

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*

徐浪^{1,2} 张芹³ 江钟立³ 王晨光³ 高婧³ 陆鹏² 林枫^{3,4}

摘要

目的:探索老年终末期肾病(end-stage renal disease, ESRD)血液透析(maintenance hemodialysis, MHD)患者随疗程延续而发生的活动和参与功能改变。

方法:纳入48例ESRD-MHD患者(老年32例和中青年16例)进行相隔6个月的先后2次ICF评估。以其中活动和参与成分的20个类目为起点,构建四种项目反应理论(item response theory, IRT)模型备选。从中选择最优模型,计算患者的个人能力水平(θ)、类目难度和区分度。对 θ 进行年龄(被试间因素)和测次(被试内因素)的混合方差分析。同时,结合 θ 对类目难度和区分度的康复意义进行解析。

结果:本研究选取了含12个ICF类目的4参数逻辑斯蒂模型(4PLM)。该模型具有最佳的模型拟合优度($P=0.0657$, M_2 统计量=42.4282)和高信度(潜在类别信度系数=0.9043)。所有条目经标记卡方检验($S-X^2$ 检验)均拟合良好(校正 P 值都 >0.05)。对4PLM估算的 θ 值进行混合方差分析发现,年龄($P=0.020$)和测次($P=0.002$)都有显著效应。在半年随访时,老年组比入组时 θ 值更低($P=0.0106$),而且显著低于同时期中青年组($P=0.0002$)。d9205(社会活动, -0.1708 Logit)和d450(步行, -0.4222 Logit)难度高于半年随访时的老年组平均能力水平(-0.5211 Logit)。

结论:基于ICF的项目反应理论模型,为评估终末期肾病血液透析患者的个人能力水平提供了新的可能。

关键词 国际功能、残疾和健康分类;终末期肾病;血液透析;项目反应理论;康复

中图分类号:R493 **文献标识码:**A **文章编号:**1001-1242(2023)-10-1385-08

ICF-based assessment of activities and participation for aged persons with end-stage renal disease and maintenance hemodialysis via item response theory modeling/XU Lang, ZHANG Qin, JIANG Zhongli, et al./Chinese Journal of Rehabilitation Medicine, 2023, 38(10):1385—1392

Abstract

Objective: To detect the change of functioning level in activities and participation for aged persons with end-stage renal disease (ESRD) and maintenance hemodialysis (MHD).

Method: Forty-eight ESRD-MHD patients ($n=32$ in older group, and 16 in younger group) completed ICF evaluations for the first time and 6-month follow-up. Twenty categories in activities and participation component of ICF were selected for generating 4 candidate IRT models. We calculated the person abilities (θ), item difficulty and discrimination for the model with best goodness-of-fit. The θ was compared via mixed analysis of variance (mixed-ANOVA) with age as between-subject factor and time as within-subject factor. The item difficulty and discrimination were interpreted for their relationships with θ .

Result: Four-parametric logistic model (4PLM) with 12 ICF categories was selected. The 4PLM had best goodness-of-fit ($P=0.0657$, M_2 statistic=42.4282) and high internal consistency (latent category reliability coefficient=0.9043).

DOI:10.3969/j.issn.1001-1242.2023.10.009

*基金项目:南京医科大学附属逸夫医院临床重点科室建设专项(YFZDXK02—7);南京医科大学科技发展基金(NMUB20210353)

1 南京医科大学康复医学院,江苏省南京市,211100; 2 南京医科大学第四附属医院; 3 南京医科大学附属逸夫医院(南京医科大学第三附属医院); 4 通讯作者

第一作者简介:徐浪,男,硕士研究生; 收稿日期:2023-01-28

cient=0.9043). All categories were tested with signed chi-square ($S-X^2$) method and were proved to be with good fitness, i.e., the adjusted P values were all >0.05 . The mixed-ANOVA showed that both age ($P=0.020$) and time ($P=0.002$) had significant effects on θ . The older group had significantly ($P=0.0106$) decreased θ values at the follow-up timepoint in comparison with the initial test. They also had significantly ($P=0.0002$) small θ values in comparison with the values of younger group at the follow-up timepoint. The d9205 (socializing, -0.1708 Logit) and d450 (walking, -0.4222 Logit) are more difficult than the average level of the person abilities in older group.

Conclusion: The Item response theory model based on ICF provides a new possibility for evaluating the individual ability level of hemodialysis patients with end-stage renal disease.

Author's address School of Rehabilitation, Nanjing Medical University, 210031

Key word International Classification of Functioning Disability and Health(ICF); end-stage renal disease; hemodialysis; item response theory; rehabilitation

终末期肾病(end-stage renal disease, ESRD)是各种慢性肾病的终末阶段,表现为不可逆的肾功能衰竭。维持性血液透析(maintenance hemodialysis, MHD)是最常用的替代疗法。虽然MHD可显著降低死亡率,但不能完全替代肾脏功能,患者病情在一定程度上依然不断进展。随着躯体功能的受损,患者的活动和参与也受到影响,生活质量下降^[1]。《国际功能、残疾和健康分类》(international classification of functioning, disability and health, ICF)强调从功能角度评价患者健康状况,为综合分析患者的功能状态提供了规范的术语体系和统一的理论框架^[2]。ICF主要包括身体结构(s)、身体功能(b)、活动和参与(d)、环境因素(e)4个成分。每个成分又细分为不同级别的功能类目(categories)。例如,d4是一级类目“活动(mobility)”,d450是二级类目“步行(walking)”。ICF经过多年发展,已经有1495个类目^[3]。除少量不适用的情况(如女性的条目不适用于男性)以外,每个人都应当具备这些类目指称的所有功能,但临床实践中对所有类目进行评估并不现实。目前常用的方法是从中提取有代表性的少数类目构成针对特定情况的“核心组合”(core set)^[4]。刘林等通过文献回顾,首先提取与ESRD-MHD患者有关的ICF类目共92条,进而对患者和医务人员进行访谈,初步提出含52个类目的ICF核心组合^[5]。

ICF用0—4级限定值(qualifier)来描述类目的障碍程度:0表示无障碍($<4\%$)、1为轻度($5\%—24\%$)、2为中度($25\%—49\%$)、3为重度($50\%—95\%$)、4为完全($96\%—100\%$)。由于分数不等距,不能直接相加获得总分来表示功能障碍程度。为了

能够从ICF评估结果中计算一个单独的数值来表示患者的总体功能状态,有大量研究使用了项目反应理论(item response theory, IRT)^[6],例如脑卒中^[7]、脊髓损伤^[8]、小儿脑瘫^[9]、血友病^[10]、ICF通用组合^[11]、康复组合^[12]等。一方面,IRT可以把患者的某种特征(例如活动和参与能力)视为潜在特质(latent trait,用 θ 表示)。这种特征不能直接测量,而需要通过问卷条目(又称“项目”,在ICF评估则称为“类目”)的得分来估算。另一方面,IRT使用统一的单位(Logit)度量个人能力和条目难度,从而在开发ICF评估工具方面提供了3种潜在用途^[7,9-10,13]:①人的能力水平比较,包括不同个体的能力高低或同一个体的能力升降;②条目的难度大小比较,从而分析条目可干预性;③人的能力与条目的难度比较,可以了解条目相对于个人能力的难度高低,从而指导选择适宜干预的功能靶点。

本研究为ESRD-MHD患者构建基于ICF的IRT模型,初步比较老年患者与中青年患者长期透析后的活动和参与功能变化,为建立基于ICF的血液透析功能结局管理提供新思路。

1 资料与方法

1.1 研究对象

本研究选取在血透专病门诊规律就诊的ESRD-MHD患者采集横断面数据和纵向数据。横断面数据来自南京医科大学第三附属医院,采集时间从2019年12月—2021年11月,仅采集1次。纵向数据来自南京医科大学第四附属医院,采集时间从2021年11月—2022年7月,分为入组时和半年后

随访共2次。

纳入标准:①年龄 ≥ 18 岁;②确诊为终末期肾病,肾小球滤过率 $< 15\text{ml/min/1.73m}^2$,已规律行维持性血液透析(3次/周,4—4.5h/次) ≥ 1 年;③病情稳定,意识清晰,能配合采集信息。排除标准:①接受手术尚未愈合者;②精神异常或患有影响认知功能的神经系统疾病(如脑卒中、阿尔茨海默症等);③患有严重的心肺疾病;④严重的失语、失用或痴呆等不能配合完成调查的情况;⑤其他由研究者判断不适于入组者。本研究得到南京医科大学第三附属医院伦理委员会批准(批号:2018-SR-017),所有受试患者均已签署知情同意书。

1.2 研究方法

1.2.1 数据采集:采用课题组前期已制备的终末期肾病血液透析专用病例报告表,包括92项与终末期肾病血液透析相关的ICF类目,含b类目39项,s类目13项,d类目20项,e类目20项^[5]。纵向数据包含人口学数据和肾病相关指标,包括肌酐、血红蛋白、尿素下降率等。横断面数据包含肾病生活质量量表(KDQOL™-36)。KDQOL™-36的前12条来自12项简明健康状况调查表(SF-12),可以计算出躯体健康得分(PCS)和心理健康得分(MCS)。KDQOL™-36的后24条为肾病专用条目,可以计算肾病总分(KSS),其涵盖3个方面:肾病负担(burdens of kidney disease,第13到第16条,肾脏疾病对患者生理、心理及经济等的压力)、肾病表现(symptoms and problems of kidney disease,第17条到第28条,肾脏疾病引起的症状和不适)、肾病影响(effects of kidney disease,第29到第36条,肾脏疾病对日常生活的影响)。由1名经培训的ICF评定专业人员对入组患者进行访谈。b、d和s按照五点计分系统登记限定值,8(未特指)和9(不适用)则编码为缺失值。本报告采用d成分20个类目用于建模。

1.2.2 IRT建模流程:采用课题组已经报道的IRT建模方法^[7,9-10,13],并根据本研究的重复测量设计对主要参数和分析过程进行调整:①数据预处理:剔除缺失值超过5%的条目。限定值0、1、2、3、4二值化转换为1、0、0、0、0,即无障碍赋值为1,有障碍赋值为0。经此转换,得分越高表示患者在活动和参与方面的独立程度或健康水平越高。继而剔除二值化后1

或0占比超过95%的条目,从而控制天花板效应与地板效应。②模型构建与选择:对于重复测量数据,本研究采用堆栈设计(stack design)进行建模^[14]。对于n例被试和m个测次,共有 $n \times m = N$ 条观测。堆栈设计在N条观测基础上建立IRT模型。构建4个逻辑斯蒂模型备选,分别是1参数逻辑斯蒂模型(1-parameter logistic model,1PLM)、2参数逻辑斯蒂模型(2PLM)、3PLM和4PLM。对每个模型进行基于 M_2 统计量的拟合优度检验,并对拟合优良的模型($P > 0.05$)进行基于似然比的模型间比较,选取贝叶斯信息准则(Bayesian information criterion,BIC)取值更小的模型作为最终IRT模型。③类目参数计算和检验:为所选模型计算对应的参数。如选择为4PLM,则需要为每个条目计算4个参数,包括区分度(条目能在多大程度上区分个人能力的高低)、困难度(要使条目得分有多难,在本研究中表示要使条目无障碍有多难)、猜测度(问卷过程中因猜测而得分的可能性有多大)、谨慎度(条目因不慎而未得分的可能性有多大)。④利用标记卡方检验($S-X^2$)来检查各条目的拟合优度, $P > 0.05$ 为拟合优良。计算患者的活动和参与水平(用 θ 值表示,单位Logit)。针对测次进行测度不变性(measurement invariance)检验,排除所选模型因测次而造成 θ 值估算偏倚。⑤模型信效度:信度分析:计算克隆巴赫 α 、古特曼 λ^2 、Molenaar-Sijtsma统计量、潜在类别信度系数。对IRT模型而言,潜在类别信度系数信度优于其他3个信度指标^[15]。效度分析:用所选模型进行量表总分与 θ 值的线性拟合,以此由计算横断面数据总分换算得到 θ 值,并以 θ 值与MCS、PCS和KSS进行Pearson相关分析。⑥模型应用分析:比较老年组与中青年组的水平,以及患者入组时与半年后随访时的功能水平。

1.3 统计学分析

采用R软件(4.2.2版)进行数据处理和统计分析^[16],统计分析均以 $\alpha = 0.05$ 为检验水准。IRT建模采用mirt工具包(1.37.1版)^[17]。人口学数据统计采用compare Groups工具包(4.0版),分类数据展示方式为“频数(占比)”,分组比较采用 χ^2 检验;非正态分布的定量数据展示方式为“中位数[95%置信区间]”,分组比较采用Kuskal-Wallis秩和检验;正态分

布的定量数据展示方式为均数±标准差,分组比较采用 Welch 双样本 *t* 检验。模型估算的 θ 值进行测次(入组时和半年后随访,为被试内因素)和年龄组(中青年组和老年组,为被试间因素)的混合设计方差分析,所用工具包为 Bruce R(0.8.9 版)。

2 结果

2.1 人口学数据

本研究共入组 54 例患者,半年后进行随访。随访期间,共脱落了 6 例患者,包括死亡 2 例,失随访 3 例,拒绝随访 1 例。最终共有 48 例患者参与 2 次评估,其中老年组 36 例(68.19±5.45 岁),中青年组 16 例(46.82±7.27 岁)。入组时(表 1),老年组血磷比中

青年组偏低($P=0.042<0.05$,效应量达到大效应:Cramer's $V=0.66>0.6$),其余指标均无明显组间差异。半年后随访时(表 2),老年组血磷仍偏低($P<0.05$),但效应量可忽略($\eta^2=1.66\times 10^{-3}<0.01$)。血磷情况提示患者血磷控制良好^[18]。老年组透析前肌酐低于中青年组($P<0.05$),效应量达到中等(Cohen's $d=0.6700<0.5$)。这可能与老年人体质成分肌量减少有关^[19]。老年组尿素下降率开始显著低于中青年组($P<0.05$),效应量显示为小效应($\eta^2=0.2920$,介于 0.01—0.05),提示在半年随访时老年组的血液透析充分性低于中青年组^[18]。

2.2 IRT 建模

建模所用重复测量数据总共有 96 条(48 例患

表 1 入组时体质和血液指标

变量	老年组(n=32)	中青年组(n=16)	统计方法	<i>P</i> 值	<i>df</i>	统计量	效应量
性别			χ^2	0.183	1	$\chi^2=1.7727$	Cv=0.1921
女性	12(37.50%)	10(62.50%)					
男性	20(62.50%)	6(37.50%)					
血透年数	3.00[1.00, 16.00]	2.50[1.00, 9.88]	K-W	0.1411	1	H=2.1654	$\eta^2=1.03\times 10^{-3}$
BMI	23.97[18.11, 29.57]	21.43[19.70, 30.37]	K-W	0.1458	1	H=2.1155	$\eta^2=5.69\times 10^{-3}$
干重(kg)	63.52±9.35	59.17±11.72	T	0.2083	24.85	$t=1.2919$	ds=0.4100
增重(kg)	3.17[- 3.13, 7.65]	3.50[- 2.57, 5.52]	K-W	0.2838	1	H=1.1486	$\eta^2=4.21\times 10^{-3}$
增重/干重	0.05[- 0.06, 0.16]	0.06[- 0.04, 0.10]	K-W	0.092	1	H=2.8384	$\eta^2=0.0499$
透析前 Cr(μmol/L)	892.75[566.96, 1171.81]	976.60[618.64, 1492.14]	K-W	0.1009	1	H=2.6905	$\eta^2=0.0256$
透析后 Cr(μmol/L)	288.25[156.57, 504.94]	318.90[206.43, 703.55]	K-W	0.2046	1	H=1.6091	$\eta^2=0.0893$
Hb(g/L)	105.00[68.95, 130.45]	106.50[76.62, 121.38]	K-W	0.7427	1	H=0.1078	$\eta^2=0.2380$
PTH(ng/L)	344.30[44.80, 1036.62]	388.45[40.22, 1019.01]	K-W	0.7363	1	H=0.1134	$\eta^2=4.26\times 10^{-3}$
K(mmol/L)	5.06±0.86	4.98±0.81	T	0.7546	31.78	$t=0.3153$	ds=0.0956
Na(mmol/L)	138.10[132.08, 142.97]	138.20[129.09, 144.19]	K-W	0.6695	1	H=0.1822	$\eta^2=1.43\times 10^{-3}$
Ca(mmol/L)	2.09[1.81, 2.58]	2.14[1.95, 2.69]	K-W	0.3412	1	H=0.9060	$\eta^2=0.0112$
P(mmol/L)	1.70±0.48	2.02±0.51	T	0.0420*	28.25	$t=- 2.1303$	ds=0.6600
URR(%)	0.69±0.07	0.69±0.07	T	0.7281	28.03	$t=- 0.3512$	ds=0.1090

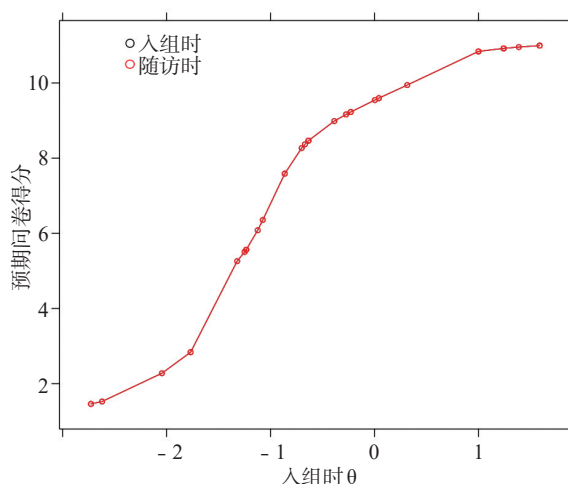
BMI:体质指数;Cr:肌酐;Hb:血红蛋白;PTH:甲状旁腺素;K:血钾;Na:血钠;Ca:血钙;P:血磷;URR:尿素下降率;*df*:自由度;ds:Cohen's *d*,以绝对值表示;Cv:Cramer's *V*。

表 2 随访时体质和血液指标

变量	老年组(n=32)	中青年组(n=16)	统计方法	<i>P</i> 值	<i>df</i>	统计量	效应量
BMI	23.72[18.97, 29.11]	22.00[20.12, 29.08]	K-W	0.2419	1	H=1.3693	$\eta^2=2.28\times 10^{-4}$
干重(kg)	62.17±8.58	59.61±10.88	T	0.4186	24.62	$t=0.8226$	ds=0.2610
增重(kg)	2.90[0.58, 14.69]	3.20[1.01, 4.82]	K-W	0.7262	1	H=0.1226	$\eta^2=2.19\times 10^{-3}$
增重/干重	0.04[0.01, 0.31]	0.05[0.02, 0.09]	K-W	0.4771	1	H=0.5055	$\eta^2=0.0345$
透析前 Cr(μmol/L)	893.88±237.90	1070.59±287.16	T	0.0435*	25.60	$t=- 2.1239$	ds=0.6700
透析后 Cr(μmol/L)	286.20[203.26, 575.24]	310.31[162.49, 693.10]	K-W	0.698	1	H=0.1505	$\eta^2=0.1070$
Hb(g/L)	110.56±17.73	117.31±20.53	T	0.2717	26.50	$t=- 1.1225$	ds=0.3520
PTH(ng/L)	330.85[65.16, 1352.64]	461.90[30.78, 735.95]	K-W	0.4059	1	H=0.6907	$\eta^2=9.11\times 10^{-6}$
K(mmol/L)	4.79[2.81, 6.94]	5.04[4.27, 6.92]	K-W	0.0964	1	H=2.7634	$\eta^2=5.03\times 10^{-3}$
Na(mmol/L)	139.95[131.57, 143.57]	140.20[135.22, 158.84]	K-W	0.8268	1	H=0.0479	$\eta^2=4.47\times 10^{-4}$
Ca(mmol/L)	2.12±0.19	2.20±0.27	T	0.3124	22.71	$t=- 1.0331$	ds=0.3340
P(mmol/L)	1.71[1.02, 2.90]	2.10[1.42, 3.44]	K-W	0.0018*	1	H=9.7885	$\eta^2=1.66\times 10^{-3}$
URR(%)	0.66[0.46, 0.76]	0.71[0.59, 0.78]	K-W	0.0099*	1	H=6.6567	$\eta^2=0.2920$

者,前后2次)观测。在20个d条目中,有8个条目在建模过程中被剔除:d440(精巧手的使用)、d475(驾驶)、d550(吃)、d570(照顾个人的健康)、d630(准备膳食)、d845(得到、保持和终止一份工作)、d850(有报酬的就业)、d9204(业余爱好)。余下12个条目用于构建4个备选模型。其中,仅有4PLM拟合良好($P>0.05$, M_2 统计量=42.4282, $df=30$)^[20]。S-X²检验(表3)显示所有条目均拟合良好($P_{adj}>0.05$)。测度不变性检验显示随访时与入组时的预期问卷得分分布曲线重叠(图1),因而测次无显著的测试功能差异(differential test functioning, DTF),提示本研究采用堆栈法所建模型可对入组和随访时的个人能力实施无偏估计。

图1 不同测次的预期问卷得分分布图



注:随访时与入组时预期问卷得分分布曲线重叠。

2.3 模型信效度分析

4PLM的信度较高:克隆巴赫 $\alpha=0.8312$ 、古特曼 $\lambda^2=0.8459$ 、Molenaar-Sijtsma统计量=0.8634、潜在类别信度系数=0.9043。量表总分(TTS)与 θ 值的线性拟合结果为 $\theta = -1.92 + 0.0732 \times \text{TTS} + 0.0145 \times \text{TTS}^2$, $R^2_{adj}=0.92$ 。以横断面数据100例作为验证集,用换算公式为验证集计算 θ 值,最后用验证集的 θ 值与PCS、MCS和KSS得分进行皮尔逊相关分析。结果发现 θ 值与PCS($r=0.49$, $P=2.09 \times 10^{-7}$)、MCS($r=0.34$, $P=5.61 \times 10^{-4}$)、KSS($r=0.36$, $P=2.57 \times 10^{-4}$)均显著正相关。

2.4 模型应用1:能力值的方差分析

老年组在入组时 $\theta = -0.0251 \pm 0.76671$ Logit,半年随访时 $\theta = -0.5211 \pm 0.7608$ Logit。中青年组入组 $\theta = 0.3361 \pm 0.7060$ Logit,半年随访 $\theta = 0.1269 \pm 0.8582$ Logit。方差分析发现(表4),年龄效应显著($P=0.020<0.05$),测次效应也显著($P=0.002<0.05$),测次与年龄无显著交互效应($P=0.191>0.05$)。简单主效应分析表明(表5),随访时的年龄组效应显著($P=0.0106<0.05$),老年组的测次效应显著($P=0.0002<0.05$)。成对比较表明(表6),半年随访时中青年组 θ 值显著高于老年组($P=0.0106<0.05$),老年组随访时 θ 值显著低于入组时($P=0.0002<0.05$)。因此,老年组的活动和参与水平随透析疗程增加而下降,且能力低于接受同样长度透析疗程的青年组。

2.5 模型应用2:个人能力与条目难度比较举例

以老年组的平均能力水平来看,相对于入组时平均功能水平(-0.0251 Logit)更难的类目仅有

表3 最终所选模型的条目参数和拟合指标

条目	条目名称	难度(Logit)	区分度	猜测度	谨慎度	S-X ² 值	df	RMSEA	P_{adj}	z-outfit	z-infit
d430	举起和搬运物体	2.8189	0.5492	0.0001	0.9984	1.9116	5	0	1	-0.0864	-0.0546
d240	控制应激和其他心理需求	0.4364	2.0459	0.5906	1	7.1771	5	0.0677	1	-0.7794	-0.1454
d9201	(体育)运动*	0.2397	3.5892	0	1	5.5846	5	0.0351	1	-1.1984	-3.501
d9205	社会活动	-0.1708	1.5062	0	1	15.5298	5	0.1489	0.0999	-1.2692	-0.4608
d450	步行	-0.4222	27.1451	0.2287	1	4.6284	5	0	1	-6.3027	-1.0462
d640	做家务	-0.7715	8.9840	0	1	1.4907	5	0	1	-1.9318	-0.4833
d220	从事多项任务	-0.7964	4.8605	0.1175	0.9727	7.0744	5	0.0661	1	0.0114	0.232
d660	帮助别人	-0.8273	13.3966	0	1	5.713	5	0.0387	1	-6.6232	-1.8336
d470	利用交通工具	-1.2928	43.5168	0.2632	0.9391	7.7197	5	0.0757	1	-0.0433	-0.1157
d465	利用设备到处移动	-1.4756	1.9997	0.0746	1	5.0284	5	0.0077	1	-0.0286	0.0378
d520	护理身体各部	-1.6182	3.9824	0	1	12.0674	5	0.122	0.4065	-0.481	-1.0066
d510	盥洗自身	-1.766	3.4217	0	1	12.3278	5	0.1242	0.3667	-0.3655	-0.987

RMSEA:近似均方根误差; P_{adj} :经Bonferroni法矫正后P值;z-outfit:离群敏感型拟合值的标准分数;z-infit:在群敏感型拟合值的标准分数。

注#:d9201原文为“sports”,中文版原译为“运动”,为避免歧义,此处添加“体育”。注*:P值<0.05。表格按照难度降序排列。

表4 混合方差分析

	均方	均方误	$F(1,46)$	P 值	η^2_p [90% CI]
年龄组 (被试间)	5.4319	0.9398	5.7796	0.020*	0.1116[0.0100,0.2664]
测次 (被试内)	2.6531	0.2490	10.6550	0.002*	0.1881[0.0473,0.3507]
交互效应 (年龄组×测次)	0.4388	0.2490	1.7621	0.191	0.0369[0.0000,0.1619]

η^2_p : 偏 η^2 , 注*: P 值<0.05。

表5 简单主效应分析

简单效应	固定水平	$F(1,46)$	P 值	η^2_p [90%置信区间]
年龄组	入组时	2.4910	0.1210	0.0514[0.0000,0.1855]
	随访时	7.1070	0.0106*	0.1338[0.0189,0.2923]
测次	老年组	15.8120	0.0002*	0.2558[0.0934,0.4170]
	中青年组	1.4070	0.2417	0.0297[0.0000,0.1489]

d430、d240 和 d9201, 到随访时功能水平下降到-0.5211Logit, 因此有两个类目在半年后有较高障碍风险: d9205(社会活动)和d450(步行)。图2左图为本研究样本所测算的个人能力 θ 值的分布柱状图。图2右图为类目按照难度从小到大的排列图, 具有相近难度的类目(d220、d640、d660)分布在同一水平线上。左右两图坐标轴对齐而构成怀特图, 从而把患者的功能水平与条目难度以Logit为统一尺度进行对照分析。以个人能力水平线为参照, 高于该线的为该患者不胜任区, 低于该线的则为可胜任区。图2显示了一名患者两次评估所得个人能力水平。

患者W.BH是一名70岁女性患者, 入组时的 θ 值为-0.2917Logit(深虚线), 随访时降至-0.9472Logit(浅虚线)。就该患者随访时的功能水平而言, d450(步行)、d220(从事多项任务)、d640(做家务)和d660(帮助别人)从可胜任区转入不胜任区。

3 讨论

本研究为终末期肾病血液透析患者初步构建了活动和参与的项目反应理论模型。该模型采用了课题组前期已经报道的“以无障碍或非依赖而计分”建模策略^[7,9-10], 即以无障碍为得1分, 有障碍记为0分, 从而对原始的ICF限定值系统进行二

值化转换。经此转换之后获得的IRT模型, 可以提供两方面信息: ①对患者可以测算个人能力水平, 即患者在活动和参与方面的功能水平高低; ②对类目可以测算难度大小, 即要使该类目无障碍或功能独立有多难。IRT模型的原理之一, 就是所用的条目足以构成区分受试者能力水平的量表。这意味着, 即使在建模过程中有一定数量的看似合理的条目被剔除, 所得结果已经可以用于区分患者功能水平高低。被剔除的条目, 并不意味着对患者并不重要, 而是在量表中增加这些条目并不能提高量表对患者能力的区分作用。本研究最终得到一个含有12条类目(表3)的活动和参与能力评估量表。为该量表配置的IRT模型还同时展示了这些类目的参数。难度最大的是d430(举起和搬运物体), 最低的是d510(盥洗自身)。区分度最大的是d470(利用交通工具), 最低的是d430(举起和搬运物体)。分析这些难度和区分度参数, 有助于为临床分析活动和参与水平的影响因素提供依据。值得注意的是, d240(控制应激和其他心理需求)和d220(从事多项任务)的猜测度均较高(分别是0.5906和0.1175), 即有较大可能凭猜测而判断患者具备该功能。这提示可依据猜测度和谨慎度参数对量表条目的评估加以改进。

本研究的主要临床发现在于, ESRD患者在接受MHD时, 活动和参与水平会出现不同程度的下

图2 怀特图和个人能力示例

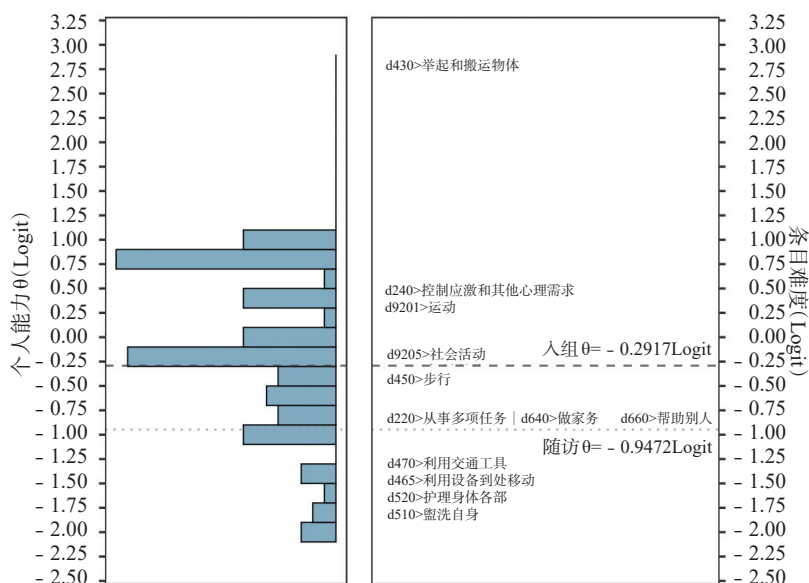


表6 成对比较表

固定水平	组间比较	均差(标准误)	t(46)	P值	Cohen's d[90%置信区间]
入组时	中青年组—老年组	0.3612 (0.2288)	1.5783	0.1214	0.5174 [-0.1425, 1.1772]
随访时	中青年组—老年组	0.6480 (0.2431)	2.6660	0.0106*	0.9282 [0.2274, 1.6290]
老年组	随访—入组	-0.4961 (0.1247)	-3.9765	0.0002*	-0.7105 [-1.0702, -0.3509]
中青年组	随访—入组	-0.2092 (0.1764)	-1.1860	0.2417	-0.2997 [-0.8084, 0.2090]

降,而老年人的这种下降程度有显著性差异(表5和表6)。Morishita等通过文献回顾,发现ESRD-MHD患者会伴随大量症状,包括疲劳、乏力、疼痛、胸闷、焦虑、抑郁、睡眠障碍等,由此出现不同程度的生理功能下降以及心理与社会认知障碍,而躯体活动的减少则与患者死亡率相关^[21]。焦乐等^[22]基于世界卫生组织残疾评定量表2.0(WHODAS2.0)分析了MHD患者的失能状况,发现轻度失能患者占70%以上,中重度失能患者占20%,绝大多数患者都存在社会参与障碍。本研究由此采用ICF对患者的活动和参与进行分析。为了解患者功能水平随时间变化的情况,本研究设置了半年随访。虽然该随访时间间隔为任意设定,但最终结果发现至少在该时间跨度上存在老年人群的显著功能下降。老年患者的活动和参与水平下降有可能与逐步出现的衰弱相关。衰弱是指多个生理系统出现累积性衰退,导致个体的生理储备减少,抗应激能力减弱^[23]。衰弱会增加MHD患者跌倒、住院和死亡等风险,导致生活质量下降^[24]。MDH患者长期血透后很容易出现蛋白质—能量消耗,肌肉蛋白合成率降低、分解代谢增加,出现体质量下降、消瘦,引起骨骼肌进行性消耗^[25]。运动的主要工作器官就是肌肉,随着患者肌肉含量的不断减少,其运动能力势必会不断下降。频繁的透析频次,较长的单次透析时间,严格的限食限水,长期的服药史等等,限制了患者的活动,改变了患者的生活方式,导致患者在社会参与方面不断退缩。以本研究所述W.BH患者为例,第一次评估时个人能力为-0.2917Logit,随访时降到-0.9472Logit。从而使d450(步行)、d220(从事多项任务)、d640(做家务)和d660(帮助别人)等多项的难度高于其随访时的能力水平。

本研究的主要应用临床价值在于,项目反应理论模型可以提供问卷条目的难度和区分度排列表,从而为康复干预提供可量化的决策依据^[7,9-10,13],即根据患者的能力水平来查询在患者可胜任及不胜任

区间内的功能类目,从而解决临床实际操作中必须面对但又往往缺少辅助决策依据的问题:如何选择需要干预的项目,并对多个选项进行排序。换言之,康复需要循序渐进,因而需要解决“(要想恢复)有多难”和“(如果干预)能进步多少”这两个基本问题。对于“有多难”,如图2所示,由于d220、d640和d660是随访时不胜任类目中距离能力线最近的类目,即难度最接近患者能力,康复干预目标可以从这3个类目中选择。对于“能进步多少”,查阅表3可知3者中的区分度最大的是d660(13.3966),其次是d640(8.9840)。在借助怀特图实现康复目标设定方面^[9-10],可能先从改善d660(帮助别人)和d640(做家务)入手,因为在3个难度相近的类目中,干预高分度类目能更大程度地提高整体能力 θ 。当然,这并不意味着直接根据参数来指定干预方案,而是为医患双方就康复方案的可能性和可行性进行充分沟通提供了量化依据。

本研究的主要技术价值在于,为运用现代测量理论探索ICF临床应用提供了思路。与以往评估方式的主要区别在于,本研究不仅评估了个人能力水平,还计算了项目难度,从而使个人能力水平与难度之间可以进行比较。另外,4PLM的猜测度和谨慎度参数还有望成为业务质控指标,即通过累积数据和更新模型参数,有望了解评估者采集和记录信息的质量。在本研究中,d240(控制应激和其他心理需求)的猜测度过高,说明可能有相当部分的患者误判为无障碍,提示该项功能需要采集更多辅助信息,改进评估方式,从而在今后模型更新中降低猜测度。本研究将在后续研究中改进以下局限性:①纳入偏倚和样本量:纳入患者主要来自南京地区,可能存在地区偏倚。样本量偏少,且年龄组分布不均,虽然本报告通过计算效应量来辅助统计推断,但未来仍需要进一步扩大样本量。②IRT模型仅使用d条目,而没有纳入其他成分。今后将纳入更多成分,以更全面反映患者的功能状态。③参数估计有待改进:在

小样本的情况下,虽然可以使用蒙特卡洛模拟或自举法重采样等方法估计参数取值范围,但持续扩增样本仍然是提高研究质量的必然要求。

4 结论

相比中青年患者,同等透析疗程的老年患者的活动参与水平在半年透析后显著下降。采用“以无障碍或非依赖而计分”的建模策略构建项目反应理论模型,有可能为评估患者个人能力水平提供新手段,而且选取难度与能力水平相适应的ICF类目可能有助于进一步探索基于类目难度的康复决策方法。

参考文献

- [1] Wilkinson TJ, McAdams-DeMarco M, Bennett PN, et al. Advances in exercise therapy in predialysis chronic kidney disease, hemodialysis, peritoneal dialysis, and kidney transplantation[J]. *Current Opinion in Nephrology & Hypertension*, 2020, 29(5): 471—479.
- [2] World Health Organization. International Classification of Functioning, Disability and Health: ICF[M]. Geneva: WHO, 2001.
- [3] Karlsson E, Gustafsson J. Validation of the International Classification of Functioning, Disability and Health (ICF) core sets from 2001 to 2019: a scoping review[J]. *Disability and Rehabilitation*, 2022, 44(14): 3736—3748.
- [4] Bickenbach JE, Cieza A, Selb M. ICF core sets: manual for clinical practice[M]. 2nd edition. Boston: Hogrefe Publishing, 2021.
- [5] 刘林, 孔雨柔, 焦乐, 等. 终末期肾病血液透析《国际功能、残疾和健康分类》核心组合的初步研究[J]. *中国康复医学杂志*, 2021, 36(7): 793—798.
- [6] Brzezińska J. Item response theory models in the measurement theory[J]. *Communications in Statistics-Simulation and Computation*, Taylor & Francis, 2020, 49(12): 3299—3313.
- [7] Feng C, Jiang ZL, Sun MX, et al. Simplified post-stroke functioning assessment based on ICF via dichotomous moken scale analysis and rasch modeling[J]. *Frontiers in Neurology*, 2022(13): 827247.
- [8] Li K, Yan T, You L, et al. Psychometric properties of the International Classification of Functioning, Disability and Health set for spinal cord injury nursing based on Rasch analysis[J]. *Disability and Rehabilitation*, 2018, 40(3): 338—345.
- [9] Jiang YE, Zhang DM, Jiang ZL, et al. ICF-based simple scale for children with cerebral palsy: Application of Moken scale analysis and Rasch Modelinig[J]. *Developmental Neurorehabilitation*, 2023.10.1080/17518423.2023.2166614.
- [10] Feng C, Geng B, Liu S, et al. Activity and participation in haemophiliacs: Item response modelling based on international classification of functioning, disability and health [J]. *Haemophilia*, 2022; 10.1111/hae.14702.
- [11] Ehrmann C, Proding B, Stucki G, et al. ICF Generic Set as new standard for the system wide assessment of functioning in China: a multicentre prospective study on metric properties and responsiveness applying item response theory[J]. *BMJ Open*, 2018, 8(12): e021696.
- [12] Gao Y, Yan T, You L, et al. Psychometric properties of the International Classification of Functioning, Disability and Health Rehabilitation Set: a Rasch analysis[J]. *International Journal of Rehabilitation Research*, 2020, 44(2): 144—151.
- [13] Feng C, Lai QL, Ferland A, et al. Mandarin stroke social network scale and item response theory[J]. *Frontiers in Stroke*, 2022, 1: 903289.
- [14] Andrich D, Marais I. A course in Rasch measurement theory measuring in the educational, social and health sciences [M]. Springer Singapore, 2019.
- [15] van der Ark LA, van der Palm DW, Sijtsma K. A Latent class approach to estimating test-score reliability[J]. *Applied Psychological Measurement*, 2011, 35(5): 380—392.
- [16] R Core Team (2022). R: A language and environment for statistical computing[M]. R Foundation for Statistical Computing, Vienna, Austria.URL <https://www.R-project.org/>.
- [17] Chalmers RP. mirt: A multidimensional item response theory package for the R environment[J]. *Journal of Statistical Software*, 2012, 48(6): 1—29.
- [18] 中国医师协会肾脏病医师分会血液透析充分性协作组. 中国血液透析充分性临床实践指南[J]. *中华医学杂志*, 2015, 95(34): 2748—2753.
- [19] Tian R, Chang L, Liu D, et al. Association of the modified creatinine index with muscle strength and mortality in patients undergoing hemodialysis[J]. *Renal Failure*, 44(1): 1732—1742.
- [20] Xu J, Paek I, Xia Y. Investigating the behaviors of M2 and RMSEA2 in fitting a unidimensional model to multidimensional data[J]. *Applied Psychological Measurement*, 2017, 41(8): 632—644.
- [21] Morishita S, Tsubaki A, Shirai N. Physical function was related to mortality in patients with chronic kidney disease and dialysis[J]. *Hemodialysis International*, 2017, 21(4): 483—489.
- [22] 焦乐, 刘林, 孔雨柔, 等. 维持性血液透析患者基于世界卫生组织残疾评定量表2.0的失能状况调查[J]. *中国康复医学杂志*, 2022, 37(5): 642—646.
- [23] Chowdhury R, Peel NM, Krosch M, et al. Frailty and chronic kidney disease: A systematic review[J]. *Archives of Gerontology and Geriatrics*, 2017, 68: 135—142.
- [24] Garcia-Canton C, Rodenas A, Lopez-Aperador C, et al. Frailty in hemodialysis and prediction of poor short-term outcome: mortality, hospitalization and visits to hospital emergency services[J]. *Renal Failure*, 2019, 41(1): 567—575.
- [25] Fouque D, Kalantar-Zadeh K, Kopple J, et al. A proposed nomenclature and diagnostic criteria for protein-energy wasting in acute and chronic kidney disease[J]. *Kidney International*, 2008, 73(4): 391—398.