·临床研究·

基于图模型的脑卒中功能变量关系结构探析*

孙铭雪[1] 江钟立[1,2] 张 芹[2] 吴亚岑[3] 侯 莹[4,5] 刘丽华[6] 刘守国[1] 纪 婕[1,2] 林 枫[1,2,7]

摘要

目的:为国际功能、残疾和健康分类(ICF)在脑卒中康复中的应用提供新思路和新依据。

方法:选取100例脑卒中患者组成的便利样本,访谈法评定国际版脑卒中ICF综合核心组合的166个条目。以条目为节点,条目间的风险相关性为连线,构建图模型。采用R软件(3.2.2版)建模,用Pajek64(5.02版)进行网络分析和可视化。

结果:在图模型总体内,存在主组元结构,并且可以从中提取具有稳定路径结构的3-核结构。主组元中还可以提取具有牢固连接关系的主岛屿结构。其中"d450步行"既属于通用组合,又属于简要核心组合,并且占据了重要位置。简要核心组合条目在总图中的子网络,以及这些条目在各级限定值上的频数分布,可以提供相互参照和相互补充的功能变量知识。

结论:ICF综合核心组合的图模型,可反映脑卒中功能变量间的复杂关系结构。该组合的频数分布,可反映卒中功能变量的属性特征。从关系结构和属性特征两个不同视角进行分析,可以为ICF的康复应用提供综合策略。

关键词 脑卒中;国际功能、残疾和健康分类;图建模;网络分析

中图分类号:R743.3, R492 **文献标识码:**A **文章编号:**1001-1242(2018)-09-1029-07

Exploratory analysis of relational structures among functioning variables in stroke by graphical modeling/ SUN Mingxue, JIANG Zhongli, ZHANG Qin, et al.//Chinese Journal of Rehabilitation Medicine, 2018, 33(9): 1029—1035

Abstract

Objective: To provide new thinking and evidence about applications of International Classification of Functioning, Disability and Health (ICF) for persons with stroke.

Method: A convenient sample of 100 stroke persons participated in this study. Investigation included a set of 166 ICF categories in the comprehensive core set for stroke. Graphical model was composed of categories as nodes and risk correlations as lines. The study used R software (Version 2.2.2) for graphical modeling and Pajek64 (version 5.02) for network analysis and visualization.

Result: There is a main component in the global structure of graphical model. The main component contains a 3-core with robust connectivity and a main island with solid connections. In the main island, a critically localized category, d450 (Walking), belongs to both generic core set and brief core set. A network of brief core set was extracted from the graphical model. For the same set of categories, this study mapped their frequency distribution on ICF scores. By comparing the brief core set network and their frequency distribution, complementary knowledge of functional variables emerged from the findings.

Conclusion: Graphical model of comprehensive core set for stroke can reveal complex relational features embedded in functional variables. Frequency distribution of the same set makes it possible to interpret attribute

DOI:10.3969/j.issn.1001-1242.2018.09.005

*基金项目:江苏省科技支撑计划(BE2012675);国家自然科学基金项目(81672255)

1 南京医科大学第一附属医院(江苏省人民医院),南京,210029; 2 南京医科大学附属逸夫医院; 3 湖南省人民医院; 4 苏州工业园区星海医院; 5 苏州市立医院; 6 江苏句容市人民医院; 7 通讯作者

作者简介:孙铭雪,女,硕士研究生,主治医师; 收稿日期:2017-10-20

features of functional variables. By analyzing different perspectives from relational and property data, these techniques may offer an integrated strategy for the applications of ICF.

Author's address The First Affiliated Hospital of Nanjing Medical University, Nanjing, 210029

Key word stroke; international classification of functioning, disability and health; graph modeling; network analysis

人类生活由大量的功能活动组成。疾病影响功能活动，从而降低生存质量。康复服务要改善功能受损，就必须要干预功能变量。这不仅需要考察单个变量，还要解析变量之间的相互关系，形成针对性的干预方案。《国际功能、残疾和健康分类》(international classification of functioning, disability and health, ICF)为描述功能变量提供了统一的理论框架和标准的术语体系[1—2]。在“身体功能”、“身体结构”、“活动和参与”和“环境和个体因素”四个大类下，ICF通过把功能活动要素化和颗粒化，细分出大量亚类(功能变量)。不同亚类的组合，就形成衡量不同功能状态的变量集。在应用层面，体现为不同疾病的ICF核心组合[3]。

脑卒中由于其高发病率和高致残率，国际ICF研究早在2004年就为其制定了ICF核心组合[4]，此后国际和国内学者不断推进卒中的ICF组合研究[4—6]。但是，传统的ICF组合研究，主要是筛选代表性的功能变量，其隐含假设是组合内的条目相互独立。因而在临床应用ICF评估时，并没有明确的参照依据，来指导考察功能变量间关系。Strobl等[7—8]提出以功能变量为节点，变量间关系为连线，构建网状的功能地形图，用于选择干预靶点、辨识干扰因素和发掘治疗路径。为探寻脑卒中功能变量之间的关系，本研究遵循Strobl等[7]提出的ICF图建模(graphical modeling)理论，借鉴课题组前期在糖尿病等疾病应用的ICF图建模方法[9—10]，为脑卒中166个功能变量建立图模型。通过对所得图模型进行探索性网络分析，为ICF在脑卒中康复中的应用提供新思路和新依据。

1 对象与方法

1.1 研究对象选择

纳入标准[5]：①经头颅CT或磁共振成像(MRI)确诊脑卒中；②脑卒中为主要诊断；③年龄≥18岁；④能够知情同意。

排除标准：①外伤或手术切口尚未愈合者；②病情危重者，如心肺功能衰竭；③ICF条目缺失值比例＞30%；④影响信息采集的其他因素，如精神病史或严重痴呆。

卒中患者便利样本100例，采集于2014年6月—2017年9月，在江苏省人民医院省级机关医院病区、南京医科大学附属逸夫医院、苏州市立医院、苏州工业园区星海医院和江苏句容市人民医院住院的脑卒中患者。男性50例，平均年龄(60.07±13.57)岁；出血性卒中18例，缺血性卒中28例，混合性卒中4例；病程(4.45±7.73)个月，身体质量指数(body mass index, BMI)为(24.26±3.10)；巴氏指数(Barthel index, BI)为(40.29±34.41)分。女性50例，平均(60.81±15.11)岁；出血性卒中14例，缺血性卒中34例，混合性卒中2例；病程(7.02±11.93)个月，BMI为(22.49±3.25)，BI为(40.13±36.46)分。

1.2 原始数据采集

本研究采用扩展版国际综合核心组合166项[5]。其中，身体功能(b)59项，身体结构(s)11项，活动和参与(d)59项，环境和个体因素(e)37项。选用含有18个条目的国际版简要核心组合[4](其中专家提名率＞50%的条目为10项)提取子网络和频数分析。参照课题组已有研究，由专业人员使用访谈法获取ICF评估数据[9—10]。允许在信息不完全而难以评估的情况下，保留缺失值(记为NA)。为避免过度填补缺失值造成结果偏倚，剔除缺失值比例＞30%者。

1.3 数据处理建模

参照Strobl等[7]和Kalisch等[11]报道的适用于来自便利样本的ICF数据图建模方法。我们前期采用该技术，分别为糖尿病[9]和多囊卵巢综合征[10]绘制了功能地形图。分为5个步骤：①粗加工阶段：正值转置为0(表示无影响)，负值转置为正值(表示不利影响程度)。②缺失值填补和二值化：使用R软件的mice工具包填补缺失值，形成10个填补集，并把数

据二值化(无障碍为0、有障碍为1)。③自举法随机重采样:为每个二值化填补集构建100个有回放的随机重采样集。④初步图建模和统计检验:采用"最小化绝对缩减与选择算子"(least absolute shrinkage and selection operator, LASSO)计算ICF条目间的条件依存关系,构建有向图(directed graph)。图中以条目为节点,一个节点指向另一个节点的带箭头的连线(arcs,弧),表示前者对后者的风险相关关系。弧值>1时,表示风险正相关。0<弧值<1时,越接近0,表示风险负相关程度越大。对于每条可能存在的弧,首先采用1000个重采样值构建置信区间。然后用10个填补集计算未受随机采样影响的实测均数m与标准差sd。最后,当m±2×sd未落入随机重采样的96%置信区间时,可以认为实测值未受随机因素影响,从而把m作为卒中ICF有向图的弧值。不符合要求的弧则删除。⑤有向图转换为无向图:删除有向图中取值小于1的弧,双向弧合并为无方向的连线(edge,边)。边值为两条弧值之和。删除剩余弧之后,形成无向图。图中两点之间的连线,表示两个ICF条目之间互为风险相关。以上步骤均采用R软件(3.2.2版)[12]。

1.4 网络分析探索

采用Pajek64(5.02版)[13]对无向图进行探索性分析。在可视化分析中,ICF大类分为四种形状:s三角、b方框、d圆形、e菱形。分析主要包括4个方面[13]。

①提取主组元(main component)。组元是具有内部连通关系的子网络,其中从任意节点出发,依循连线关系,必定能找到至少一条路径通向其他节点。主组元是节点数量最多的组元。它反映了最大的相互联系的变量集群,有助于把握功能变量的整体结构特征。

②k核解析(k-core decomposition)。k核是指每个节点至少与k个其他节点相连的子网络。k值越大,意味着如果随意去除某个节点,对核团的连通结构影响越小,即核团结构越稳定。

③提取主岛屿(main island)。根据网络中的连线值,从高到低截取子网络。其中含有节点数量最多的主峰,就是网络的主岛屿。它反映了风险相关性最高的变量集群,有助于把握功能变量的重要结构特征。

④关系数据与属性数据的对比分析。从总体图模型中提取简要核心组合中的18个条目以及条目之间的连线(下简称为简要网络),以此反映简要核心条目之间的关系。为观察条目本身的属性,本研究还采用R软件(3.2.2版)的分类数据可视化工具包[12],绘制这些条目在限定值上的频数分布马赛克图。简要网络(关联分析)体现关系视角,而马赛克图(频数分析)则体现属性视角。

2 结果

2.1 组元分析和k核解析

图模型整体含有166个节点和215条连线(图1左上角方形虚线框所示)。主组元含有126个项目和212条连线(图1左上角圆形虚线框所示)。从主组元中,可以剥出一个3-核结构(图1主体所示),其中含有49个条目和89条连线。在3-核中,每个条目至少与三个其他条目相连,从而构成了脑卒中功能变量关系结构中最稳固的变量集群。图1还突出显示了从"d170"到"d440"的一条最短路径。

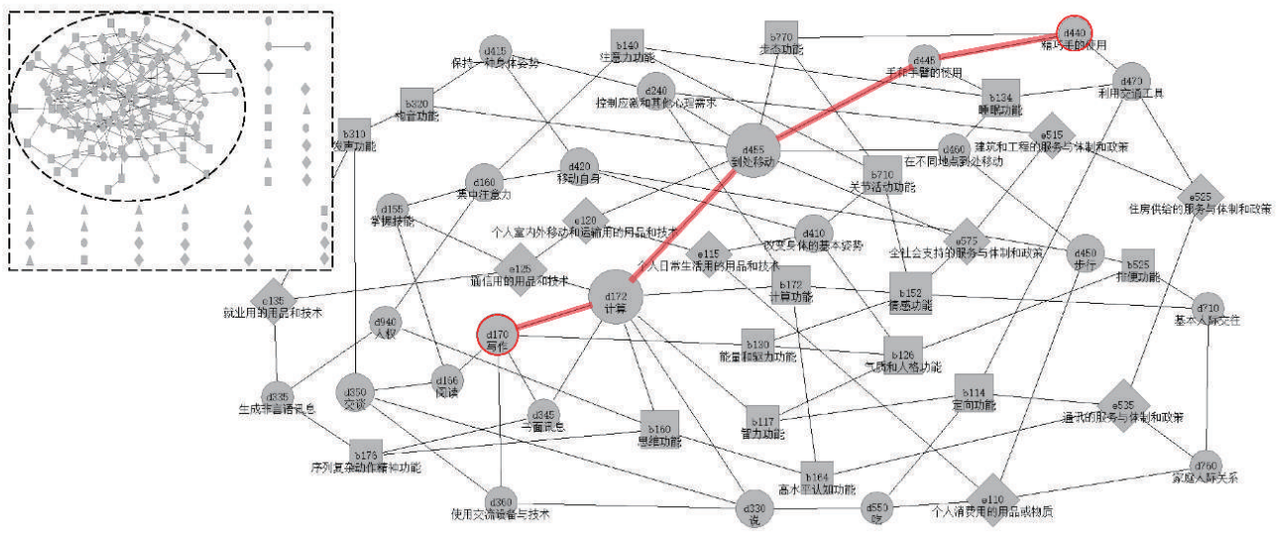
2.2 岛屿结构探索

图2显示了含有80个条目的主岛屿结构。简要核心组合的18个条目中,有8项位于主岛屿内(图2所示绿色和粉色节点)。从左下角开始,按照顺时针观察,可以从图中发现五个分支模块,大致对应于感知觉、人际交往、日常作业、语言交流和环境适应。各个分支模块在整体网络上的接入点有所不同。感知觉模块的接入点在"b164高水平认知功能"。人际交往模块的接入点是"b114定向功能"。日常作业的接入点是"d450步行"。语言交流的接入点"d455到处移动"。环境适应模块则形成了环路结构。现有通用组合和简要核心组合中,占据模块接入点位置的是d455(通用组合条目)、b114(简要核心组合条目)和d450(通用组合与简要核心组合共享条目)。

2.3 ICF组合条目分析

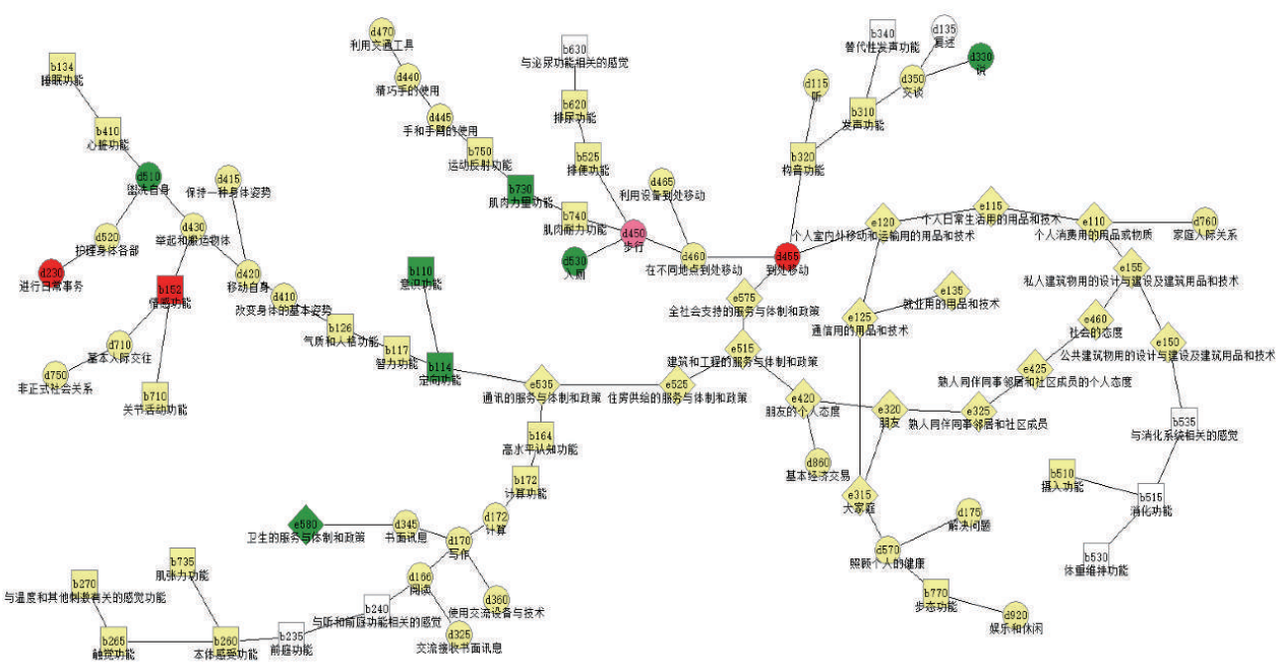
图3分别从关系视角和属性视角,展示了国际版简要核心组合的分析结果。b110和b114存在极高的风险相关(左图:两者间连线值为1535.52),但这种相关性主要在于两者同时无障碍(右图:两者均

图1　ICF图模型总图、组元结构和3–核结构



左上角方形虚框内为图模型总图，圆形虚框内为主组元。图主体部分展示3-核结构。图中节点的大小，与节点拥有的连线数量呈正比。

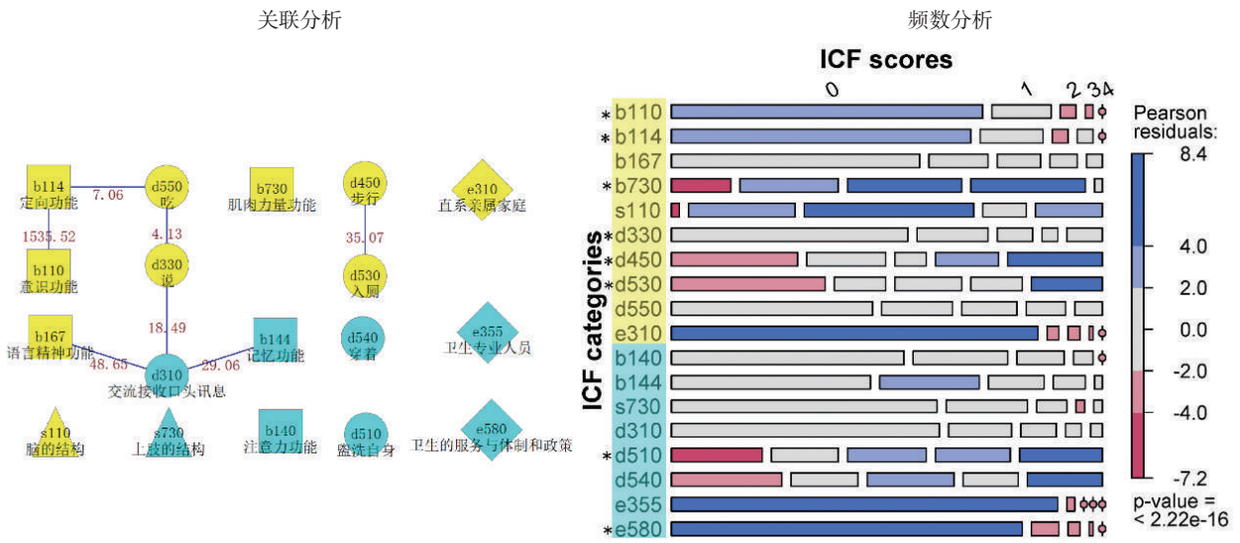
图2　主岛屿



图中用不同色系表示条目所属的组合。红色来自通用组合，绿色属于国际版18条简要核心组合，黄色属于2004年发布的130条综合核心组合，白色属于2005年增补的36条扩展综合核心组合条目。d450既属于通用组合，又属于简要核心组合，以粉色表示。

在0级显著高）。d450和d530两个也存在风险相关关系（左图：两者间存在值为35.07的连线），但这种相关性主要在于两者容易同时受损（右图：两者均在4级显著高）。如左图所示，d310虽然同时影响b167（线值48.65）、d330（线值18.49）和b144（线值29.06），但右图显示d310本身在各限定值级别上无

图3 国际版卒中ICF简要核心组合条目的关联分析和频数分析



左图:从总图中提取简要核心条目组成的子网络。线值表示条目之间的风险相关性。值越大,风险相关性越高。在2004发布的国际版简要核心组合18个条目中,被半数以上专家认可的10个条目[4],用黄色标记。

右图:各简要核心条目在限定值上的频数分布(e条目限定值已转为0表示非不利因素,1—4为不利程度)。各矩形面积越大,表示其中所分布的患者数越多。P值表示独立性χ^2检验的水平。此处$P<2.22\times10^{-16}$提示ICF条目与限定值相关,从而可以用限定值来衡量ICF条目。用从红色向蓝色过渡的色条来表示皮尔逊残差检验的显著性水平。蓝色表示显著高于预期值,红色表示显著低于预期值。深色表示残差$r>4$,对应于显著性水平$P<0.0001$。浅色表示残差值$2<r<4$,对应的显著性水平$P<0.05$。灰色表示频数分布无显著性差异。各条目左侧的星号,表示该条目出现在主岛屿结构中。

明显偏倚。右图还显示b730、s110、d510和d540在中重度障碍水平显著高分布。b144在轻度障碍水平有较高分布。e355和e580则主要在无障碍水平呈现高分布。

3 讨论

作为统一的术语体系,ICF可以综合描述人类基本功能活动。但是,它所含条目多达1495项。在有限的时间内,难以评估所有ICF条目,从而限制了它的临床应用。出于实用目的,研究者为各种常见疾病制定了核心组合,并且发布了适用于多种疾病的通用组合。如果从康复实践角度来考察各种组合的应用,那就必然面临两个基本问题:①哪些条目最重要(要选哪些条目进行干预);②条目之间有什么关系(要按照什么次序来进行干预)。前者决定哪些条目要评,而后者则可以帮助制定干预路径。以往在ICF组合研究中的条目取舍,注重的是问题①。常用方法是从临床调查和专家调研结果中,根据频数统计和聚类分析,提取最重要且最具有代表性的条目纳入核心组合。这种遵循"变量独立性假设"传统分析方法,最终给出的ICF组合,不包含变量之间的风险相关信息[7]。例如,传统的评定方法在评定"d170写作"和"d440精巧手的使用"时,仅能得到两个散在的评定值,而不能提示两者之间存在如图1所示的通路:d170—d172—d455—d445—d440。图建模为ICF评估提供了另一条分析进路[7,9—11,14—16]。从条目之间的关系来衡量条目本身的地位,从而有助于同时解决条目选取和排序问题。本研究以脑卒中为例,从关系型数据(relational data,本研究基于风险相关性)和属性数据(property data,基于频数分布)两个视角[17],对ICF组合条目进行探索,从而为ICF在康复实践中的应用提供新思路和新依据。

本研究发现(图1),脑卒中的扩展版ICF综合核心组合可以构建条目间的风险相关图模型。根据连

线关系判断，模型中存在结构稳定的功能变量核团，并且可以从中发掘变量间的最短联系路径（以图1红色路径为例）。这一方面提示，遵循图模型，可能优化功能恢复路径。另一方面还提示，在制定功能恢复策略时，存在多种备选路径，从而为功能代偿提供了更多可能。组元分析和k核解析，是根据连线有无来探查路径不同，而岛屿分析则是从连线强度来考察路径差异。前者反映的是功能恢复通路的可能性，而后者则反映了功能变量结构的现实性。本研究对图模型的岛屿结构分析发现（图2），可以根据风险相关的强度（连线值），提取出紧密联系的功能变量子群，从中观察到变量依循连接强度所形成的分布路径、功能模块和关键节点。一方面提示，脑卒中对功能状态的影响，存在固有的风险相关分布模式。例如，“d450步行”—“b525排便功能”—“b620排尿功能”—“b630与泌尿功能相关的感觉”，构成了与排泄功能相关的高关联路径。另一方面也提示，一些ICF条目占据了重要地位，是评估和干预的重点。例如，通用组合与简要核心组合共享条目“d450步行”。一方面，它与“d530如厕”有较高风险相关（图2中央绿色节点与粉色节点相连）。另一方面，它还占据了关键接驳位置，影响着肌肉耐力、力量、反射、手与手臂、交通工具，以及排尿与排便等多种作业相关的功能变量（图2正中偏上的网络分支）。

借助关系视角和属性视角的并置分析（图3），本研究还发现d450与d530在卒中患者容易同时受损。这一方面进一步证实d450的重要性。另一方面提示，从关系数据与属性数据中发掘的知识存在交集。另外，通过对两个视角的并置分析，本研究还发现，每个视角都有一些从另一个视角无法解释或推测的知识。例如，b110、b114、d550、d330、d310和b167构成一条风险相关路径。路径中的风险相关强度经历了从高到低递减，而后又递增的过程。但是这种连线关系难以根据属性视角的频数分布模式推断得知。与此类似，e355和e580在马赛克图上有着较为相似的分布模式，但是在网络图中并没有直接的风险相关关系。因此，综合使用关系数据和属性数据，有可能更全面地考察疾病对功能变量的影响。

已有研究表明，图建模能够反映已有临床知识的功能变量结构。例如，Kalisch等[11]对脊髓损伤的研究，Becker等[14]对头颈部肿瘤的研究，以及我们前期对糖尿病和多囊卵巢综合征的研究[9—10]。董艺等[15]分析了30例卒中后吞咽障碍患者的60项ICF条目。采用了与本研究相同的方法构建图模型，但未区分关系和属性分析视角，也未在各级限定值上进行频数统计检验，并且样本量较小。尤萌等[16]采用与本研究规模相当的脑卒中样本（106例），并且采用了LASSO法。但是只纳入了b类条目，并且存在方法错误：①没有采用自举法重采样和置信区间检验；②误把LASSO的系数β直接作为连线值；③误把限定值（等级数据）做均数分析。本研究最大限度地纳入了卒中相关ICF的条目，并且严格遵循了便利样本的统计建模方法，为脑卒中提供了更为完整的功能地形图。

通过对功能地形图进一步实施k-核解析和岛屿分析，以及与频数分布图相互参照，可以印证临床知识和指导临床实践。此处举一例说明，“d450步行”同时属于通用组合和简要核心组合（图2），其障碍频数显著偏高（图3右图），而且与“d530如厕”显著风险相关（图3左图）。这符合临床知识：恢复步行活动是患者常见的诉求之一，也往往是首要康复目标，还是影响如厕的重要因素。但是，从指导临床实践角度来看，通用组合的另一个条目“d455到处移动”占据着卒中患者功能网络的重要位置。①它影响范围最大。如图1所示，是3-核结构中仅有的两个影响范围最大的变量之一，而且还直接与另一个最大影响范围变量“d172计算”相连。另外七个与d455相连的变量是“e575全社会支持的服务与体制和政策”、“d460在不同地点到处移动”、“d445手和手臂的使用”、“b770步态功能”、“d240控制应激和其他心理需求”、“b320构音功能”和“e120个人室外移动和运输用的用品和技术”。②d455是多个功能模块的枢纽。如图2所示，围绕该节点从左上角顺时针观察，日常作业、语言交流和环境适应等模块如果要连通为整体，都需要经过d455。③d455可以通过“d460在不同地点到处移动”影响到“d450步行”。如果把该通路从图2中删除，功能变量就散落成多个互不相连的模块。据此，一方面可以建议把d455添加到卒中简要核心组合，或者在卒中ICF评

定时采用通用组合。另一方面提示,在卒中康复方面,可以围绕步行和移动方面的活动开展治疗,从而对功能产生整体影响。

本课题还有许多方面有待深入研究。①在模型构建方面,需要进一步扩大样本,从样本分层(如性别、卒中类型等因素)探讨不同因素的功能影响。②在模型分析方面,还可以从中心性(centrality)、社团发掘和三方谱(triadic census)等角度[13],深入分析功能变量的关系结构。③需要寻找有效的手段,把图模型用于指导康复实践。本研究在这方面进行了图模型和马赛克图并置分析的尝试。但是,还有待进一步整合关系和属性两大视角,借助现代数据处理技术,探索实用新型的ICF临床应用策略。

参考文献

[1] Kostanjsek N.Use of The International Classification of Functioning, Disability and Health (ICF) as a conceptual framework and common language for disability statistics and health information systems[J]. BMC Public Health, 2011, 11 (Suppl 4):S3.

[2] Silva SM, Corrêa FI, Faria CD, et al. Evaluation of post-stroke functionality based on the International Classification of Functioning, Disability, and Health: a proposal for use of assessment tools[J]. J Phys Ther Sci, 2015, 27(6):1665—1670.

[3] Yen TH, Liou TH, Chang KH, et al. Systematic review of ICF core set from 2001 to 2012[J]. Disabil Rehabil, 2014, 36(3):177—184.

[4] Geyh S, Cieza A, Schouten J, et al. ICF core sets for stroke [J]. J Rehabil Med, 2004, 36(44 Suppl):135—141.

[5] Glässel A, Coenen M, Kollerits B, et al. Validation of the extended ICF core set for stroke from the patient perspective using focus groups[J]. Disabil Rehabil, 2012, 34(2): 157—166.

[6] 曹蓉,许光旭,丁晓晶,等.脑卒中国际功能残疾和健康分类简要核心组合的信度与效度研究[J].中国康复医学杂志,2011,26(8):714—719.

[7] Strobl R, Stucki G, Grill E, et al. Graphical models illustrated complex associations between variables describing human functioning[J]. J Clin Epidemiol, 2009, 62(9):922—933.

[8] Strobl R, Grill E, Mansmann U. Graphical modeling of binary data using the LASSO: a simulation study[J]. BMC Med Res Methodol, 2012, 12(1):16.

[9] 林枫,江钟立,吴亚文,等.糖尿病国际功能、残疾和健康分类综合核心组套的图建模[J].中国康复医学杂志,2012,27(4):308—314.

[10] 林枫,江钟立,王萍,等.超重和肥胖多囊卵巢综合征功能变量关系结构解析[J].中国康复医学杂志,2012,27(8):713—719.

[11] Kalisch M, Fellinghauer BA, Grill E, et al. Understanding human functioning using graphical models[J]. BMC Med Res Methodol, 2010, (10):14.

[12] Core R, Team R. A language and environment for statistical computing[M]. Vienna, Austria: R Foundation for Statistical Computing, 2015.

[13] 沃特·德·诺伊, 安德烈·姆尔瓦, 弗位迪米尔·巴塔盖尔吉.蜘蛛:社会网络分析技术[M].林枫,译.第2版.北京:世界图书出版公司北京公司,2014.

[14] Becker S, Strobl R, Cieza A, et al. Graphical modeling can be used to illustrate associations between variables describing functioning in head and neck cancer patients[J]. J Clin Epidemiol, 2011, 64(8):885—892.

[15] Dong Y, Zhang CJ, Shi J, et al. Clinical application of ICF key codes to evaluate patients with dysphagia following stroke[J]. Medicine (Baltimore), 2016, 95(38):e4479.

[16] 尤萌,姜竹青,王旭,等.ICF脑卒中核心分类组合功能的图形建模研究[J].中国法医学杂志,2016,31(4):336—340.

[17] 林枫,江钟立.网络思维:基于点线符号的认知图式和复杂性范式[J].自然辩证法通讯,2011,33(1):29—35.