童的家长参加, 要求其照顾被调查儿童生活起居的监护人, 并对其生活习惯特别是睡眠状况有深入了解。由专业耳鼻咽喉科医师对调查问卷所涉及内容进行逐一讲解, 现场调查时研究人员将每次问答时间控制在 5~10 min。

1.4 统计学方法 采用 SAS JMP Clinical 6.0 统计学软件进行数据分析, 计量资料符合正态分布以 $(\bar{x} \pm s)$ 表示, 不符合正态分布以 $M(P_{25}, P_{75})$ 表示; 计数资料的统计描述采用相对数表示。统计分析采用 IRT 建立研究对象对条目的反应与潜在变量 (本研究为睡眠障碍) 之间的非线性关系, 采用二参数 Logistic 函数建立二者的项目特征函数 (item characteristic

表 1 PSQ 中英文版内容

Table 1 PSQ in Chinese and English version

维度	题号	源问卷	中文版
呼吸	A1	While sleeping, does your child snore more than half the time?	睡觉打鼾超过睡眠时间的一半
	A2	While sleeping, does your child always snore?	睡觉总是打鼾
	A3	While sleeping, does your child snore loudly?	睡觉鼾声重
	A4	While sleeping, does your child have “heavy” or loud breathing?	夜间睡觉时呼吸沉重或呼吸声大
	A5	While sleeping, does your child have trouble breathing, or struggle to breathe?	夜间睡觉时呼吸困难或呼吸费力
	A6	Have you ever seen your child stop breathing during the night?	睡眠时是否观察到有呼吸停止
	A7	Does your child tend to breathe through the mouth during the day?	白天愿意张嘴呼吸
	A8	Does your child have a dry mouth on waking up in the morning?	早晨睡醒时有口干
	A9	Does your child wake up with headaches in the morning?	早上起床时伴头痛
睡眠	B1	Does your child have a problem with sleepiness during the day?	白天睡意的困扰或总爱睡觉
	B2	Has a teacher or other supervisor commented that your child appears sleepy during the day?	老师或其他看护人评价孩子白天表现为昏昏欲睡
行为	C1	This child often does not seem to listen when spoken to directly	在跟孩子讲话时他/她好像心不在焉或没有在听
	C2	This child often has difficulty organizing task and activities	较难安排活动或任务
	C3	This child often is easily distracted by extraneous stimuli	易被外部的刺激因素分散注意力
	C4	This child often fidgets with hands or feet or squirms in seat	坐下时总是手脚乱动不安或浑身不适
	C5	This child often is ‘on the go’ or often acts as if ‘driven by a motor’	常表现的忙个不停或高谈阔论
	C6	This child often interrupts or intrudes on others (e.g. butts into conversations or games)	平时总爱打断别人或者抢话说
其他	D1	Does your child occasionally wet the bed?	睡觉偶尔尿床
	D2	Does your child wake up feeling unrefreshed in the morning	晨起时感觉未得到充分休息
	D3	Is it hard to wake your child up in the morning?	早上很难叫醒
	D4	Did your child stop growing at a normal rate at any time since birth?	曾有过生长发育停滞
	D5	Is your child overweight?	是否超重

function), 并绘制项目特征曲线, 二参数 Logistic 模型表达为:

$$P_i(\theta) = \frac{1}{1 + e^{(-D_{ni}(\theta - b_i))}}$$
 式中, θ 表示受试者的能力 (本研究为是否存在睡眠障碍), $P_i(\theta)$ 表示能力为 θ 的受试者在项目上正确回答的概率 (即如果存在睡眠障碍, 在条目中体

现为睡眠障碍的概率), a_i 表示项目 i 的区分度参数 (即能够区分是否睡眠障碍), b_i 表示项目 i 的难度参数, e 表示自然对数之底, $e=2.718\ 28$, D 为常数, $D=1.7$; 同时还计算了 PSQ 的项目信息函数和测验信息函数, 用来反映该问卷在估计受试者能力所提供的信息量大小, 基于每个条目的项目信息量

函数定义为: $I_i(\theta)=\frac{P'_i(\theta)^2}{P_i(\theta)\times[1-P_i(\theta)]}$, 式中, $P_i(\theta)$ 表示项目 i 的被试反应函数, $P'_i(\theta)$ 表示项目反应函数对 θ 的一阶导函数, 测验信息量函数定义为每个项目信息函数的累积和, 即: $I(\theta)=\sum_{i=1}^m I_i(\theta)$ 。以 $P<0.05$ 为差异有统计学意义。

2 结果

2.1 基本情况 共发放问卷 11 420 份, 回收 10 743 份, 回收率为 94.07%, 经数据清洗后得到有效问卷 9 198 例, 有效回收率为 85.62%。其中男 4 736 例 (51.49%), 女 4 462 例 (48.51%); 年龄 3~14 岁, 平均年龄 (8.8 ± 3.3) 岁; PSQ 总分的中位数为 3 (2, 5) 分; 呼吸、睡眠、行为、其他 4 个维度的得分中位数分别为 0 (0, 1) 分、0 (0, 0) 分、2 (1, 3) 分、0 (0, 1) 分。

2.2 各条目的项目反应函数和项目特征曲线 从难度参数来看, 多数项目的难度参数均 ≥ 1.000 , 只有 C1 (在跟孩子讲话时他/她好像心不在焉或没有在听)、C3 (易被外部的刺激因素分散注意力)、C5 (常表现得忙个不停或高谈阔论)、C6 (平时总爱打断别人或者抢话说) 这 4 个项目的难度参数 <1.000 (分别为 0.843、-0.346、0.436、0.722), 提示这 4 个项目过于容易, 多数人容易“通过”, 即使无睡眠障碍的人也倾向于做出阳性回答。这 4 个条目中回答“是”的比例分别为 30.34%、57.74%、39.89%、33.38%, 远高于其他条目回答“是”的比例。从区分度参数来看, D1 (睡觉偶尔尿床) 和 D5 (是否超重) 的区分度最低 (分别为 0.441 和 0.124), 提示这两个项目难以有效区分睡眠障碍与否 (见表 2)。

项目特征曲线显示, 多数峰值集中在 1~3 之间, 提示这些条目对于倾向于有睡眠障碍的儿童贡献的信息量最多。而 D1 和 D5 这两个条目的项目信息量始终非常低, 即无论被测对象是否具有睡眠障碍的倾向, 这两个条目贡献的信息量均较低 (见图 1)。

2.3 测验信息函数 由于 D1 和 D5 提示效果不佳, 本研究尝试将这个条目删除, 并计算删除前 22 个条目和删除后的 20 个条目的测验信息函数, 剔除这 2 个条目后, 测验信息量几乎没有任何变化, 提示这两个条目可能对睡眠障碍的区分效果不大 (见图 2)。

3 讨论

目前国内对儿童 OSAS 的筛查尚无较好的工具, 美国 2012 年儿童 OSAS 指南^[2]强调 CHERVIN 等^[5]开发的 PSQ 在筛查儿童 OSAS 方面优于其他儿童睡眠问卷。该问卷在美国儿童的研究结果显示, 该问卷的 22 个条目中, 存在 7 项以上症状的儿童患 OSAS 的可能性增加, 其灵敏度和特异度分别为 85% 和 81%^[5]。在此之后, 关于葡萄牙版本和西班牙版本 PSQ 问卷的评估和应用研究相继发表, 其良好的信度、效度和临床筛查能力在不同国家和种族的儿童中得到多次证

表 2 各条目的区分度和难度参数估计结果

Table 2 The estimation of difficulty and discrimination for each item

条目	回答“是”的比例 (%)	难度参数估计			区分度参数估计		
		参数估计值	标准误	P 值	参数估计值	标准误	P 值
A1	2.40	2.675	0.097	<0.001	1.953	0.126	<0.001
A2	9.80	1.910	0.056	<0.001	1.600	0.075	<0.001
A3	5.72	2.164	0.064	<0.001	1.867	0.097	<0.001
A4	14.80	1.609	0.047	<0.001	1.473	0.063	<0.001
A5	2.82	2.349	0.069	<0.001	2.412	0.149	<0.001
A6	2.79	3.081	0.142	<0.001	1.432	0.096	<0.001
A7	9.33	2.938	0.143	<0.001	0.871	0.051	<0.001
A8	34.14	1.220	0.069	<0.001	0.581	0.031	<0.001
A9	6.74	3.808	0.245	<0.001	0.750	0.055	<0.001
B1	5.79	3.681	0.229	<0.001	0.836	0.061	<0.001
B2	3.59	3.485	0.198	<0.001	1.090	0.079	<0.001
C1	30.34	0.843	0.028	<0.001	1.319	0.046	<0.001
C2	19.36	1.773	0.066	<0.001	0.942	0.041	<0.001
C3	57.74	-0.346	0.025	<0.001	1.105	0.040	<0.001
C4	14.91	1.529	0.04	<0.001	1.607	0.063	<0.001
C5	39.89	0.436	0.024	<0.001	1.258	0.044	<0.001
C6	33.38	0.722	0.027	<0.001	1.264	0.045	<0.001
D1	13.44	4.378	0.364	<0.001	0.441	0.039	<0.001
D2	7.09	2.603	0.102	<0.001	1.199	0.063	<0.001
D3	14.39	2.247	0.091	<0.001	0.913	0.044	<0.001
D4	0.17	3.509	0.254	<0.001	2.713	0.426	<0.001
D5	31.80	6.172	1.364	<0.001	0.124	0.027	<0.001

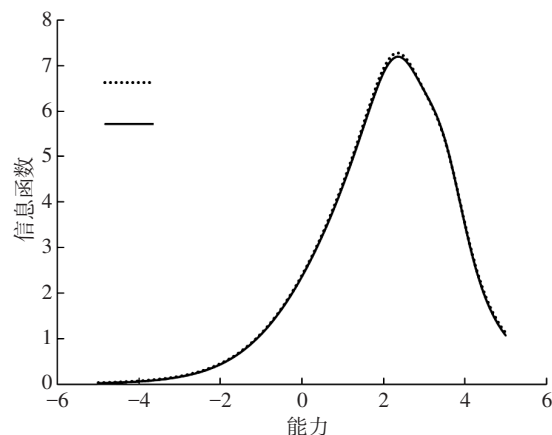


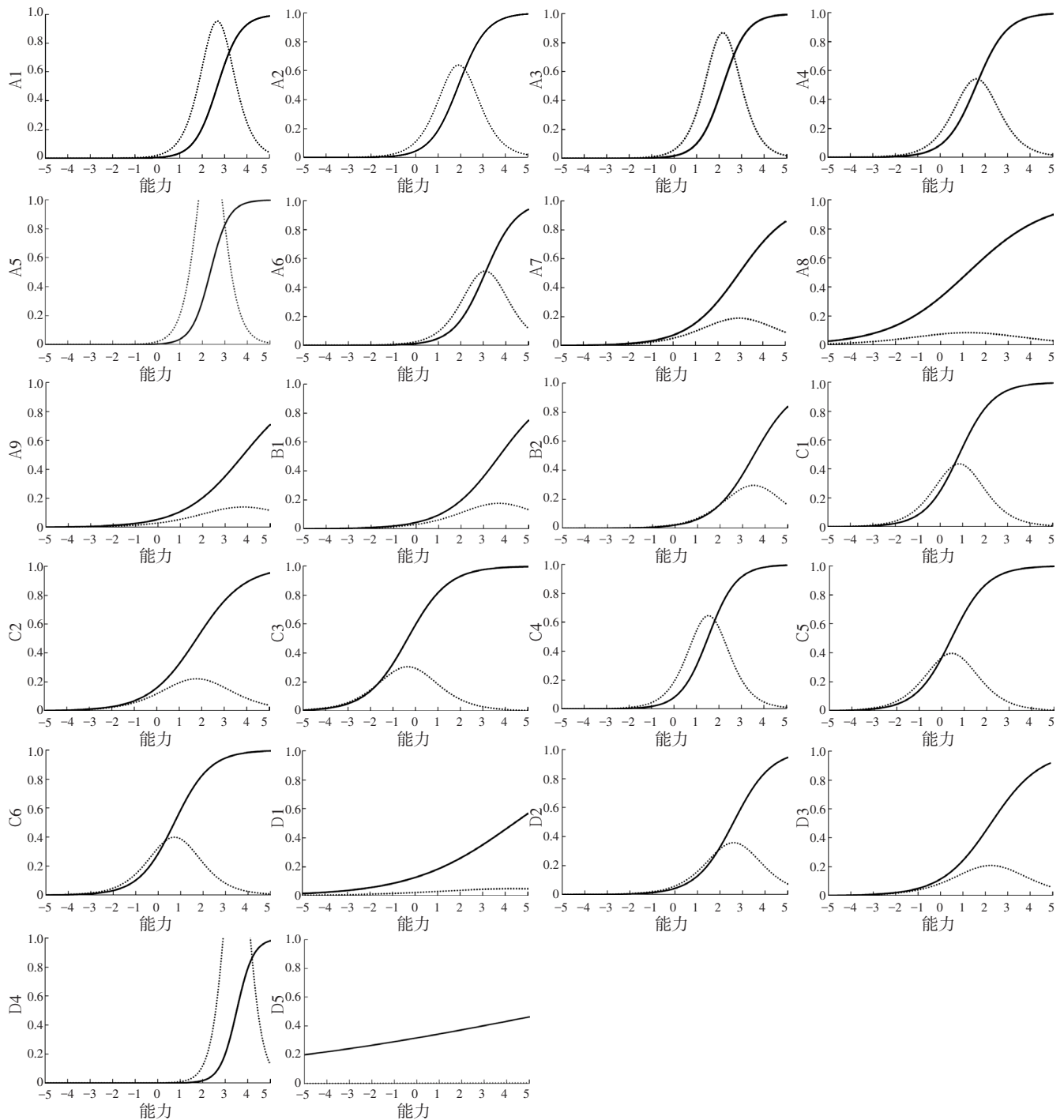
图 2 22 个条目和 20 个条目的测验信息函数比较

Figure 2 The comparison of test information function between 22 items and 20 items

明^[6-8]。而目前我国尚未完全引入 PSQ 问卷, 仅李晓丹等^[4]和 WANG 等^[9]分别对 389 例和 173 例儿童对该问卷进行了评价, 结果认为该问卷具有较好的信度和效度。

本研究在大规模人群的基础上, 采用 IRT 对该问卷进行了验证, 通过对 9 198 例儿童的分析, 认为 PSQ 问卷中多数条目具有较好的信度, 而且综合来看其评价效果较好。测验信息函数呈“瘦高型”曲线, 提示问卷具有较高的区分度; 而且最高峰值 >1 , 说明对于可能存在睡眠障碍的受试者, 该问卷可以贡献最高的信息量。

但本研究也发现问卷中个别条目可能需要调整。正常情况下, 区分度参数取值范围为 (0, 3) 之间, 而难度参数在 (-3, 3) 之间较为合理^[10-11]。综合这两个参数可以发现, “睡



注：横坐标表示潜在的“能力”（本研究为睡眠障碍的倾向），横坐标 ≤ 0 倾向于无睡眠障碍， >1 倾向于有睡眠障碍；纵坐标为回答该项目的概率。虚线为项目信息函数，反映了每一条目所提供的信息量；实线为项目特征曲线，其斜率反映了各条目的区分度

图1 各条目的项目特征曲线和项目信息函数

Figure 1 Characteristic curves and information function for each item

觉偶尔尿床”和“是否超重”这两个条目的区分度参数过低，同时难度过高，项目信息量也未发现随横坐标增加而变化的趋势，而且删除这两个条目后，测验信息量几乎没有任何变化。这些结果均提示，实际中可以考虑简化问卷，删除这两个条目。同时，美国提出的以存在7项以上症状作为PSQ阳性的建议^[12]，也需要同时修改。

IRT是20世纪50年代由REISE等^[13]提出并发展起来

的一种测量理论，目前主要在心理学、教育学等领域应用较多^[14-16]。传统的测量理论存在一定的局限：首先，其测量误差受多方面因素的影响，不同被测对象甚至同一对象的不同时间测量有时差别较大；而IRT则更为稳定，因为其不是计算具体的信度值，而是通过项目特征曲线来反映被测对象的能力（如是否具有睡眠障碍）与某一项目回答的关系。A1条目随着横坐标的增加，纵坐标的值迅速升高，说明被测对象

越倾向于存在睡眠障碍,其选择“是”的概率越高;而D5条目这种趋势则不明显。其次,传统测量理论主要是针对量表的总体特征,如Cronbach's α 系数评价的是整个量表的信度;而IRT则关注构成量表的每一条目。项目特征曲线给出了每一条目的项目信息量,其含义类似于传统测量理论中的信度。最后,IRT对不同能力受试者的灵敏度更高。从测验信息函数来看,信息函数最高峰值对应的横坐标约为2.5。提示,对于倾向于睡眠障碍的受试者,此时该问卷提供的信息函数最高;如果受试者倾向于不存在睡眠障碍,此时该问卷贡献的信息函数越来越低。

目前IRT中比较常用的项目特征曲线函数有Rasch模型、Logistic模型等,而且根据不同的需求,可采用单参数、二参数或三参数模型。本研究采用的是基于区分度和难度两个参数的二参数Logistic模型,而不是基于区分度参数、难度参数、猜测参数的三参数模型。这主要是因为睡眠障碍是一种真实存在的症状,正常情况下不会受随机作答的影响,因此不考虑猜测参数,仅估计区分度参数和难度参数。

总之,本研究通过大规模人群的调查,认为PSQ在评价儿童睡眠障碍中可以提供较好的初筛效果,个别条目可删除,国内应用时可以考虑对其进行简化,重新设定适合国内儿童的临界值。

作者贡献:邵隽进行文章的构思与设计,结果的分析与解释,撰写论文;邵隽、许志飞、马军、胡佩瑾进行研究的实施与可行性分析;李晓丹、杜江南、王桂香进行数据收集;李晓丹、杜江南、郑莉进行数据整理;阎小妍、冯国双进行统计学处理;陈俊、冯国双进行论文的修订;邵隽、张杰、张亚梅、刘原虎、赵靖负责文章的质量控制及审校;倪鑫对文章整体负责,监督管理。

本文无利益冲突。

参考文献

- [1] SHEN Y, XU Z, SHEN K. Urinary leukotriene E4, obesity, and adenotonsillar hypertrophy in Chinese children with sleep disordered breathing [J]. *Sleep*, 2011, 34 (8): 1135-1141. DOI: 10.5665/SLEEP.1178.
- [2] MARCUS C L, BROOKS L J, DRAPER K A, et al. Diagnosis and management of childhood obstructive sleep apnea syndrome [J]. *Pediatrics*, 2012, 130 (3): e714-755.
- [3] LI A M, SO H K, AU C T, et al. Epidemiology of obstructive sleep apnoea syndrome in Chinese children: a two-phase community study [J]. *Thorax*, 2010, 65 (11): 991-997. DOI: 10.1136/thx.2010.134858.
- [4] 李晓丹, 邵隽, 许志飞, 等. 简体中文版儿童睡眠问卷应用于北京地区阻塞性睡眠呼吸暂停综合征儿童筛查的信度和效度评估 [J]. *中华耳鼻咽喉头颈外科杂志*, 2016, 51 (11): 812-818.
LI X D, TAI J, XU Z F, et al. The validity and reliability of simplified Chinese version of the Pediatric Sleep Questionnaire for screening children with obstructive sleep apnea syndrome in Beijing [J]. *Chinese Journal of Otorhinolaryngology Head and Neck Surgery*, 2016, 51 (11): 812-818.
- [5] CHERVIN R D, HEDGER K, DILLON J E, et al. Pediatric Sleep Questionnaire (PSQ): validity and reliability of scales for sleep-disordered breathing, snoring, sleepiness, and behavioral problems [J]. *Sleep Med*, 2000, 1 (1): 21-32.
- [6] TOMAS VILA M, MIRALLES TORRES A, BESELER SOTO B. Spanish version of the Pediatric Sleep Questionnaire (PSQ). A useful instrument in investigation of sleep disturbances in childhood. Reliability analysis [J]. *An Pediatr (Barc)*, 2007, 66 (2): 121-128.
- [7] BERTRAN K, MESA T, ROSSO K, et al. Diagnostic accuracy of the Spanish version of the Pediatric Sleep Questionnaire for screening of obstructive sleep apnea in habitually snoring children [J]. *Sleep Med*, 2015, 16 (5): 631-636.
- [8] CERTAL V, DE LIMA F F, WINCK J C, et al. Translation and cross-cultural adaptation of the Pediatric Sleep Questionnaire into Portuguese language [J]. *Int J Pediatr Otorhinolaryngol*, 2015, 79 (2): 175-178. DOI: 10.1016/j.ijporl.2014.12.002.
- [9] WANG C H, YANG C M, HUANG Y S. The validation and reliability of Chinese version of the Pediatric Sleep Questionnaire for patients with sleep breathing problem [J]. *Taiwanese Journal of Psychiatry (Taipei)*, 2012, 3 (26): 177-186.
- [10] 罗照盛. 项目反应理论基础 [M]. 北京: 北京师范大学出版集团, 2012.
LUO Z S. Item response theory [M]. Beijing: Beijing Normal University Publishing Group, 2012.
- [11] 丁树良, 罗芬, 涂东波, 等. 项目反应理论新进展专题研究 [M]. 北京: 北京师范大学出版集团, 2012.
DING S L, LUO F, TU D B, et al. Some topics of advanced development in item response theory [M]. Beijing: Beijing Normal University Publishing Group, 2012.
- [12] CHERVIN R D, HEDGER K, DILLON J E, et al. Pediatric Sleep Questionnaire (PSQ): validity and reliability of scales for sleep-disordered breathing, snoring, sleepiness, and behavioral problems [J]. *Sleep Med*, 2000, 1 (1): 21-32.
- [13] REISE S P, RODRIGUEZ A. Item response theory and the measurement of psychiatric constructs: some empirical and conceptual issues and challenges [J]. *Psychol Med*, 2016, 46 (10): 2025-2039. DOI: 10.1017/S0033291716000520.
- [14] BUATOIS S, RETOUT S, FREY N, et al. Item response theory as an efficient tool to describe a heterogeneous clinical rating scale in de novo idiopathic parkinson's disease patients [J]. *Pharm Res*, 2017, 34 (10): 2109-2118.
- [15] 钟铁, 季晓辉. 基于经典测量理论和项目反应理论的试卷质量分析方法的比较 [J]. *中国医学教育杂志*, 2015, 35 (1): 155-160. DOI: 10.3760/cma.j.issn.1673-677X.2015.01.047.
ZHONG Y, JI X H. Compare the similarities and differences between CTT and IRT in the quality analysis of the test paper [J]. *Chinese Journal of Medical Education*, 2015, 35 (1): 155-160. DOI: 10.3760/cma.j.issn.1673-677X.2015.01.047.
- [16] 陈新林, 古模发, 童琴, 等. 项目反应理论在SF-36心理测量学特征评价中的应用 [J]. *中国慢性病预防与控制*, 2010, 18 (5): 444-447.
CHEN X L, GU M F, TONG Q, et al. Application of item response theory to the characteristics evaluation of SF-36 scale [J]. *Chinese Journal of Prevention and Control of Chronic Non-communicable Diseases*, 2010, 18 (5): 444-447.

(收稿日期: 2018-06-01; 修回日期: 2018-06-26)

(本文编辑: 陈素芳)