

基于中国 1 500 万余例次住院病例的 121 种罕见病现况分析

石鑫森 刘徽 王琳 王朝霞 董冲亚 王燕芳 姚晨 詹思延 丁洁 李岩

【摘要】 目的 运用病案首页数据库,探讨分析我国第一批罕见病目录的 121 种疾病在我国三级甲等医院住院病例的基础数据,了解我国罕见病现状。方法 基于北京医学会罕见病分会前期罕见病调研的数据,对第一批罕见病目录中罕见病在全国 96 家三级甲等医院 2014—2015 年罕见病住院患者现况进行对比及描述性研究分析。结果 经与调研清单比对,第一批罕见病目录中有 19 种罕见病不在调研清单中,其余 102 种罕见病总例次为 54 468 例次,占同期住院总例次的 0.35%。前 10 位最常见罕见病总例次为 37 977 例次,占同期住院总例次的 0.25%。前 10 位最少见罕见病总例次为 24 例次,占同期住院总例次的 0.000 16%。住院人群诊治罕见病种类最多的前 5 个省/直辖市分别为北京、湖南、上海、山东、广东,被收治的住院罕见病例次最多的前 5 个省/直辖市为北京、上海、广东、山东、湖北。年龄分布显示,25~64 岁年龄组占 45.8%,0~14 岁年龄组占 28.6%。罕见病前 10 位重复住院率为 28.42%~64.88%。结论 本研究初步探索了第一批罕见病目录中 121 种罕见病在全国三级甲等医院住院人群中的例次、种类、省/直辖市分布、年龄分布及重复住院率等,为进一步建立全国范围内罕见病注册数据库及制定罕见病相关政策提供基础数据支撑。

【关键词】 罕见病; 数据库; 三级甲等医院

基金项目:国家重点研发计划精准医学研究重点专项课题(2016YFC0901505);儿科遗传性疾病分子诊断与研究北京市重点实验室(BZ0317);北京市科技计划课题(Z151100003915126)

Study on the current situation of China's First List of Rare Diseases based on 15 million hospitalizations Shi Xinmiao*, Liu Hui, Wang Lin, Wang Zhaoxia, Dong Chongya, Wang Yanfang, Yao Chen, Zhan Siyan, Ding Jie, Li Yan. *Department of Pediatrics, Peking University First Hospital, Beijing 100034, China

Corresponding authors: Ding Jie, Email: djnc_5855@126.com; Li Yan, Department of Hospital Administration of Peking University Health Science Centre, Peking University, Beijing 100191, China, Email: liyan8290@163.com

【Abstract】 Objective To investigate and analyze the distribution of 121 diseases of China's First List of Rare Diseases based on hospitalized patients of tertiary hospitals and to explore the current situation of rare diseases in China. **Methods** Based on previous data of study from Beijing Society of Rare Diseases, a comparison between China's First List of Rare Diseases and the survey list from the pre-study was performed. Descriptive analysis was carried out on the current situation of rare diseases on hospitalizations in 96 tertiary hospitals from year of 2014 to 2015. **Results** Nineteen out of 121 diseases on China's First List of Rare Diseases were not included in the rare diseases survey list of Beijing Society of Rare Diseases. The total number of other 102 rare disease cases was 54 468, accounting for 0.35% of the inpatients during the same period. The top ten most and least cases with rare disease were demonstrated in this study. The number of the top ten most cases was 37 977, accounting for 0.25% of the inpatients during the same period. The number of the top ten least cases was 24, accounting for 0.000 16% of the inpatients during the same period. The top most five types of rare diseases counted on the provinces and municipalities were Beijing, Hunan, Shanghai, Shandong and Guangdong. The top five most cases of rare diseases counted on the provinces and municipalities were Beijing, Shanghai, Guangdong, Shandong and Hubei. The age

DOI:10.3760/cma.j.issn.0376-2491.2018.40.012

作者单位:100034 北京大学第一医院儿科(石鑫森、丁洁),神经内科(王朝霞),医学统计室(董冲亚、姚晨);北京大学医学信息学中心(刘徽);北京医学会罕见病分会(王琳);北京大学临床研究所(王燕芳);北京大学公共卫生学院流行病与卫生统计学系(詹思延);北京大学医学部医院管理处(李岩)

通信作者:丁洁,Email:djnc_5855@126.com;李岩,Email:liyan8290@163.com

distribution showed that the cases with rare diseases aged 25 - 64 years accounted for 45.8%, and the cases in children aged 0 - 14 accounted for 28.6%. The top ten readmission rate ranged from 28.42% to 64.88%. **Conclusions** This study preliminarily investigates the number, type, province and municipality distribution, age distribution, and readmission rate of 121 rare diseases from China's First List of Rare Diseases in the hospitalized patients of tertiary hospitals, which provides important data for registration study, medical and drug policy making and other relevant work on rare diseases in China in the future.

【Key words】 Rare diseases; Databases; Tertiary care centers

Fund program: The National Key Research and Development Program of China (2016YFC0901505); Beijing Key Laboratory of Molecular Diagnosis and Study on Pediatric Genetic Disease (BZ0317); Beijing Municipal Science and Technology Project (Z151100003915126)

罕见病又称“孤儿病”,指发病率极低,患病人口数量极少的一大类疾病。2016 年 Orphanet 数据库年度报告指出目前罕见病 6 900 种以上^[1]。

2018 年 5 月,由国家卫生健康委员会、科学技术部、工业和信息化部、国家药品监督管理局、国家中医药管理局五部门联合发布了《第一批罕见病目录》,包括 121 种罕见病(以下统称为“目录”)。此目录对于罕见病的临床诊治、药物研发、医疗救助、科研、医保和社会保障相关政策的制定等都具有重要的推进意义,在我国罕见病事业发展史上具有里程碑意义。

然而,我国罕见病领域的临床和研究尚属起步阶段,目前尚无全国范围内的罕见病流行病学调查数据库,严重制约开展相关工作^[2]。为此,本研究基于国家卫生健康委员会医疗管理服务指导中心的住院病案首页数据库(以下统称为“病案首页数据库”),对我国三级甲等医院住院病例中《第一批罕见病目录》中 121 种罕见病的现况进行研究和分析。

对象与方法

一、对象

1. 数据来源:病案首页数据库是基于住院人群、涵盖全国 24 个省/直辖市、96 家三级甲等医院的 2014—2015 年期间 15 458 065 住院例次病案首页的数据。该数据库基于数据管理平台系统框架,自下而上包括数据集成层、数据存储与管理层、数据分析与挖掘层及分析结果展示层,经过病案首页上传、清洗、入库、查询输出、分析挖掘、结果报告等步骤输出研究结果,且全程数据经过安全控制及质量控制^[3]。本研究由数据库专业管理人员根据本研究目的及研究设计进入有限开放权限的数据库进行数据提取。

2. 研究对象:本研究以病案首页数据库中出院诊断符合《第一批罕见病目录》中疾病的患者为研

究对象。本研究经作者所在医院伦理委员会批准。

二、方法

1. 从调研清单中确认目录中疾病:调研清单综合了原国家卫计委专家建议的 151 种罕见病清单、北京医学会罕见病分会专家建议的 121 种罕见病清单、上海儿童医学中心陈静^[4]主编的《可治性罕见病》中的 117 种罕见病清单、“国民经济和社会发展第十三个五年规划”国家重点研发计划精准医学研究重点专项项目罕见病临床队列研究的 66 种罕见病清单(2016YFC0901500)及山东省罕见病病例信息登记工作的 68 种罕见病清单,将上述汇集的罕见病去重以及逐一对照英文校对各个罕见病中文名称,继之运用“会议共识法”手段,组织进行了两轮“专家共识会”,通过公开讨论、轮流发言、无记名投票等形式,对罕见病初步调研清单中纳入的疾病进行筛选,并对疾病中英文名称进行修订、完善,最终形成“罕见病调研清单”(以下简称“调研清单”)。以目录中的 121 种罕见病为研究对象,若目录中疾病在调研清单中,则进行描述性统计分析。

2. 数据提取与清理:对调研清单中每个疾病,使用多视角卷积神经网络算法在病案首页数据库中进行疾病匹配^[5]。以匹配结果为基础,以疾病为单位建立子数据库,每个疾病数据库由该疾病领域专家进行疾病去重复并根据每一种罕见病的中文译文名称代表的疾病定义对提取结果进行整理,确保提取数据的准确性。

3. 统计学方法:对目录中罕见病住院患者数据进行基本统计描述,包括:2014 年例次、2015 年例次、总例次、罕见病种类、省/直辖市分布、年龄分布及重复住院率。连续变量采用 $\bar{x} \pm s$ 进行统计描述,分类变量采用频数及百分比进行统计描述。数据提取与统计分析均运用统计学软件 R (3.5.1)。

结 果

1. 目录与调研清单比对结果:目录中有 19 种

罕见病不在调研清单中,其余102种罕见病在调研清单中。

2. 目录中 102 种罕见病的例次及省/直辖市分布:(1)罕见病例次及省/直辖市分布:目录中 102 种罕见病总例次为 54 468 例次,占同期住院总例次的 0.35% (54 468/15 458 065)。2014 年总例次占同期住院患者的 0.36% (26 475/7 429 813),2015 年总例次占同期住院患者的 0.35% (27 993/8 028 252)。

目录中有 14 种罕见病未被检索到,另外 88 种罕见病中,前 10 位最多例次罕见病总例次占 102 种罕见病总例次的 69.7% (37 977/54 468),占同期住院总例次的 0.25% (37 977/15 458 065)。前 10 位最少例次罕见病总例次仅占 102 种罕见病总例次的 0.04% (24/54 468),占同期住院总例次的 0.000 16% (24/15 458 065)。各疾病对应的 2014—2015 年总例次见表1。(2)罕见病种类、总例次及省/直辖市

表 1 目录中 88 种罕见病在病案首页数据库中的例次分布

目录编码	疾病名称	总例次	目录编码	疾病名称	总例次
54	特发性肺动脉高压	9 536	109	脊延髓肌萎缩症(肯尼迪病)	101
60	朗格汉斯组织细胞增生症	7 657	120	X-连锁肾上腺脑白质营养不良	93
4	肌萎缩侧索硬化	4 057	92	卟啉病	92
55	特发性肺纤维化	3 764	101	囊性纤维化	87
112	系统性硬化症	3 252	15	原发性肉碱缺乏症	86
37	肝豆状核变性	3 000	82	尼曼匹克病	86
102	视网膜色素变性症	1 852	51	低血磷性佝偻病	85
68	马凡综合征	1 751	27	法布雷病	69
46	纯合子家族性高胆固醇血症	1 714	99	丙酸血症	58
23	先天性脊柱侧弯	1 394	18	瓜氨酸血症	55
88	阵发性睡眠性血红蛋白尿	1 337	64	淋巴管肌瘤病	47
91	POEMS 综合征	1 213	5	Angelman 综合征	42
35	糖原累积病(I 型、II 型)	1 178	20	先天性高胰岛素性低血糖血症	40
25	先天性纯红细胞再生障碍性贫血	1 113	67	枫糖尿症	40
114	结节性硬化症	1 064	28	家族性地中海热	32
98	进行性肌营养不良	926	94	原发性联合免疫缺陷	31
53	特发性低促性腺激素性腺功能减退症	904	95	原发性遗传性肌张力不全	27
56	IgG4 相关性疾病	900	117	威廉姆斯综合征	20
8	非典型溶血尿毒综合征	532	62	Leber 遗传性视神经病变	19
73	黏多糖贮积症	507	26	Erdheim-Chester 病	18
71	甲基丙二酸血症	470	22	先天性肌强直(非营养不良性肌强直综合征)	17
86	成骨不全症(脆骨病)	469	38	遗传性血管性水肿	15
110	脊髓性肌萎缩症	389	52	特发性心肌病	15
29	范可尼贫血	350	84	Noonan 综合征	15
2	白化病	333	6	精氨酸酶缺乏症	14
87	帕金森病(青年型、早发型)	315	115	原发性酪氨酸血症	14
16	Castleman 病	301	107	Silver-Russell 综合征	13
43	遗传性痉挛性截瘫	298	121	X-连锁淋巴增生症	12
33	Gitelman 综合征	266	21	先天性肌无力综合征	10
17	腓骨肌萎缩症	264	30	半乳糖血症	10
3	Alport 综合征	250	104	重症先天性粒细胞缺乏症	8
90	苯丙酮尿症	246	13	生物素酶缺乏症	7
111	脊髓小脑性共济失调	168	19	先天性肾上腺发育不良	7
59	卡尔曼综合征	152	50	低碱性磷酸酶血症	7
89	黑板息肉综合征	149	61	莱伦氏综合征	7
74	多灶性运动神经病	139	118	湿疹血小板减少伴免疫缺陷综合征	5
69	McCune-Albright 综合征	136	97	进行性家族性肝内胆汁淤积症	4
49	高苯丙氨酸血症	126	75	多种酰基辅酶 A 脱氢酶缺乏症	2
78	肌强直性营养不良	125	34	戊二酸血症 I 型	1
31	戈谢病	123	39	遗传性大疱性表皮松解症	1
47	亨廷顿舞蹈症	117	40	遗传性果糖不耐受症	1
119	X-连锁无丙种球蛋白血症	109	41	遗传性低镁血症	1
105	婴儿严重肌阵挛癫痫(Dravet 综合征)	104	58	异戊酸血症	1
93	Prader-Willi 综合征	102	85	鸟氨酸氨甲酰基转移酶缺乏症	1

注:目录编码为目录中的疾病编码

分布:在病案首页数据库涵盖的 24 个省级行政单位中,住院人群诊治罕见病种类最多的前 5 个省/直辖市分别为北京、湖南、上海、山东、广东,其对应罕见病种类分别为 76 种、69 种、62 种、61 种及 59 种,被收治的住院罕见病患者最多的前 5 个省/直辖市为北京、上海、广东、山东、湖北,其对应罕见病例次分别为 13 770 例次、3 875 例次、3 726 例次、2 815 例次及 2 637 例次。(3)罕见病总例次的性别及年龄分布:两年所有罕见病患者总例次为 54 468 例次,其中男性为 29 655 例次,占 54.4%;各年龄段例次及占比见表 2。(4)罕见病重复住院率:重返例次是指将每种罕见病在病案首页数据库中的所有住院患者在同一医院重复住院次数总和。重复住院率是指重返例次占住院例次的百分比。重复住院率前 10 位疾病重复住院率范围为 28.42% ~ 64.88%,详情见表 3。

讨 论

本研究提示总例次最多的前 10 位罕见病例次范围为 1 394 ~ 9 536 例次,表明有些罕见病在我国总的发病人数并不少见,应引起足够的重视。这些疾病应编入疾病目录并进行归类,充分认识这类疾病,有利于给予病人及时恰当的治疗。

根据国内罕见病相关文献报道^[6-17],对于前 10 位最少见罕见病,多个文献报道的数量总和,甚至某些单中心报道的例数多于本研究结果的例次,此结果可能与某些种类罕见病仅在门诊就诊而不收住院治疗有关,也可能与病案首页数据库未囊括全国所有医院有关。

由于患者的籍贯等信息属于保密信息,本文的省市分布为诊治医院分布情况。省/直辖市分布显示,罕见病患者大多数集中分布在北京、上海、广东及山东,这提示更多的罕见病患者到这些省市就医,另外,也间接反映这些省市的三甲医院更具备诊治罕见病的能力。

两年数据显示罕见病诊治例次占同期住院患者例次的比例基本相同(0.36% 和 0.35%),提示罕见

表 3 目录中重复住院率前 10 位罕见病

目录编码	疾病名称	重返例次	总例次	重复住院率(%)
60	朗格汉斯组织细胞增生症	4 968	7 657	64.9
6	精氨酸酶缺乏症	9	14	64.3
119	X-连锁无丙种球蛋白血症	44	109	40.4
118	湿疹血小板减少伴免疫缺陷综合征	2	5	40.0
8	非典型溶血性尿毒症	205	532	38.5
16	Castleman 病	111	301	36.9
25	先天性纯红细胞再生障碍性贫血	365	1 113	32.8
101	肺囊性纤维化	28	87	32.2
91	POEMS 综合征	390	1 213	32.2
88	陈发性睡眠性血红蛋白尿	380	1 337	28.4

注:目录编码为目录中的疾病编码

病住院诊治情况在我国基本平稳。年龄分布显示,25 ~ 64 岁年龄组占 45.8%,提示劳动力人群占极大比例,反映了罕见病对家庭、社会及医疗的影响;0 ~ 14 岁年龄组占 28.6%,提示儿童是罕见病住院患者的重要组成部分。总罕见病住院患者例次以及各省/直辖市分布、年龄分布等情况有助于在制定相关医保、救助等罕见病政策时作为参考数据。此外,本研究从病案首页数据库中提取的罕见病重复住院率的数据,尽管目前病案首页数据库尚不能提供相应的医疗支出详情,但仍然有助于分析了解相关罕见病的医疗负担。

在大数据策略指导下,基于疾病数据库中的疾病信息进行数据挖掘,是疾病流行病学的有效方法^[16]。我国罕见病调研缺乏全国性的流行病学数据,此调研工作以全国三级甲等医院的 1 545 万例次住院病例的病案首页为基础,经过严谨的数据提取及清理,由于某些罕见病方面没有准确的 ICD-10 编码或多种疾病名称,因此对数据必须进行人工二次核查,以确保数据信息的准确性。本研究可提供第一批罕见病目录中 84.3% 的疾病的例次及省/直辖市分布情况,为各级政府机构制定医疗保障、大病救助、孤儿药研发等相关政策^[17-19]提供基础数据资料。

本研究存在一些局限性。罕见病病例数可能被

表 2 目录中 88 种罕见病在病案首页数据库中总例次的年龄分布(例)

年份	总例次	0~4 岁	5~14 岁	15~24 岁	25~34 岁	35~44 岁	45~54 岁	55~64 岁	65~74 岁	75~84 岁	>85 岁
2014	26 475	4 158	3 075	2 189	2 672	2 688	3 379	3 807	2 655	1 529	284
2015	27 993	4 717	3 577	2 183	2 626	2 618	3 506	3 626	2 775	1 846	470
合计	54 468	8 875	6 652	4 372	5 298	5 306	6 885	7 433	5 430	3 375	754

注:88 例次患者年龄数据缺失

低估,首先,数据库中患者为住院患者,而很多罕见病患者可能仅在门诊就诊。其次,尽管数据库内医院为三级甲等医院,且均为省级行政单位的委属(管)医院、卫/教共建医院或大型公立医院,很可能代表了具有诊治疑难重症能力的医疗机构,但并未囊括所有三级甲等医院。另外,由于数据库中患者的身份信息属于保密信息,无法实现跨医院个体识别,因此得到的重返例次仅为在同一医院重复住院的例次。

总之,本研究通过比对第一批罕见病目录与北京医学会罕见病分会前期罕见病调研清单的疾病种类,初步探索了第一批罕见病目录中罕见病在全国三级甲等医院住院人群中的例次、种类、省/直辖市分布、年龄分布及重复住院率,为进一步建立全国范围内罕见病注册数据库及制定罕见病相关政策提供基础数据支撑。

志谢 国家卫生计生委医疗管理服务指导中心;北京医学会罕见病分会

参 考 文 献

- [1] Orphanet. 2016 Activity report, Orphanet Report Series, Reports Collection, (V1. 2) [EB/OL]. (2016) [2017-07]. <http://www.orpha.net/orphacom/cahiers/docs/GB/ActivityReport2016.pdf>.
- [2] Yang L, Su C, Lee AM, et al. Focusing on rare diseases in China: are we there yet? [J]. Orphanet J Rare Dis, 2015, 10: 142. DOI: 10.1186/s13023-015-0361-3.
- [3] 包小源, 俞国培, 李岩, 等. 病案首页数据分布式集成管理平台的设计与应用 [J]. 中国医院管理, 2015, 34(5): 30-32.
- [4] 陈静. 可治性罕见病 [M]. 1 版. 上海: 上海交通大学出版社, 2017.
- [5] Luo Y, Song G, Li P, et al. Multi-Task Medical Concept Normalization Using Multi-View Convolutional Neural Network [R]. New Orleans: National Conference on Artificial Intelligence, 2018: 17066.
- [6] 何雨蔚. 20 例 Wiskott-Aldrich 综合征病例分析 [D]. 重庆医科大学, 2015: 1-28.
- [7] 刘源, 夏强, 张建军, 等. 活体肝移植治疗进行性家族性肝内胆淤积症 [J]. 中华小儿外科杂志, 2015, 36(10): 758-760. DOI: 10.3760/cma.j.issn.0253-3006.2015.10.010.
- [8] 李溪远, 华瑛, 丁圆, 等. 新生儿期发病的经典型异戊酸血症四例分析 [J]. 中华围产医学杂志, 2015, 18(3): 188-194. DOI: 10.3760/cma.j.issn.1007-9408.2015.03.005.
- [9] 刘琦, 陈益平, 陈伟. 四例戊二酸血症 I 型患儿 GCDH 基因的突变研究 [J]. 中华医学遗传学杂志, 2015, 32(2): 187-191. DOI: 10.3760/cma.j.issn.1003-9406.2015.02.008.
- [10] 陈占玲, 温鹏强, 王国兵, 等. 鸟氨酸氨甲酰基转移酶缺乏症患儿 OTC 基因突变分析 [J]. 中华医学遗传学杂志, 2014, 31(5): 565-569. DOI: 10.3760/cma.j.issn.1003-9406.2014.05.005.
- [11] 王艳, 刘欣, 吴虹林, 等. 鸟氨酸氨甲酰基转移酶缺乏症的临床特征与代谢谱及 OTC 基因突变分析 [J]. 中华医学遗传学杂志, 2014, 31(2): 148-151. DOI: 10.3760/cma.j.issn.1003-9406.2014.02.005.
- [12] 施晓容, 柯钟灵, 郑爱东, 等. 一个戊二酸血症 I 型患者家系的临床分析及基因突变研究 [J]. 中华医学遗传学杂志, 2014, 31(5): 608-611. DOI: 10.3760/cma.j.issn.1003-9406.2014.05.015.
- [13] 刘琦, 陈益平. 儿童戊二酸血症 I 型 2 例报告 [J]. 中国当代儿科杂志, 2014, 16(1): 87-90. DOI: 10.7499/j.issn.1008-8830.2014.01.020.
- [14] 徐凯峰, 吴珠明, 王卫凯. 新生儿异戊酸血症 2 例并文献分析 [J]. 国际儿科学杂志, 2014, 41(2): 215-216. DOI: 10.3760/cma.j.issn.1673-4408.2014.02.031.
- [15] 付溪, 高洪杰, 吴婷婷, 等. 异戊酸血症 2 例患儿的临床研究并文献复习 [J]. 中华实用儿科临床杂志, 2014, 29(8): 599-604. DOI: 10.3760/cma.j.issn.2095-428X.2014.08.011.
- [16] 曲艳吉, 詹思延. 罕见疾病患病率的流行病学研究方法 [J]. 中华儿科杂志, 2015, 53(4): 309-312. DOI: 10.3760/cma.j.issn.0578-1310.2015.04.019.
- [17] 张冀. 建立中国罕见病医疗保障体系的几点探讨 [J]. 哈尔滨医药, 2014, 34(2): 85-86.
- [18] 刘慧敏, 孟锐. 关于加快建立我国孤儿药政策的建议 [J]. 中国药物评价, 2012, 29(2): 176-179.
- [19] 刘菲, 周静, 胡明. 我国罕见病用药医疗保障政策及医保目录收录情况分析 [J]. 中国卫生经济, 2018, 37(3): 71-76. DOI: 10.7664/CHE20180320.

(收稿日期:2018-06-25)

(本文编辑:陈新石)

· 读者 · 作者 · 编者 ·

本刊“临床医学影像”栏目征稿

医学影像检查是临床常用的诊断手段。影像学改变是病理改变的反映,但不同的病理改变往往有相似的影像学表现,这给诊断带来很大困难。为了促进临床影像诊断经验的交流和诊断、鉴别诊断水平的提高,中华医学杂志自2001年第1期开辟“临床医学影像”栏目,为特殊的、少见的、但具有临床启发意义的影像学表现提供一个展示园地,使局部的、个人的经验尽快地为广大临床医师借鉴,为临床医学影像诊断积累宝贵的第一手资料。本栏目是一个以图片展示

为主的栏目,要求提供高质量的影像图片,图片必须清晰、对比度好、病变特征显示明确。每篇文章可提供2~4幅不同影像技术的图片,如X线、CT、磁共振成像、超声、核素显像或病理图片等。文字部分则宜简练,描述患者的简要病史,主要影像学表现,经病理或临床科学手段确定的最后诊断结果,不进行讨论,不引用参考文献,字数在400字以内。欢迎踊跃投稿。