

doi: 10.3969/j.issn.1672-4933.2025.03.013

基于项目反应理论的汉语版西方失语症成套测试初探

Preliminary Investigation to Western Aphasia Battery-Chinese: Based on Item Response Theory

高婧 林枫

GAO Jing, LIN Feng

【摘要】目的 开发基于项目反应理论的汉语失语症简版评估工具,验证其心理测量学特性。**方法** 纳入116例受试者,从汉语版西方失语症成套测试精选100项测试任务,运用摩肯(Mokken)量表分析构建非参数型模型,筛选符合参数型项目反应理论建模的题目,并建立四参数(区分度、难度、猜测度和谨慎度)模型。通过模型拟合比较选出最优模型,得到语言能力评分转换表。**结果** 最终由66项题目构建四参数逻辑斯蒂模型和简版言听测试。该模型具有良好的内部一致性信度(Cronbach's $\alpha=0.971$)和高潜在类别信度(LCRC=0.981)。模型估算的语言功能水平(θ 值)与原始分数显著相关($r=0.85, P<0.001$),且与失语商呈显著相关($r=0.39, P<0.001$)。人项图实现了治疗靶点可视化定位,为制订个性化干预方案提供依据。**结论** 简版言听测试具有良好的心理测量学特征,可用于失语症临床评估。四参数项目反应模型实现了评估治疗衔接的精准对接。

【关键词】 失语症;摩肯量表分析;西方失语症成套测试;项目反应理论

【中图分类号】 G766

【文献标识码】 A

【文章编号】 1672-4933(2025)03-0281-06

【Abstract】 Objective To develop a shortened WAB-C version (PNAT) using IRT for efficient aphasia assessment. **Methods** We selected 100 items from WAB-C (yes/no question, word recognition, and object naming) and then applied Mokken scale analysis to building a Monotone homogeneity model (MHM). Optimized 4-parametric IRT model was selected through model-data fitting and model-model comparison. **Results** The sample contained 116 individuals. The optimum pIRM was a 4-parametric logistic model (4PLM) composed of 66 items. The dichotomous 66-item form had good reliability with Cronbach's $\alpha=0.9707$ and latent class reliability coefficient (LCRC)=0.9817. The values of person ability (θ) were estimated. Clinical insights from the item-person mapping were discussed with a specific case. **Conclusion** The 66-item shortened WAB-C offers a new option for the efficient application in clinical scenarios. The IRT-driven 4PLM bridges evaluation to personalize therapy planning.

【Key words】 Aphasia; Mokken scale analysis; Western aphasia battery; Item response theory

1 引言

失语症是获得性语言障碍,表现为语言符号理解与运用受损。临床评估工具主要采用双维度架构:(1)综合性量表,如汉语版西方失语症成套测试(western aphasia battery-Chinese, WAB-C)^[1,2]、汉语失语症检查法^[3,4]、中国康复医学研究中心汉语标准失语症检查量表^[5],涵盖听、说、读、写语言能力评估;(2)专项评估模块,针对特定语言功能研发,如汉语语法量表^[6]、汉语失写检查法^[7]、功能性沟通能力检查法^[8]等。

西方失语症成套测试(WAB)源于波士顿诊断性失语检查法的改良^[9],英语原版具有良好的信效度^[10],粤语版验证了跨语言适用性^[11]。WAB修订版(western aphasia battery-revised, WAB-R)^[12]现已广泛应用于临床研究^[13],作为其他失语症评定量表的效标^[15,14-16]。WAB-C在部分

子测试中题目区分效度不足^[17],但该研究样本量小,可能存在抽样误差^[18]。现行评估体系存在双重困境:(1)时间成本,WAB测试约需1小时^[19],导致患者配合度下降;(2)治疗衔接,现有工具难以量化患者潜在语言能力阈值,影响阶梯式康复方案的制订。

项目反应理论(item response theory, IRT)较经典测试理论优势(classic test theory, CTT)在于可获得个人潜在特质(latent trait)^[20,21]。本研究创新性地构建了WAB-C精简版言语评估模块(parsimonious naming and auditory comprehension test based on WAB-C, PNAT),基于IRT框架精选听理解(是否判断/听词辨认)与命名(物品命名)核心题目,旨在优化评估效率的基础上,生成个体化能力剖面(θ 值),并定位治疗靶点(题目难度-能力梯度匹配)。

基金项目:国家自然科学基金资助项目(81672255);南京医科大学科技发展基金(NMUB20210350)

作者单位:南京医科大学附属逸夫医院 南京 211100

作者简介:高婧 硕士 主管技师;研究方向:失语症评估与治疗

通讯作者:林枫, E-mail: peterduus@njmu.edu.cn

2 方法

2.1 研究对象

本研究遵循最大变异(maximum variation sampling)原则^[22],选取2017年4月~2022年12月就诊于南京医科大学附属逸夫医院康复医学科接受WAB-C测试的受试者121例。采集受试者姓名、性别、出生年月、受教育程度、脑卒中类型、失语症类型、失语商(aphasia quotient, AQ)、简易精神状态检查量表(mini-mental status exam, MMSE)、蒙特利尔认知评估量表(Montreal cognitive assessment, MoCA)、代币测验(token test)得分等信息。5位受试者因人口学信息缺失而剔除,纳入样本116例。人口学数据均作为首次评估基线,见表1。

2.2 研究方法

选取WAB-C听理解任务(是/否问题和听词辨认)与物品命名任务构成100道测试题。是/否问题可反映失语症患者功能性交流能力,且题目涉及个人信息、即时环境及患者已习得信息^[23]。听词辨认与物品命名分别评估患者听理解与口语表达水平,与治疗紧密相关。评分方法:(1)是/否问题,正确回答记3分,反复提问后回答正确记2分,错误回答记0分;(2)听词辨认,正确回答指认记1分,错误指认或无反应记0分;(3)物品命名,正确回答记3分,音素错误2分,经过提示回答正确记1分,错误回答或无反应记0分。

2.3 IRT建模^[24-29]

数据预处理:①将原始得分进行二值化转化,0表示错误或无反应,1表示正确,得分越高,患者能力越强。②剔除缺失值比率>30%的题目,并填补缺失值^[30]。非参数型建模与参数型模型参考已发表综述^[30],进行模型验证。信度评估采用Cronbach's α 、Guttman's λ 、Molenaar Sijtsma 统计量、潜在类别信度系数(latent category reliability coefficient, LCRC)。效度评估通过比较最优模型估算的 θ 值与非二值化原始总分、AQ值相关性。测试功能差异(differential test functioning, DTF)分析,确保模型的个人能

力估计不受测次影响,避免产生偏倚。模型稳定测试采用蒙特卡洛模拟,以检验采样过程对项目难度参数的影响。

2.4 统计分析

采用R软件4.2.0进行数据处理和统计分析,以 $\alpha=0.05$ 为检验水准。非参数IRT建模采用mokken包,参数型IRT分析和蒙特卡洛模拟采用mirt包。蒙特卡洛模拟共执行1000次,每次模拟5000个复本。

3 结果

3.1 非参数型IRT建模:摩肯量表分析

100条测试题目经摩肯量表分析(mokken scale analysis, MSA)的自动条目筛选程序(automatic item selection procedure, AISP)精简为66条,涵盖是否问题2项、听词辨认48项和物品命名16项。其中听词辨认涵盖原WAB-C的10种类型:实物指认(6)、图片指认(5)、图形指认(3)、汉字指认(5)、数字指认(3)、颜色指认(3)、家具指认(5)、身体部位指认(6)、手指指认(5)和左右分辨(7)。66个测试题目的整体同质化系数 $H=0.496$,条目间同质化系数 $H_{ij}>0$ ^[31],局部独立性检验的总体违例为4.98%(符合局部独立性假设)^[32],单调性检验未发现显著违例。66个测试项目构成单调同质模型(monotone homogeneity model, MHM)模型,该模型内部一致性较好:Cronbach's $\alpha=0.9715$ 、Guttman's $\lambda=0.9730$ 、Molenaar Sijtsma 统计量=0.9747、LCRC=0.9826。模型总分与WAB-C原始分数显著相关($P<0.001, r=0.93$)。

3.2 参数型IRT建模:逻辑斯蒂模型分析

在MHM模型基础上析出四参数逻辑斯蒂模型(4 parameter logistic model, 4PLM)。该模型的拟合度指标:近似均方根误差估计值(root mean square error of approximation, RMSEA)=0.047、比较拟合指数(comparative fit index, CFI)=0.985、Tucker-Lewis 指数(TLI)=0.986,提示模型适应性好。项目参数拟合度判断发现无拟合度不达标条目,见表2。模型信度指标好:Cronbach's $\alpha=$

表1 人口学数据分布(n=116)

项目	结果	项目	结果[n(%)]
年龄(岁)	52.00 [18.88; 73.12]	失语症类型	
受教育年限(年)	10.00 (0.00; 21.85)	完全性失语	10 (8.62)
MMSE(分)	17.51 (6.14)	经皮质混合性失语	6 (5.17)
MoCA(分)	11.00 [5.00; 27.25]	Broca失语	28 (24.14)
Token(分)	16.38 (7.96)	经皮质运动性失语	16 (13.79)
AQ(分)	61.80 [6.25; 95.55]	Wernicke失语	6 (5.17)
脑卒中类型:n(%)		经皮质感觉性失语	8 (6.90)
出血性卒中	38 (32.76)	传导性失语	6 (5.17)
缺血性卒中	49 (42.24)	命名性失语	30 (25.86)
脑外伤	11 (9.48)	健康组	6 (5.17)
其他类型	18 (15.52)	流畅度分型	
		流畅型	56 (48.28)
		非流畅型	60 (51.72)

表2 项目参数估计

条目编号	条目	难度	区分度	猜测度	谨慎度	χ^2	RMSEA	P
a59	左肘_LAR	0.896	21.898	0.146	1.000	47.579	0.040	0.022*
a58	右腕_LAR	0.661	5.865	0.100	1.000	49.693	0.044	0.827
a56	左膝_LAR	0.371	2.031	0.111	1.000	52.828	0.050	0.000*
a50	无名指_FNG	0.342	0.979	<0.001	0.999	36.068	0.000	0.000*
a60	右眼_LAR	0.231	3.628	0.149	1.000	39.717	0.017	0.037*
a57	左踝_LAR	0.196	1.304	<0.001	1.000	39.914	0.018	0.000*
a19	一_CHC	0.066	31.697	0.379	1.000	33.629	0.010	0.121
a21	小_CHC	0.065	42.994	0.408	0.983	44.820	0.034	0.002*
a54	右耳_LAR	0.059	2.191	0.136	1.000	66.875	0.070	0.043*
a20	人_CHC	0.034	40.517	0.394	1.000	30.816	0.000	0.015*
a55	右肩_LAR	-0.003	1.687	0.001	1.000	33.395	0.000	0.007*
a26	数字61_DGT	-0.166	7.538	0.447	1.000	38.577	0.010	0.028*
a51	食指_FNG	-0.192	1.191	<0.001	0.999	49.492	0.044	0.080
a38	椅子_FRT	-0.193	24.513	0.304	1.000	37.920	0.000	<0.001*
a34	绿色_CLR	-0.199	3.109	0.235	0.948	67.744	0.071	0.999
a29	数字32_DGT	-0.211	24.835	0.528	0.974	51.522	0.048	0.040*
a24	不_CHC	-0.219	6.090	0.152	0.957	60.867	0.062	0.999
a47	脖子_BDP	-0.242	1.587	0.075	0.999	47.336	0.040	0.028*
a46	胸口_BDP	-0.271	1.995	0.182	0.999	42.601	0.028	0.392
a18	圆柱体_ICN	-0.277	2.518	0.084	1.000	56.756	0.056	0.134
a33	红色_CLR	-0.327	25.380	0.473	0.988	43.892	0.032	0.022*
a14	三角形_ICN	-0.346	3.760	0.280	1.000	30.476	0.000	0.103
a36	黑色_CLR	-0.346	3.270	0.359	1.000	34.838	0.000	0.068
a53	中指_FNG	-0.354	1.882	<0.001	1.000	33.367	0.000	0.350
a48	下巴_BDP	-0.427	1.834	<0.001	1.000	37.547	0.000	0.015*
a41	门_FRT	-0.428	14.041	0.346	1.000	43.573	0.031	NA
a15	圆形_ICN	-0.545	1.626	<0.001	0.999	40.713	0.021	0.014*
a25	数字5_DGT	-0.632	8.1600	0.585	0.999	41.055	0.023	0.046*
a22	口_CHC	-0.654	1.758	<0.001	1.000	73.911	0.078	0.089
n18	透明胶带_THO	-0.702	15.930	0.000	0.749	47.393	0.040	0.210
a39	写字台_FRT	-0.741	3.124	<0.001	1.000	21.293	0.000	0.022*
n16	橡皮筋_THO	-0.787	34.737	0.000	0.747	973.309	0.397	0.116
n9	门锁_SHO	-0.798	36.234	0.000	0.868	746.923	0.346	0.026*
n14	香烟_SHO	-0.806	34.678	0.000	0.825	3697.691	0.786	0.021*
n11	螺丝刀_SHO	-0.810	8.616	0.000	0.829	30.260	0.000	0.214
n20	火柴_THO	-0.825	40.320	0.000	0.933	3531.953	0.768	0.002*
n8	橡皮_SHO	-0.826	39.948	0.000	0.840	7559.426	1.126	0.035*
a40	电灯/灯管_FRT	-0.830	3.364	<0.001	0.987	42.647	0.028	<0.001*
n7	牙刷_SHO	-0.843	44.794	0.000	0.916	4045.331	0.822	0.015*
a45	眼睛_BDP	-0.845	2.799	0.000	1.000	40.487	0.021	0.015*
n12	钥匙_SHO	-0.862	18.460	0.000	0.963	146.070	0.135	0.103*
a44	鼻子_BDP	-0.863	2.938	0.147	1.000	30.427	0.000	0.021*
a43	耳朵_BDP	-0.872	2.41	0.368	1.000	37.522	0.000	0.002*
n6	锤子_SHO	-0.884	12.004	0.000	0.920	67.651	0.071	0.248
n17	汤勺_SHO	-0.886	56.578	0.000	0.919	3034.456	0.711	0.031*
n19	筷子_SHO	-0.900	63.476	0.000	0.976	3648.371	0.780	NA
a52	小指_FNG	-0.902	1.988	<0.001	1.000	28.275	0.000	0.031*
n2	球_THO	-0.904	18.800	0.000	0.966	138.002	0.130	0.018*
n1	手枪_SHO	-0.918	2.766	0.000	1.000	38.244	0.006	0.113
n10	铅笔_SHO	-0.946	47.479	0.000	0.984	373.660	0.238	NA
a37	窗户_FRT	-0.988	1.960	<0.001	1.000	29.602	0.000	<0.001*
a7	火柴_PIC	-1.172	2.050	<0.001	1.000	41.482	0.024	0.031*
a49	拇指_FNG	-1.402	1.198	<0.001	0.999	50.298	0.046	0.068
a8	茶杯_PIC	-1.529	1.883	<0.001	1.000	37.316	0.000	0.048*
a2	火柴_OBJ	-1.549	2.226	<0.001	1.000	48.131	0.041	0.002*
a4	花_OBJ	-1.566	2.477	0.232	1.000	27.900	0.000	NA
a9	梳子_PIC	-1.630	1.73	<0.001	1.000	39.372	0.015	0.002*
a6	螺丝刀_OBJ	-1.715	1.552	0.002	1.000	61.588	0.063	<0.001*

注：a: auditory comprehension, 听理解检查; n: naming, 命名检查; LAR: left and right, 左右指认; FNG: finger, 手指指认; CHC: character, 汉字指认; DGT: digital, 数字指认; FRT: furniture, 家具指认; CLR: color, 颜色指认; BDP: body part, 身体部位指认; ICN: iconic, 图形指认; PIC: picture, 图片指认; OBJ: object, 实物指认; THO: two-hand operation, 双手操作; SHO: single-hand operation, 单手操作, *P<0.05

0.971、Guttman's $\lambda=0.972$ 、Molenaar Sijtsma 统计量 = 0.974、LCRC=0.982。模型估算的个人能力 θ 值与AQ值 ($P<0.001, r=0.39$)以及原始得分 ($P<0.001, r=0.85$)均显著相关。

本研究部分受试者接受WAB-C复测,图1显示两次检测的预期分数曲线之间存在重叠,所建模型在不同测试时间点上具有一致性,表明模型能跨时间点为同组患者提供稳定的评估结果。

采用蒙特卡洛模拟获得项目难度预期值。模型的实测项目难度值与预期难度值的Kendall秩有显著相关 ($P=4.77 \times 10^{-27}, r=0.98$),表明采样过程对项目难度无显著影响。基于预期值的中位数绘制人项图(图2),显示了个人能力(θ 值)与条目难度(b 值)之间的关系。人项图分为个人能力分布(上部)和条目难度分布(下部),由下到上表

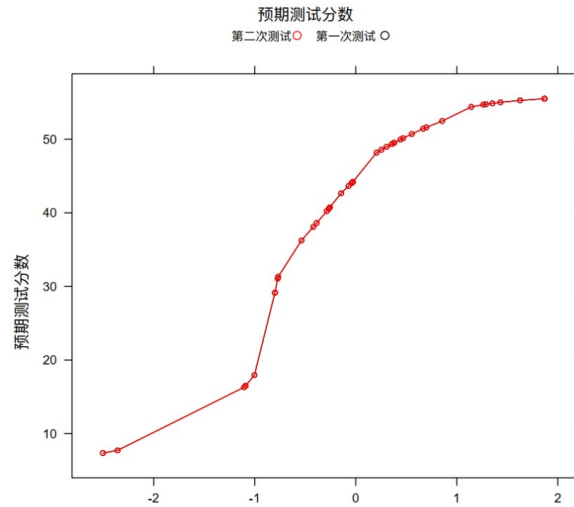


图1 检测秩次的DTF曲线

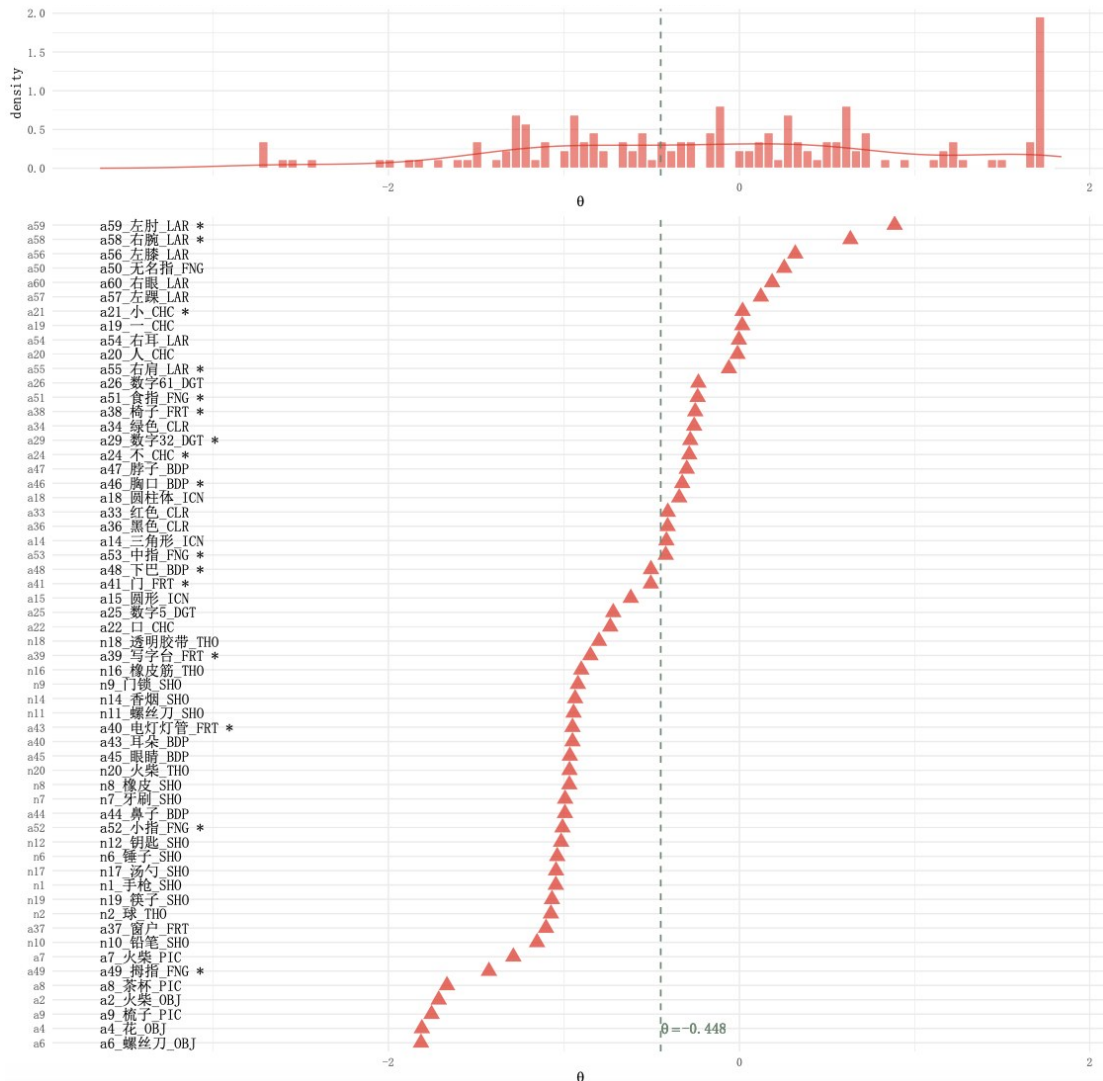


图2 人项图

注:*表示患者错误应答条目

示题目难度递增。以患者Y为例,其能力值为 $\theta = -0.448$ logit(图中垂直虚线)。左侧胜任区内,患者正确应答6项:a49(拇指)、a52(小指)、a40(电灯/灯管)、a39(写字台)、a41(门)、a48(下巴);右侧挑战区内,患者错误应答10项:a53(中指)、a46(胸口)、a24(不)、a29(32)、a38(椅子)、a51(食指)、a55(右肩)、a21(小)、a58(右腕)和a59(左肘)。难度最低的受损条目为a6(螺丝刀),患者能力临界点为a53(中指)。

4 讨论

IRT是一种心理测验方法^[33],强调个体对测试项目的反应受个体能力和题目质量的影响。本研究首次建立失语症评估的双阶段IRT框架,采用非参数型IRT对WAB-C进行条目筛选,随后应用参数型IRT进行个人能力-测试项目探索。MSA已应用于健康相关问卷调查^[34],本研究前期采用MSA分析老年人功能状态^[35]、脑卒中患者功能^[36]、终末期肾病患者血液透析患者活动参与水平^[24]。本研究运用MSA从WAB-C100项中筛选出66条测试题目。这些题目构建的MHM符合IRT基本假设单维性、局部独立性和单调性,且模型信效度验证良好。

Ivanoca等^[37]系统分析20种常用量表,确立标准化开发流程(含Rasch模型)。IRT应用于语言测试始于上世纪80年代初^[38]。IRT逐步应用于失语症评估,涵盖词汇加工^[39-43]、句法^[44]、交流能力^[45]及WAB^[46]。IRT能准确提供反映患者能力的测试题目。

本研究通过参数型IRT量化个人能力 θ 值与试题难度参数 b 的对应关系,并借助人项图实现 θ - b 的空间可视化,提供被试能力值与试题难度值的相对位置^[47,48]。受试者Y为1例经皮质感觉性失语患者,此型特点为流畅性口语,听理解严重障碍,命名障碍,复述能力好,常伴有严重失读和失写^[49]。患者在其他认知检测中表现出定向力减退、复杂图形辨别障碍、复杂指令无法完成、偏盲等症状。人项图揭示患者听理解方面受损,其中左右分辨、汉字指认和数字指认项目难度值超出患者能力值,因此PNAT可反映患者的功能状态。

Tyson等^[50]研究揭示了脑卒中患者及其照顾者对康复评估的认识,发现部分患者不清楚评估目的,对没有得到充分解释的重复测试感到不满。此外,大部分患者希望获得关于康复预后的信息。PNAT可缩短评估时间,可视化的人项图使治疗目标一目了然,有助于临床医生与患者实时了解治疗进展,从而制订精准、有效的康复计划。

本研究存在以下局限性:(1)参数型项目反应理论(parametric item response theory, pIRT)模型与AQ之间相关

性为轻中度, $r_{\text{Pearson}} = 0.39$, $CI_{95\%} [0.25, 0.52]$,这可能表明pIRT模型在预测AQ方面存在局限性。WAB自发言语对AQ贡献值最大(约30%)^[51],而PNAT主要以听理解测试为主,可能影响与AQ的相关性。(2)本研究纳入失语症类型差异较大,且包括健康人群,可能对目标人群结果产生偏倚。此外,本研究未探讨IRT与失语症分型诊断的关系,仅选取WAB-C中部分测试项目,因此无法进行完整的分型诊断。未来研究将增加受试者数量,全面分析WAB-C测试条目,建立全面的评估模型,为开发适合失语症患者的计算机自适应测试奠定基础。

参考文献

- [1] 王荫华. 西方失语症成套测验(WAB)介绍(一)[J]. 中国康复理论与实践, 1997, 3(2): 87-89
- [2] 王荫华. 西方失语症成套测验(WAB)介绍(二)[J]. 中国康复理论与实践, 1997, 3(3): 135-140
- [3] 王新德, 高素荣, 胡超群, 等. 汉语失语症检查法(草案)[J]. 中华神经精神科杂志, 1988, 21(4): 252-253.
- [4] 高素荣, 褚月发, 石舜琴, 等. 汉语失语症检查法标准化的研究[J]. 中国心理卫生杂志, 1992, 6(3): 125-128, 144.
- [5] 张庆苏, 纪树荣, 李胜利, 等. 中国康复研究中心汉语标准失语症检查量表的信度与效度分析[J]. 中国康复理论与实践, 2005, 11(9): 703-705.
- [6] 赵丽丽, 李承晏, 毛善平, 等. 汉语语法量表的制定和标准化[J]. 卒中与神经疾病, 2002, 9(5): 296-298.
- [7] 刘晓加. 汉语失写检查法及其在失写症检查中的临床应用[J]. 第一军医大学学报, 1995, 15(3): 253-255.
- [8] 王小荣, 卓大宏. 功能性语言沟通能力检查法及其在中风失语症检测中的应用-功能性语言沟通能力检查方案的制订[J]. 中国康复医学杂志, 1992, 7(6): 248-250.
- [9] Kertesz A, Poole E. The aphasia quotient: the taxonomic approach to measurement of aphasic disability[J]. The Canadian journal of neurological sciences. Le journal canadien des sciences neurologiques, 2004, 31(2): 175-184.
- [10] Shewam CM, Kertesz A. Reliability and validity characteristics of the Western Aphasia Battery(WAB)[J]. The Journal of speech and hearing disorders, 1980, 45(3): 308-324.
- [11] Yiu Edwin M-L. Linguistic assessment of Chinese-speaking aphasics: Development of a Cantonese aphasia battery[J]. Journal of Neurolinguistics, 1992, 7(4): 379-424.
- [12] Kertesz A. Western Aphasia Battery-Revised[M]. San Antonio, TX: The Psychological Corporation, 2007.
- [13] Kertesz A. The Western Aphasia Battery: a systematic review of research and clinical applications[J]. Aphasiology, 2022, 36(1): 21-50.
- [14] Yang H, Tian S, Flamand-roze C, et al. A Chinese version of the Language Screening Test (CLAST) for early-stage stroke patients[J/OL]. PloS One, 2018, 13(5): e0196646.
- [15] Sun M, Zhan Z, Chen B, et al. Development and application of a Chinese Version of the Language Screening Test (CLAST) in post-stroke patients[J/OL]. Medicine, 2020, 99(37): e22165.
- [16] 谢敬晓, 沈抒, 赵玉华, 等. Token测验在失语症诊断中的作用[J]. 中国

- 康复医学杂志,2004,19(7):509-511.
- [17] 李嘉洋,梁韵瑾,欧阳嘉璐,等. WAB 汉化版失语症诊断量表的信度与效度分析[J/OL]. *Modern Linguistics*, 2022, 10(5):821-821.
- [18] 戴晓阳,曹亦薇. 心理评定量表的编制和修订中存在的一些问题[J]. *中国临床心理学杂志*, 2009, 17(5):562-565.
- [19] 李胜利. 言语治疗学[M]. 第二版. 北京:华夏出版社, 2014:35.
- [20] 李晓铭. 项目反应理论的形成与基本理论假设[J]. *心理发展与教育*, 1989, 1: 25-31.
- [21] 苏昭,刘燕. 项目反应理论研究综述[J]. *江西科技学院学报*, 2013, 8(1): 45-48.
- [22] Palinkas LA, Horwitz SM, Green CA, et al. Purposeful Sampling for Qualitative Data Collection and Analysis in Mixed Method Implementation Research[J]. *Administration and policy in mental health*, 2015, 42(5):533-544.
- [23] Gray L, Hoyt P, Mogil S, et al. A comparison of clinical tests of yes/no questions in aphasia[C]. *Clinical Aphasiology: Proceedings of the Conference*. BRK Publishers, 1977:265-268.
- [24] 徐浪,张芹,江钟立,等. 基于 ICF 和项目反应理论建模评估老年终末期肾病血液透析患者的活动和参与的研究[J]. *中国康复医学杂志*, 38(10):1385-1392.
- [25] Feng C, Geng BF, Liu SG, et al. Activity and participation in haemophiliacs: Item response modelling based on international classification of functioning, disability and health[J/OL]. *Haemophilia: The Official Journal of the World Federation of Hemophilia*, 2023, 29(1):308-316.
- [26] Feng C, Jiang ZL, Sun MX, et al. Simplified Post-stroke Functioning Assessment Based on ICF via Dichotomous Mokken Scale Analysis and Rasch Modeling[J/OL]. *Frontiers in Neurology*, 2022, 13:827247.
- [27] Chen JJ, Zhu ZY, Bian JJ, et al. Nutrition-associated health levels in persons with cancer: item response modelling based on the International Classification of Functioning, Disability and Health[J/OL]. *European Journal of Physical and Rehabilitation Medicine*, 2023, 59(5):593-593.
- [28] Jiang YE, Zhang DM, Jiang ZL, et al. ICF-Based simple scale for children with cerebral palsy: Application of Mokken scale analysis and Rasch modeling - PubMed[J/OL]. *Dev Neurorehabil*, 2023, 26(2):71-88.
- [29] Gikaro JM. Impact of Pain and Pain Related Conditions on Active Living[D]. Nanjing, China: Nanjing Medical University, 2023.
- [30] Feng C, Liu SG, Zhou L, et al. Research Paradigm of the International Classification of Functioning, Disability and Health (ICF) With Item Response Theory: Clarification, Classification, and Challenge[J]. *Physiother Res Int*, 2025, 30(1) :e70021.
- [31] Sijtsma K, Van Der Ark LA. A tutorial on how to do a Mokken scale analysis on your test and questionnaire data[J/OL]. *The British Journal of Mathematical and Statistical Psychology*, 2017, 70(1):137-158.
- [32] Straat JH, Van Der Ark LA, Sijtsma K. Using conditional association to identify locally independent item sets[J/OL]. *Methodology: European Journal of Research Methods for the Behavioral and Social Sciences*, 2016, 12(4):117-123.
- [33] Michael Furr R(美), 李英武, 董妍, 著. 心理测量学=Psychometrics an introduction[M]. 北京:北京师范大学出版社, 2021:351.
- [34] Stochl J, Jones PB, Crudace TJ. Mokken scale analysis of mental health and well-being questionnaire item responses: a non-parametric IRT method in empirical research for applied health researchers[J/OL]. *BMC medical research methodology*, 2012, 12(74):1-16.
- [35] Zhou L, Feng C, Zhong LJ, et al. Laying the Foundation for Developing an Item Bank Measuring Presby-function Based on the International Classification of Functioning, Disability and Health[J/OL]. *Journal of Gerontology and Geriatrics*, 2023, 71(4):245-259.
- [36] Feng C, Jiang ZL, Sun MX, et al. Simplified Post-stroke Functioning Assessment Based on ICF via Dichotomous Mokken Scale Analysis and Rasch Modeling[J/OL]. *Frontiers in Neurology*, 2022, 13:827247.
- [37] Ivanova MV, Hallowell B. A tutorial on aphasia test development in any language: Key substantive and psychometric considerations[J/OL]. *Aphasiology*, 2013, 27(8):891-920.
- [38] 韩江红,刘畅. 项目反应理论在语言测试中的应用研究[J]. *江苏师范大学学报(哲学社会科学版)*, 2018, 44(2):141-146.
- [39] Kendall D, Del Toro C, Nadeau S, et al. The development of a standardized assessment of phonology in aphasia[C/OL]. *Clinical Aphasiology Conference*. Isle of Palms, SC, USA, 2010.
- [40] Del TCM, Bislick LP, Comer M, et al. Development of a Short Form of the Boston Naming Test for Individuals With Aphasia[J/OL]. *Journal of Speech, Language, and Hearing Research*, 2011, 54(4):1089-1100.
- [41] Fergadiotis G, Kellough S, Hula WD. Item Response Theory Modeling of the Philadelphia Naming Test[J/OL]. *Journal of Speech, Language, and Hearing Research*, 2015, 58(3):865-877.
- [42] Alons E, Dijkhuis L, Tuijl P van, et al. Development and Diagnostic Accuracy of a Shortened Dutch Naming Test for People with Aphasia Using Item Response Theory[J/OL]. *Archives of Clinical Neuropsychology*, 2022, 37(8):1735-1748.
- [43] Fergadiotis G, Casilio M, Dickey MW, et al. Item Response Theory Modeling of the Verb Naming Test[J/OL]. *Journal of Speech, Language, and Hearing Research*, 2023, 66(5):1718-1739.
- [44] Milman LH, Dickey MW, Thompson CK. A psychometric analysis of functional category production in English agrammatic narratives[J/OL]. *Brain and Language*, 2008, 105(1):18-31.
- [45] Baylor C, Oelke M, Bamer A, et al. Validating the Communicative Participation Item Bank (CPIB) for use with people with aphasia: an analysis of differential item function (DIF)[J/OL]. *Aphasiology*, 2017, 31(8):861-878.
- [46] Hula W, Donovan NJ, Kendall DL, et al. Item response theory analysis of the Western Aphasia Battery[J/OL]. *Aphasiology*, 2010, 24(11): 1326-1341.
- [47] Lunz ME. Using The Very Useful Wright Map[EB/OL]. *Measurement Research Associates Test Insights*. <https://www.rasch.org/mra/mra-01-10.htm>. 2010.
- [48] 余民宁. 量表编制与发展: Rasch 测量模型的应用[M]. 新北:心理出版社股份有限公司, 2020:9.
- [49] 高素荣. 失语症[M]. 第二版. 北京:北京大学医学出版社, 2006:235.
- [50] Tyson SF, Burton LJ, Megovern A, et al. Service users' views of the assessment process in stroke rehabilitation[J/OL]. *Clinical Rehabilitation*, 2014, 28(8):824-831.
- [51] Ellis C, Peach RK, Rothermich K. Relative Weight Analysis of the Western Aphasia Battery: Aphasiology[J/OL]. *Aphasiology*, 2021, 35(10): 1281-1292.

收稿日期 2024-03-13

责任编辑 薛 静