

## • 专论:用户画像在医疗健康领域的应用 •

编者按: 用户画像作为一种重要分析工具, 能有效利用海量数据挖掘用户特征、预测用户行为和发现用户潜在需求。在“健康中国”战略下, 全民健康意识大幅提升, 用户健康信息需求个性化、多样化特征愈发显著, 用户画像在医疗健康领域的应用也越来越广泛。通过用户画像对人体特征信息、健康状况和健康行为等可视化展现, 可以为个性化健康管理提供更完整丰富的信息, 还可以根据患者健康需求变化进行动态更新, 具有精准度高、时效性强等特点, 在精准医疗推荐、健康管理、疾病预警和疾病诊治等方面应用前景广阔。本期专论着眼于用户画像在医疗健康领域的应用, 所载论文包括我国大城市居民参与互联网在线问诊的行为画像分析、基于文本挖掘的互联网医疗平台用户画像模型构建、健康画像的数据模型设计与智能化应用等, 以期为提高医疗质量和效果、满足用户健康需求提供参考。

# 我国大城市居民参与互联网在线问诊的行为画像分析\*

吕英杰<sup>1</sup> 马子婷<sup>1</sup> 胡雅欣<sup>1,2</sup>

(<sup>1</sup> 北京化工大学经济管理学院 北京 100029 <sup>2</sup> 北京中医药大学东方医院 北京 100078)

[摘要] 目的/意义 研究大城市居民选择互联网医疗服务时的典型特征, 建立患者行为画像, 为改进互联网医疗服务提供决策依据。方法/过程 以北京市患者为研究对象, 从个体属性特征、在线问诊行为以及择医需求偏好 3 个维度构建患者群体画像, 通过聚类方法对在线患者群体分类。结果/结论 在线患者可划分为价值潜力型、平台忠诚型、潜在转化型、需求强烈型和服务转移型 5 类不同特征群体。总体上来自大城市的在线患者病例级别偏低且活跃度不高, 通常使用价格较低的图文问诊方式, 倾向于选择来自三甲医院、职称较高且在线服务质量较高的医生。5 类患者群体各具典型特征, 需要分别采取针对性措施满足其个性化医疗需求。

[关键词] 互联网医疗; 大城市居民; 用户画像

[中图分类号] R-058 [文献标识码] A [DOI] 10.3969/j.issn.1673-6036.2024.06.001

## Analysis of Behavioral Portraits of Urban Residents Participating in Internet Online Consultations in China

LYU Yingjie<sup>1</sup>, MA Ziting<sup>1</sup>, HU Yaxin<sup>1,2</sup>

<sup>1</sup>School of Economics and Management, Beijing University of Chemical Technology, Beijing 100029, China; <sup>2</sup>Dongfang Hospital, Beijing University of Chinese Medicine, Beijing 100078, China

[Abstract] **Purpose/Significance** To study the typical characteristics of urban residents when they choose internet medical services, and to establish patient behavior portraits, so as to provide decision-making basis for improving internet medical services. **Method/Process** Taking patients in Beijing as the research object, patient group portraits are constructed from three dimensions: individual attrib-

[修回日期] 2024-06-18

[作者简介] 吕英杰, 博士, 教授, 发表论文 30 余篇。

[基金项目] 北京市社会科学基金项目 (项目编号: 19GLC064)。

ute characteristics, online consultation behavior and medical demand preference, and online patient groups are classified by the clustering method. **Result/Conclusion** Online patients are classified into five distinct groups: value potential, platform loyalty, potential conversion, strong demand, and service transfer. Overall, online patients from major cities tend to have lower case severity and lower levels of activity. They typically prefer lower - cost text - based consultations and show a preference for doctors from top - tier hospitals with higher professional titles and superior online service quality. These five different patient groups possess their own typical characteristics, and specific measures need to be taken to meet their personalized medical needs.

[**Keywords**] internet medical service; residents in major cities; user portrait

## 1 引言

近年来“互联网+”与医疗健康行业的深度融合,有效缓解了医疗资源供需矛盾及配置结构失衡问题<sup>[1]</sup>。在线医疗服务打破了传统就医模式时间和地点的限制,有效避免患者就医过程中可能面临的交叉感染风险,成为许多人看病就医的优选方式<sup>[2]</sup>,在线医疗服务平台用户访问量及问诊量呈现激增趋势<sup>[3]</sup>。中国互联网络信息中心第 53 次《中国互联网络发展状况统计报告》显示,截至 2023 年 12 月我国互联网医疗用户规模达 4.14 亿人,占我国网民规模总量的 37.9%<sup>[4]</sup>。

当前互联网医疗存在地区发展不均衡等问题<sup>[5]</sup>,其中大城市医疗水平相对较高,尤其以三甲医院为代表的优质医疗资源更为丰富,是互联网医疗服务的主要提供者<sup>[6-7]</sup>。同时大城市互联网应用更加普及,居民医疗健康素养较高,成为互联网医疗服务的主要使用者<sup>[8]</sup>。因此有必要研究在大城市传统医疗资源和线上医疗资源均较为发达的情况下,居民在选择互联网医疗服务时具有哪些行为特征,这不但有利于医疗服务平台为用户提供更优质的服务,提升用户体验,同时也有利于更好地解决医疗资源供需矛盾,缓解大城市医疗资源供给不均和优质医疗资源浪费等问题。

本文选取北京市为我国大城市的典型代表,以参与互联网医疗服务的北京市患者作为研究对象,通过收集患者参与互联网医疗服务的行为数据,了解患者行为特征,统计分析患者个人属性特征及择医需求,并利用用户画像技术构建北京市患者行为画像,以便进一步明确大城市居民参与互联网在线问诊服务时的行为特征与需求,助力互联网医疗服务健康发

展,对健康中国建设也具有重要价值和意义。

## 2 资料来源和方法

### 2.1 资料来源

研究资料来自好大夫在线医疗服务平台,好大夫在线为患者提供在线咨询问诊、医院/医生信息查询、线下门诊预约、疾病知识科普等各种互联网医疗服务,截至 2024 年 6 月,该平台收录来自全国 10 561 家正规医院的 929 190 万名医生,其中有 28 万名公立医院医生实名注册,累计服务超过 8 400 万例患者。本研究数据来自于该平台提供的数据开放平台服务,该服务提供 2019 年患者在该平台参与在线问诊的 104 万条脱敏数据记录。从中筛选所在地为北京市的患者及其在线问诊信息,共得到 10.7 万条样本数据。由于疾病类型众多,根据《医疗机构诊疗科室目录》中的 37 个一级诊疗科目,选取问诊量排名前 10 位的疾病为研究对象,最终得到 14 145 条记录作为实验样本。各疾病类型人数分布情况,见图 1。

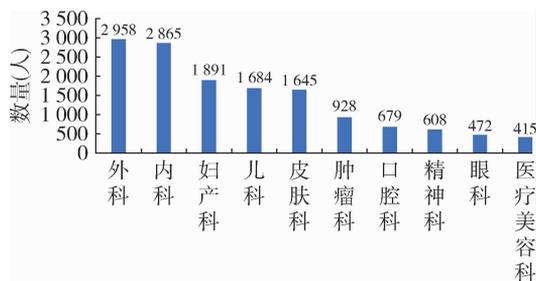


图 1 患者疾病类型分布

### 2.2 研究方法

为构建参与互联网问诊服务的患者用户群体画像,从个体属性特征、在线问诊行为及择医需求偏

好3个维度进行分析。每个维度包含的各项指标及说明,见表1。其中个体属性特征包括患者病例级别和在线活跃度两个指标;在线问诊行为会因平台运营模式和服务功能不同而存在显著差异<sup>[9]</sup>,相关文献<sup>[10]</sup>指出平台提供的问诊功能类型、是否提供免费服务以及收费问诊价格,都会显著影响患者在线问诊决策,因此在该维度选取问诊形式、收费类型和问诊价格3个关键指标深入剖析;择医需求偏好分为医生个体属性及医生在线行为两方面,医生个

体属性包括医生职称和医院级别,医生在线行为包括医生提供在线服务的数量和质量。进一步通过聚类分析方法,基于个体属性特征和在线问诊行为两个维度下的5个指标,对患者群体进行特征分类,并分析不同类别群体问诊需求和在线问诊行为典型特征。由于本研究实验所选聚类指标既包含分类变量又包含连续变量,样本容量较大,也缺乏事先明确的聚类数目,因此采用二阶聚类法处理实验样本。

表1 患者群体画像各维度指标

维度	指标名称	数据类型	描述
个体属性特征	病例级别	有序分类变量	疾病严重程度由轻到重分为4个等级,分别用1、2、3、4表示
	活跃度	连续变量	患者在线问诊次数
在线问诊行为	问诊形式	无序分类变量	1=图文问诊,2=电话问诊,3=预约门诊
	收费类型	二分类变量	1=免费,2=收费
	问诊价格	连续变量	在线问诊支付价格均值
择医需求偏好	医生职称	二分类变量	1=高级职称,2=非高级职称
	医院级别	二分类变量	1=三甲医院,2=非三甲医院
	服务数量	连续变量	医生在线问诊量
	服务质量	连续变量	医生诊后报到患者数

### 3 结果

#### 3.1 患者群体画像特征

统计分析3个维度各项指标数据后得到北京市患者参与互联网问诊服务的用户群体画像,见图2。

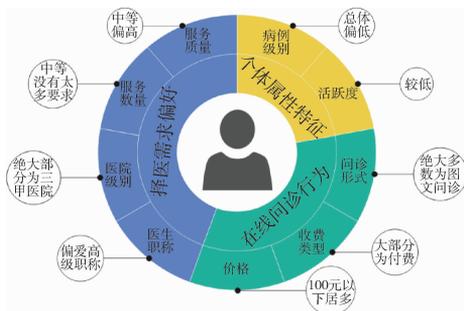


图2 北京市患者参与互联网问诊服务的用户画像

3.1.1 个体属性特征方面 患者病例级别总体偏低且活跃度不高。病例级别为第2等级的患者占比最高,病例级别为1和3的次之,轻微症状患者占样本总体的66%,与生活在北京等一线城市居民

实际情况相符。由于大城市居民生活压力较大、节奏较快,医疗资源总量虽然较大但人均依然不足,患者病情较轻时难以付出足够时间和精力及时到医院就诊,更倾向于选择在线医疗服务平台问诊。患者参与在线问诊服务活跃度方面,仅有16%的患者为活跃人群,表明大多数居民并不频繁使用在线问诊服务,也从侧面反映出在线医疗服务还未被广泛接受和认可,取代传统医疗模式还需要在可信度和满意度等服务质量方面进一步提升。

3.1.2 在线问诊行为方面 患者倾向于选择收费较低的图文问诊形式。图文问诊占有所有问诊形式的75%,由于该类服务不受空间限制,医患双方采用异步沟通方式,具有很大灵活性和便利性,与大城市快节奏、高效率的生活更相匹配。而电话问诊(12%)和预约门诊(13%)人数相对较少。因为电话问诊更适用于情况较为紧急时的医患沟通,往往费用较高。预约门诊功能使用较少则与大城市医疗资源较为丰富有关,居民获取优质医疗资源往往更为便利,因此对于平台提供的预约门诊功能需求

并不强烈。分析在线平台用户付费情况发现，只有 1/4 患者选择免费问诊，充分说明大部分患者对于在线医疗服务有一定认可度并愿意为之支付费用。但从付费金额分布来看，超过一半的患者支付费用在 50 元以下，约 80% 患者付费金额在 100 元以下，表明大部分城市居民对于选择在线医疗服务仍有各种顾虑，经过综合考虑各种因素后，其通常不太愿意支付高昂费用使用在线医疗服务。

3.1.3 择医需求偏好方面 患者倾向于选择来自三甲医院和具有较高职称且在线服务质量较高的医生。选择三甲医院的医生占比高达 85%。这一方面说明大城市居民对优质医疗资源具有更高的需求和期待，另一方面也体现了由于优质医疗资源可及性更好，其对在线医疗资源也抱有更高的标准和要求。分析医生提供的在线问诊服务时，发现患者更看重医生

过往提供在线服务的质量而非数量。特别是患者诊后报到数等指标反映了医生在既往服务中获得其他患者的认可情况，患者更倾向于选择已在过去展现出高专业水准的医生，以获得更可靠的医疗建议。

### 3.2 基于聚类分析的患者群体分类

通过二阶聚类分析发现，当聚类数目为 6 时，具有较小的贝叶斯信息准则 (Bayesian information criterion, BIC) (23 633.377)，较大的 BIC 变更比例 (0.283)，以及最大的距离测量值比例 (1.768)。因此可以得出最佳聚类数目为 6，聚类品质为优。通过进一步分析每个聚类群体在各聚类指标上的显著特征，可以总结各类别人群特征倾向，其中由于类 1 和类 5 各方面属性特征较为类似，将二者合并，最终形成 5 类用户群体，不同类型群体典型特征，见图 3。

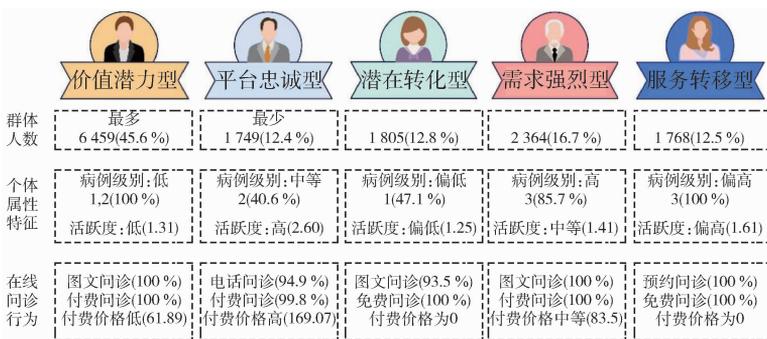


图 3 各类别患者群体用户画像

群体 1 被定义为“价值潜力型”，占比 45.6%，典型特征：样本数量最多，虽然患病程度最低，但仍选择在线付费问诊，表明这类群体对自己健康状况比较关注，对在线问诊服务消费潜力较大，具有较高潜在价值。群体 2 被定义为“平台忠诚型”，占比 12.4%，典型特征：付费平均值以及活跃度均为所有群体中最高的，说明其对在线医疗服务的忠诚度和依赖度较高，对在线医疗服务持续发展至关重要。群体 3 被定义为“潜在转化型”，占比 12.8%，典型特征：在所有在线问诊形式中皆选择免费问诊，存在从免费问诊向付费问诊转化的潜力。群体 4 被定义为“需求强烈型”，占比 16.7%，典型特征：相比于其他群体来说病情普遍较重，大部分病例级别为 3、4，并且选择的问诊方式是图文

问诊，对在线问诊服务具有较强烈的需求。群体 5 被定义为“服务转移型”，占比 12.5%，典型特征：通过在线问诊方式和医生交流的主要目的是预约线下门诊服务，可以作为线上问诊服务的潜在用户，通过采用一定措施和手段可以吸引该类用户实现从线下门诊到线上问诊服务的转移。

## 4 讨论

### 4.1 “价值潜力型”群体

该类群体患者人数约占样本的一半，表明该类人群在以北京市为代表的大城市患者中分布最广泛，如果能够吸引该类型群体持续使用在线医疗服务，将会为互联网医疗持续健康发展带来示范效

应。该群体病例级别较低,表明所患疾病严重程度不高,但在问诊形式上均选择付费问诊,说明其对在线医疗服务具有一定信任和认可,未来也很有可能继续付费使用在线医疗服务,具有较大消费潜力。但值得注意的是当前该群体平均付费金额较低,一方面是由于病例级别较低,消费金额与病例级别相匹配;另一方面也表明该群体对在线医疗服务还有不少顾虑,不希望为在线问诊服务支付较高价格。在问诊功能选择上偏好使用图文问诊,这种异步沟通方式也更适用于病例级别较低等病情不急情况。此外,该群体的在线活跃度也处于低等水平,在线问诊频率较低,说明对在线医疗服务的依赖程度不强。针对这类群体,在线平台可以进一步研究其择医需求偏好,实现个性化定制推送,改善用户体验,激发其价值潜力。在线平台在运营过程中也可以针对该类用户发放优惠券和推送折扣信息,激励其提升在线活跃度的同时提高消费金额。

#### 4.2 “平台忠诚型”群体

该类群体在总样本中占比最少,一定程度上说明患者总体上对在线医疗服务的认可度和忠诚度相对较低,在线医疗服务在以北京市为代表的一线城市中也远未建立较好的用户基础。该群体付费平均值最高且活跃度达2.60,也是所有群体中最活跃的一类,充分表明其愿意为在线医疗服务花费金钱,也意味着其对于自身健康问题非常重视,并且相信在线平台提供的医疗服务水平,对在线医疗平台具有较高的忠诚度和依赖度。该群体病例级别平均为中等,其中病例级别为2的患者最多,占该群体人数的40.6%,其余各疾病等级人群在该类中均有分布,也从侧面反映无论病情是否严重都会存在一定比例的患者对在线医疗服务具有较高的信赖。问诊形式方面,该类群体采用电话问诊比例达到94.9%,说明其更喜欢通过直接沟通方式与医生立即联系并交流病情,这也与其更重视自身病情且擅长使用在线服务有显著关联。

针对这类群体,在线医疗服务平台可以采取措施进一步提高其对平台的认可度,不断优化用户体验,如通过积分、等级、设置VIP等方式提高用户

的在线身份价值,激励这些用户持续使用在线问诊服务,保持、继续提升忠诚度。

#### 4.3 “潜在转化型”群体

该类群体主要参与免费问诊服务,意味着其不愿意为在线医疗服务付费,原因可能是刚接触在线医疗服务,对在线平台信任度不高,或者对平台提供的各项服务功能还不够了解。也可能是对所患疾病认识不足,只想通过平台提供的免费问诊服务进一步了解自身病情,以便进一步决定是否继续使用在线付费问诊服务或另做打算。在病例级别方面,该类人群中病例级别为1的占47.1%,其次是病例级别为2(43.7%)的,说明该群体病例级别总体偏低,病情较轻。此外,该群体偏好使用图文问诊,在线活跃度也偏低,原因可能是其自身病情并不严重因而没有过多重视,这与该群体倾向于选择免费问诊而不愿意支付费用的行为特征密切相关。针对这类群体,在线医疗服务平台应尽可能挖掘并满足用户问诊需求,通过设置平台使用引导等方式帮助其更好地了解平台功能。还可以通过各种运营手段触达用户,为其提供体验高质量付费医疗服务的机会,或者发放定向优惠及折扣等,促进该群体实现从免费用户到付费用户的转化。

#### 4.4 “需求强烈型”群体

该类群体均为付费用户,虽然付费平均值在所有群体中处于中等水平,但在问诊价格定位较低的图文问诊群体中,该群体的付费平均值是最高的,说明其对于付费问诊的接受度很高,对在线医疗服务具有较高信任度并且愿意为这种服务支付更高费用。由于该群体中有高达85.7%的患者病例级别为3,病例级别为4的占14.7%,且样本中绝大部分病例级别为4的患者都归入该群体,说明该群体是所有群体中病例级别最高的一类,因此更注重自身的健康状况,愿意为在线问诊服务投入更多成本,以期能够获得高水平医生的高质量诊疗建议,从而更好地解决自身疾病问题。鉴于该群体在病情较严重情况下仍然选择在线问诊服务,并愿意为此支付较高价格,充分说明其对在线医疗服务具有较为

强烈的需求。针对这类群体，在线平台应将服务重点放在如何满足患者对高质量医疗资源的迫切需求上，采取措施为患者推送更适合其病情的医生资源，提高医患匹配度，并提供更为完善的沟通交流工具为双方高质量在线沟通提供保障，这些措施将进一步增强患者对在线医疗服务的信心，提升患者满意度水平，促使其更愿意为在线医疗服务付费。

#### 4.5 “服务转移型”群体

该类群体在线问诊行为主要是为线下预约门诊服务，其在线行为表现为免费问诊且消费金额为 0。该类人群之所以没有选择线上问诊，一方面可能是因为其在线医疗平台的具体功能并不了解；另一方面也可能是因为其线上问诊服务仍有顾虑，认为线上问诊的可信度不高，无法真正满足其医疗需求。该群体中所有人的病例级别均为 3，处于病情较严重水平，更可能出于疾病的风险考虑而偏好选择到医院线下就诊。该群体活跃度为 1.61，在所有群体中属于较高水平，也反映出其非常关注自身病情，且愿意花费较多时间与医生在线交流，因此具备从传统门诊服务转移到线上问诊服务的潜力。针对这类群体，平台应不断完善自身功能设计，提供更便捷的在线问诊功能、更有效的医患沟通工具和渠道，吸引用户多参与在线问诊服务，如针对患者病情特点和需求偏好推送健康科普知识、推荐医生直播讲座等，不断增强患者对在线平台的信任和依赖感，推动其将医疗需求从线下服务转移到线上服务。

## 5 结语

本文以北京市患者为研究对象分析大城市居民选择互联网医疗服务的典型特征，研究成果不仅有利于医疗服务平台为大城市患者提供更优质的服务，提升患者体验，优化完善服务，也有助于更好地发挥互联网医疗服务在解决大城市医疗资源供需矛盾、缓解大城市医疗资源供给不均和优质医疗资源浪费等问题中的重要价值。本文存在一定不足之处有待完善。首先，在样本选择方面，后续考虑纳入北京市之外更多大城市和地区的患者用户，

提高研究结果代表性和推广性。其次，考虑采用多种不同类型在线医疗服务平台，拓展研究范围和数据来源，获得更准确和全面的实验结果。最后，受数据可用性限制，仅基于 2019 年患者实验样本开展研究，未来将考虑纳入最新数据，反映患者在线问诊行为的趋势变化。

**作者贡献：**吕英杰负责提出研究思路，设计研究框架，论文修订及定稿；马子婷负责确定研究方案，分析处理实验数据，对模型进行实验验证，撰写初稿；胡雅欣负责文章质量控制与审查，参与文章修订。

**利益声明：**所有作者均声明不存在利益冲突。

## 参考文献

- 1 宋玉娟, 伍林生. 健康中国战略背景下互联网医院发展的 SWOT 分析 [J]. 中国卫生事业管理, 2021, 38 (7): 499-502.
- 2 王涛, 翟兴. 突发公共事件应急管理情报需求支持机制研究——以互联网医疗健康业务模式演进为例 [J]. 情报科学, 2020, 38 (7): 3-8, 44.
- 3 刘刊, 周宏瑞, 侯月婷. 共享医疗平台如何实现价值共创? ——一个探索性单案例研究 [J]. 管理评论, 2022, 34 (11): 337-352.
- 4 中国互联网络信息中心. 第 53 次《中国互联网络发展状况统计报告》[EB/OL]. [2024-03-22]. <https://www.cnnic.net.cn/n4/2024/0322/c88-10964.html>.
- 5 邓朝华, 邓子豪, 樊国睿, 等. 在线健康平台能否减少健康不公平? ——基于价值共创视角 [J]. 管理世界, 2024, 40 (1): 119-140.
- 6 许荣华, 张婷婷, 周根贵. 我国互联网医疗资源的分布及区域差异研究 [J]. 浙江工业大学学报 (社会科学版), 2021, 20 (3): 319-324.
- 7 熊丹妮, 夏晨曦, 李燊, 等. 中国在线医疗服务医疗机构及医生的分布与启示 [J]. 中国卫生事业管理, 2018, 35 (7): 481-484, 506.
- 8 吴江, 周露莎. 医疗信息资源跨地区流动: 在线医疗社区优化医疗资源配置作用的研究 [J]. 信息资源管理学报, 2017, 7 (4): 58-65, 85.
- 9 马骋宇. 在线医疗社区服务利用及转化研究——以好大夫在线为例 [J]. 中国卫生政策研究, 2016, 9 (11): 70-73.
- 10 曹仙叶, 刘嘉琪. 基于服务多样性视角的在线医疗社区患者选择决策行为 [J]. 系统管理学报, 2021, 30 (1): 76-87.