

基于 LASSO-logit 回归模型的内蒙古流动人口卫生服务利用影响因素分析*

内蒙古医科大学卫生管理学院(010100) 胡嘉靖 刘思佳[△]

【摘要】 目的 研究内蒙古流动人口卫生服务利用状况及其影响因素,为政府管理部门推进优化政策措施提供建议。方法 使用 2017 年全国流动人口卫生计生动态监测调查数据,采用卡方检验、 t 检验、秩和检验及 LASSO-logit 回归等方法进行影响因素分析。结果 内蒙古流动人口医疗服务利用率为 49.98%。通过 LASSO-logit 回归筛选,民族、家庭月收入、是否工作、是否独居、居民健康档案建立、接受健康教育状况评分、居住地到最近医疗机构距离、健康状况等因素对流动人口卫生服务利用产生影响。进一步使用二元 logit 回归分析显示,是否工作、家庭月收入、居住地到最近医疗机构距离、接收健康教育状况评分、健康状况对流动人口的医疗服务利用具有显著影响。结论 内蒙古流动人口卫生服务利用行为亟待改善,使能资源能够有效改善流动人口医疗服务利用行为,健康状况对于流动人口医疗服务利用具有显著影响,建议政府和社会重视流动人口健康状况,通过加强流动人口社会保障覆盖、继续开展健康教育宣传、提供健康咨询和指导等手段改善卫生服务利用行为。

【关键词】 流动人口 卫生服务利用 影响因素 内蒙古

【中图分类号】 R195.1 **【文献标识码】** A **DOI** 10.11783/j.issn.1002-3674.2025.01.018

改革开放以来,中国城镇化快速发展,但户籍人口城镇化率和常住人口城镇化率之间仍然有较大差异^[1],即存在很大一部分的人户分离群体。第七次人口普查数据显示,2020 年我国流动人口达到 3.76 亿人,比 2010 年增加了 1.55 亿,研究显示,我国人口迁移流动的整体规模和强度今后还有相当大的上升空间^[2]。流动人口常常从事劳动密集型工作,生活环境较差,存在较大的健康风险^[3]。与一般居民相比,流动人口人户分离的特征使其难以享受公平的健康保障^[4],改善流动人口健康水平成为推动中国式现代化进程中的重要任务。

医疗卫生服务的充分利用是降低人群发病率和死亡率的关键因素之一^[5]。内蒙古自治区历来就是人口迁移流动比较频繁的地区之一^[6],流动人口占总人口比例较高,因此有必要对内蒙古自治区流动人口卫生服务利用状况进行研究。本研究以安德森卫生服务利用模型为基础,使用 LASSO-logit 回归方法对内蒙古流动人口的医疗卫生服务利用状况及其影响因素进行分析,为政府管理部门推进优化政策措施提供参考。

资料和方法

1. 资料来源

数据资料来源于国家卫生健康委 2017 年全国流动人口动态监测调查数据(China migrants dynamic survey, CMDS)。CMDS 是一项大规模全国流动人口抽样调查,覆盖全国 31 个省市和新疆生产建设兵团,

采用分层、多阶段、与规模成比例的 PPS 方法进行抽样,纳入的调查对象为在流入地居住一个月以上且年龄在 15 周岁以上的流入人口。2017 年内蒙古自治区共调查流动人口 5000 人,以参与回答问题“最近一次患病(负伤)或身体不适时,您首先去哪里看的病/伤?”对样本进行筛选,剔除未回答该问题的样本,最终纳入研究样本数为 2349 人。

2. 方法

(1) 指标选取与解释

调查问卷由国家卫生健康委印制,调查过程及其数据经过严格的培训、自查、审核和质量控制。本研究的医疗服务利用情况作为因变量,考察最近一次患病或负伤时是否就诊。自变量指标选取采用安德森卫生服务利用模型为指导^[7],纳入倾向特征、使能资源、需求因素和流动因素等 4 个方面,其中倾向特征包括性别、年龄、受教育年限、户口性质、是否工作、婚姻状况、民族、是否独居、住所是否稳定;使能资源包括家庭月收入、基本医疗保险、居住地到最近的医疗服务机构距离、健康档案是否建立、接受健康教育状况;需求因素包括自评健康、慢性病患病状况。流动因素包括流动范围、流动时间。

(2) LASSO 回归

传统的回归模型通常使用最小二乘法进行估计,在纳入相关因素较多时,由于建立的回归模型较为复杂,容易出现模型过拟合问题^[8],同时不同自变量之间可能存在多重共线性,使得模型稳健性不足。LASSO 回归在模型中使用函数系数的绝对值函数对回归系数进行惩罚,使一些与因变量的效应非常弱的系数变小^[9]。因此 LASSO 回归可以对数据进行降维以达到筛选自变量的作用^[10]。

* 基金项目:内蒙古自然科学基金青年项目(2022QN07001);内蒙古自治区哲学社会科学规划项目(2021ZZB051);内蒙古医科大学青年培育项目(YKD2021QN038);内蒙古医科大学面上项目(YKD2023MS069)

[△]通信作者:刘思佳,E-mail:liusijawuhai@163.com

3. 统计分析

采用 Stata 17.0 软件进行统计分析。定性资料采用率及构成比[$n(\%)$]进行描述,组间使用卡方检验。定量资料使用平均数及标准差进行描述($\bar{x}\pm s$),对数据进行正态性检验和方差齐性检验,服从正态分布且方差齐次的采用独立样本 t 检验,不服从正态分布的采用秩和检验。建立 LASSO-logit 模型对自变量进行筛选,使用 10 折交叉验证(10-fold cross-validation)选择最优的调和参数 λ 。使用经过筛选后的自变量建立二元 logit 回归分析流动人口卫生服务利用的影响因素。检验水准取 $\alpha=0.05$ 。

结果

1. 流动人口基本情况及医疗服务利用的单因素分析

内蒙古流动人口中利用卫生服务的比例是 50.02% (1175 人),未就诊的比例为 49.98% (1174 人),两组流动比例均衡。在个人特征中,男女比例差异较小,年龄以 50 岁以下的中青年为主,占比超过

80%,少数民族人数较少(9.37%),总体教育程度较低,平均受教育年限仅为 9.62 年,农业户籍人口占绝大多数(82.50%),拥有稳定住房的人口约 50%。使能资源方面,家庭月收入均值为 5449.02 元,中位数为 5000 元,未参加基本医疗保险的较少(6.77%),参加城乡居民医疗保险的流动人口占比高(79.14%),大多数被调查者居住地 15 min 以内均可抵达最近的医疗服务机构(78.12%),但超过 60%的流动人口未建立居民健康档案,对 9 类疾病和保健知识的健康教育进行调查发现,29.80%的被调查者从未接受健康教育。需求因素方面,慢性病患病率较低(13.37%),自评健康中 89.27%认为身体健康或基本健康。流动因素方面,跨省流动人口相对较少(22.39%),省内跨市流动相对较多(41.46%),流动时间在 6 年以上的占比超过一半(54.02%),6 年以下的分布较为均匀。通过单因素分析发现,就诊组与未就诊组年龄、教育年限、是否工作、居住地到最近医疗机构距离、居民健康档案、接受健康教育状况、慢性病患病状况、自评健康的差异存在统计学意义(表 1)。

表 1 内蒙古流动人口医疗服务利用行为的单因素分析

| 变量 | 未就诊组 (n=1174) | 就诊组 (n=1175) | 统计量 | 变量 | 未就诊组 (n=1174) | 就诊组 (n=1175) | 统计量 |
|------------------------|--------------------|--------------------|----------------------|---------------------------|------------------|------------------|----------------------|
| 性别 | | | $\chi^2=0.021$ | 基本医疗保险 | | | $\chi^2=1.155$ |
| 男 | 589(50.17%) | 586(49.87%) | | 居民保险 | 924(78.71%) | 935(79.57%) | |
| 女 | 585(49.83%) | 589(50.13%) | | 职工保险 | 164(13.97%) | 167(14.21%) | |
| 年龄(median, IQR) | 36.17(29.42~45.77) | 37.58(30.08~48.50) | $z=-2.957^{**}$ | 未参保 | 86(7.33%) | 73(6.21%) | |
| 教育年限($\bar{x}\pm s$) | 9.79 \pm 3.59 | 9.45 \pm 3.60 | $t=2.306^*$ | 居住地到最近医疗机构距离 | | | $\chi^2=13.876^{**}$ |
| 户口性质 | | | $\chi^2=0.485$ | 15 min 以内 | 951(81.01%) | 884(75.23%) | |
| 农业 | 975(83.05%) | 963(81.86%) | | 15 min~30 min | 201(17.12%) | 257(21.87%) | |
| 非农业 | 199(16.95%) | 212(18.04%) | | 30 min~1 h | 17(1.45%) | 31(2.64%) | |
| 是否工作 | | | $\chi^2=16.789^{**}$ | 1 h 以上 | 5(0.43%) | 3(0.26%) | |
| 是 | 844(71.89%) | 752(64.00%) | | 居民健康档案建立 | | | $\chi^2=10.100^{**}$ |
| 否 | 330(28.11%) | 423(36.00%) | | 是 | 437(37.22%) | 513(43.66%) | |
| 婚姻状况 | | | $\chi^2=0.174$ | 否 | 737(62.78%) | 662(56.34%) | |
| 在婚 | 1052(89.61%) | 1059(90.13%) | | 接受健康教育状况 (median, IQR) | 2.00(0.00~5.00) | 2.00(0.00~6.00) | $z=-2.737^{**}$ |
| 非在婚 | 122(10.39%) | 116(9.87%) | | 慢性病患病状况 | | | $\chi^2=9.889^{**}$ |
| 民族 | | | $\chi^2=3.090$ | 是 | 131(11.16%) | 183(15.57%) | |
| 汉族 | 1067(90.89%) | 1062(90.38%) | | 否 | 1043(88.84%) | 992(84.43%) | |
| 蒙古族 | 83(7.07%) | 76(6.47%) | | 自评健康 | | | $\chi^2=63.887^{**}$ |
| 其他 | 24(2.04%) | 37(3.15%) | | 健康 | 877(74.70%) | 703(59.83%) | |
| 是否独居 | | | $\chi^2=1.255$ | 基本健康 | 214(18.23%) | 303(25.79%) | |
| 是 | 36(3.07%) | 46(3.91%) | | 不健康但生活可以自理 | 81(6.90%) | 164(13.96%) | |
| 否 | 1138(96.93%) | 1129(96.09%) | | 生活不能自理 | 2(0.17%) | 5(0.43%) | |
| 住房是否稳定 | | | $\chi^2=0.584$ | 流动范围 | | | $\chi^2=2.881$ |
| 是 | 615(52.39%) | 597(50.81%) | | 跨省 | 254(21.64%) | 272(23.15%) | |
| 否 | 559(47.61%) | 578(49.19%) | | 省内跨市 | 507(43.19%) | 467(39.74%) | |
| 家庭月收入(median, IQR) | 5.00(3.79~6.50) | 5.00(3.20~6.50) | $z=0.908$ | 市内跨县 | 413(35.18%) | 436(37.11%) | |
| | | | | 流动时间(Median, IQR) | 6.33(3.42~11.25) | 6.75(3.08~11.08) | $z=-0.442$ |

注: * $P<0.05$; ** $P<0.01$

2. LASSO 回归筛选变量

将所有 18 个自变量纳入 LASSO 回归进行变量筛选,构建系数分布图(图 1a)和交叉验证图(图 1b)。在系数分布图中,随着横轴所代表的惩罚系数 λ 增加,纵轴所代表的影响因素的系数不断被压缩至 0,即选择的自变量个数减少。在交叉验证图

中,在 $\lambda=0.008$ 附近纵轴所代表的交叉验证误差走势变得平坦,因此选择该值作为最优模型,此时 LASSO 回归模型共纳入 8 个自变量,包括民族、家庭月收入、是否工作、是否独居、居民健康档案建立、接受健康教育状况、居住地到最近医疗机构距离、健康状况。

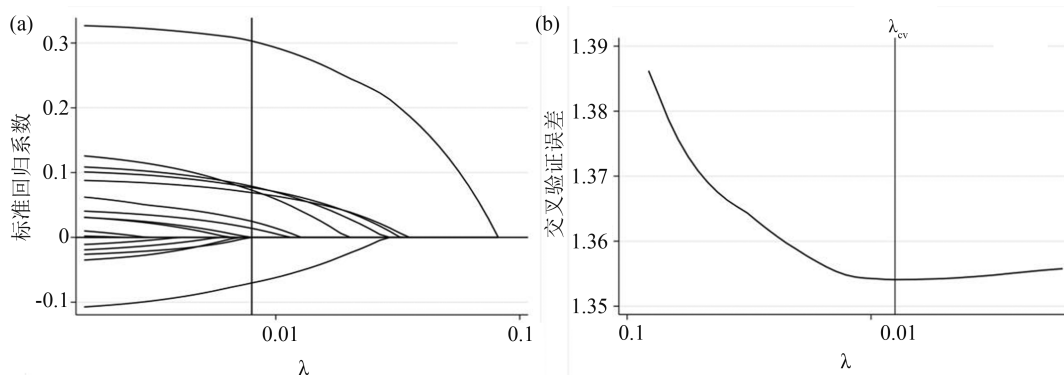


图 1 LASSO-logit 回归中影响因素的系数分布图和交叉验证图

3. 流动人口医疗服务利用的多因素分析

根据 LASSO 回归的分析结果,以流动人口医疗服务利用为因变量(未就诊=0,就诊=1),将筛选出的 8 个影响因素纳入二元 logit 回归。结果显示,是否工作、家庭月收入、居住地到最近医疗机构距离、接受健康教育状况、健康状况对流动人口的医疗服务利用具有显著影响(表 2)。其中,与无工作的人群相比,有工作的流动人医疗服

务利用率较低($OR=0.791$);家庭月收入越高,医疗服务利用率越高($OR=1.040$);居住地到最近医疗机构在 15~30 min 的居民比距离更近的居民利用率更高($OR=1.368$);接受较多健康教育可以提升医疗服务利用率($OR=1.034$);健康状况较差的人群的医疗服务利用率更高,与自评为健康的人群相比,基本健康和不健康但生活能自理的利用率分别提升了 79.8%和 156.5%。

表 2 内蒙古流动人口医疗服务利用行为的多因素分析

| 变量 | β | S.E. | Wald | P | OR | 95% CI |
|----------------------------|---------|-------|--------|-------|-------|--------------|
| 民族(汉族为参照) | | | | | | |
| 蒙古族 | -0.041 | 0.162 | -0.240 | 0.807 | 0.960 | 0.689~1.336 |
| 其他 | 0.443 | 0.423 | 1.630 | 0.103 | 1.557 | 0.914~2.651 |
| 是否独居(否为参照) | 0.373 | 0.342 | 1.580 | 0.114 | 1.452 | 0.915~2.304 |
| 是否工作(否为参照) | -0.234 | 0.075 | -2.460 | 0.014 | 0.791 | 0.657~0.953 |
| 家庭月收入 | 0.039 | 0.014 | 2.880 | 0.004 | 1.040 | 1.013~1.069 |
| 居住地到最近医疗机构距离(15 min 以内为参照) | | | | | | |
| 15 min~30 min | 0.313 | 0.148 | 2.890 | 0.004 | 1.368 | 1.106~1.691 |
| 30 min~1 h | 0.466 | 0.500 | 1.480 | 0.138 | 1.593 | 0.861~2.947 |
| 1 h 以上 | -0.699 | 0.372 | -0.930 | 0.350 | 0.497 | 0.115~2.152 |
| 居民健康档案建立(未建立为参照) | 0.187 | 0.118 | 1.910 | 0.056 | 0.187 | 0.995~1.460 |
| 接受健康教育状况 | 0.033 | 0.016 | 2.190 | 0.029 | 1.034 | 1.003~1.065 |
| 健康状况(健康为参照) | | | | | | |
| 基本健康 | 0.587 | 0.189 | 5.580 | 0.000 | 1.798 | 1.463~2.210 |
| 不健康,但生活能自理 | 0.942 | 0.396 | 6.090 | 0.000 | 2.565 | 1.894~3.472 |
| 生活不能自理 | 1.120 | 2.602 | 1.320 | 0.187 | 3.066 | 0.581~16.179 |

讨 论

1. 内蒙古流动人口医疗服务利用行为亟待改善

就诊率是衡量医疗服务利用行为的重要指标。内蒙古流动人口的患病就诊率较低,患病后选择不就医的比例达到 49.98%,远低于国内其他地区^[11-12],主动利用医疗服务的意愿亟待改善。多数流动人口的文化水平较低、缺乏专业技能,往往流动过程中从事劳动保护不足、工作环境较差的工作,因此流动人口的患病风险相对较高^[13]。流动人口的快速增长作为快速城镇化过程中产生的社会现象,其健康状况值得关注。内

蒙古流动人口中接近 80%为省内流动,但内蒙古地理上东西跨度大,中西部地区的社会特点接近晋语区的山西、陕西等省,东部地区与东北三省较为相似,生活习惯与社会风俗存在较大差异^[14],导致流动人口的生活适应和社会融入存在一定难度,进而影响医疗服务的利用。因此建议地方政府加强对于流动人口的重视,以基层居民自治组织和基层医疗卫生机构为着手点,通过家庭医生随访、开展健康知识宣传讲座等方式加强对流动人口健康状况的关注。

2. 使能资源能够有效改善流动人口医疗服务利用行为

通过提供必要的使能资源可以促进患者更好地利用医疗服务。研究结果显示,是否工作、家庭月收入、医疗机构距离和健康教育对流动人口医疗服务利用行为具有显著影响。首先,有工作的流动人口医疗服务利用率相对较低。虽然工作可以带来更高的收入,但也意味着更高的时间机会成本,若工作的人群选择因病请假往往会影响到其收入或晋升^[15],而多数流动人口不固定的收入会进一步降低其就医的可能性。此外,流动人口受限于其文化程度,从事体力劳动的比例较高。内蒙古的第三产业相对欠发达,在煤炭、钢铁等传统工业中提供了大量工作岗位,这类行业存在工作时间长的特点,进而压缩了就医的可能性。建议工作单位可以提供合理的工作时间和更好的福利政策,政府和社会可以给予灵活就业人员更多的社会保障,鼓励健康的工作习惯和减少工作压力,在落实每周最长工作时间、双休日等法律制度的基础上,鼓励带薪休假制度。其次,家庭月收入与流动人口医疗服务利用存在正相关关系。低收入家庭在遇到健康问题时会出于经济问题的考虑选择忽视治疗^[16],在一定程度上加剧了基本医疗服务的不均衡。因此,建议继续深化推进基本医疗保险覆盖和质量,加强异地就医结算的范围和宣传力度,探索城乡居民基本医疗保险开展门诊统筹的合理性,降低流动人口就医负担。

居住地与医疗机构的距离可能会影响医疗服务利用行为。通常情况下医疗机构距离越远,患者越不愿意前往就医^[17]。然而距离在不同患者主观判断中的价值不同,可能存在某个临界值。在该范围内不会显著影响医疗服务利用意愿,超出范围就医意愿则会大幅降低。内蒙古地广人稀,医疗机构难以有效覆盖全部居民,可以通过互联网医疗、家庭医生签约等形式增加医疗服务利用率。最后,健康教育会显著提升流动人口医疗服务利用率。一方面流动人口的健康意识相对较低,健康教育可以提升其医疗知识,促使其关注健康;另一方面有效的健康教育可以促进预防和早期治疗^[18]。因此,政府可以通过各类渠道加强健康宣传和教

3. 健康状况对于流动人口医疗服务利用具有显著影响

研究结果显示,健康状况对于医疗服务利用行为有显著的影响。健康状况差的人需要更频繁地使用医疗服务^[19],以维持其健康状况和生活质量。相反,健康状况良好的人则需要较少的医疗服务。但与此同时,健康状况还会影响医疗服务利用的能力。健康状况差的人可能由于身体状况不佳而难以使用医疗服务。相反,健康状况良好的人则更容易使用医疗服务。

此外,健康状况还会影响健康意识。健康状况差的人更倾向于寻求医疗服务,这可能是由于其对健康的关注更高,对健康问题的敏感度更高^[20]。相反,健康状况良好的人则更倾向于避免使用医疗服务,可能是由于其对健康状况的信心更强,对健康问题的担忧程度较低。因此,在制定医疗服务政策和计划时,需要考虑到健康状况对于医疗服务利用的影响,特别是对于那些健康状况较差的人,可以通过增加医疗资源、提高医疗服务质量等措施为他们提供更多的医疗服务和支持,以确保他们能够获得所需的医疗服务。对于健康状况较好的人,鼓励其积极参与健康促进活动,从而提高健康水平和减少疾病风险。

参 考 文 献

- [1] 李国平,孙瑀.以人为核心的新型城镇化建设探究[J].改革,2022,346(12):36-43.
- [2] 朱宇,林李月,柯文前.国内人口迁移流动的演变趋势:国际经验及其对中国的启示[J].人口研究,2016,40(5):50-60.
- [3] 姚强,李寒旋,杨菲.医疗保险参保地对我国流动人口卫生服务利用和健康状况影响及对策研究:一个范畴综述[J].中国卫生事业管理,2022,39(9):666-671+720.
- [4] 尹上岗,王晨歌,马志飞,等.中国流动人口医疗保险参保空间格局及其影响因素:基于全国流动人口动态监测调查数据的分析[J].地域研究与开发,2019,38(1):155-161.
- [5] 赖玉清,汪洋,毛菊,等.西藏地区中老年人就医行为及其影响因素研究[J].现代预防医学,2023,50(1):134-138.
- [6] 冯利英,吴新娣.十年来内蒙古地区迁移流动人口研究[J].内蒙古财经学院学报,2005(4):105-109.
- [7] 陈鸣声.安德森卫生服务利用行为模型演变及其应用[J].南京医科大学学报(社会科学版),2018,18(1):5-8.
- [8] 秦瑶,韩红娟,陈杜荣,等.基于LASSO logistic 回归模型的轻度认知障碍逆转预测模型[J].中国卫生统计,2022,39(5):653-658.
- [9] 张秀秀,王慧,田双双,等.高维数据回归分析中基于LASSO的自变量选择[J].中国卫生统计,2013,30(6):922-926.
- [10] 孙红卫,杨文越,王慧,等.惩罚logistic 回归用于高维变量选择的模拟评价[J].中国卫生统计,2016,33(4):607-611.
- [11] 张小娟,郝春彭.居民医保门诊保障水平对医疗服务利用和费用的影响[J].卫生经济研究,2022,39(11):23-26.
- [12] 许新鹏.城乡医保统筹对农村流动人口医疗服务利用的影响:基于2017年CMDs数据的分析[J].中国卫生政策研究,2022,15(3):53-60.
- [13] 张培芳,郑思思,王宇潇,等.2013—2018年我国流动人口两周患者未就诊率的meta分析[J].职业卫生与病伤,2018,33(3):136-139.
- [14] 杨春燕,赖彦斌,刘目兴.内蒙古现代化进程因子分析与地域差异[J].海南师范学院学报(自然科学版),2002(Z1):96-100.
- [15] 王振,曾智.退休对医疗卫生服务利用的影响:基于断点回归的实证研究[J].现代预防医学,2022,49(16):2992-2997.
- [16] 张瀚月,李卉,杨超,等.泸州市农村高血压患者就医行为及影响因素分析[J].重庆医学,2019,48(2):305-309.
- [17] 谢铮,徐玲.基于计划行动理论的我国农村居民就医行为选择分析[J].北京大学学报(医学版),2010,42(3):270-274.
- [18] 徐畅.医疗可及性对老年人健康以及医疗服务利用的影响研究[D].北京:首都经济贸易大学,2019.
- [19] 周钦,田森,潘杰.均等下的不公:城镇居民基本医疗保险受益公平性的理论与实证研究[J].经济研究,2016,51(6):172-185.
- [20] 王翌秋,张兵.农村居民就诊单位选择影响因素的实证分析[J].中国农村经济,2009,290(2):77-85.

(责任编辑:邓妍)