

# 健康画像的数据模型设计与智能化应用\*

范春<sup>1</sup> 马洁<sup>1</sup> 徐安琪<sup>1</sup> 徐一涵<sup>1</sup> 史金龙<sup>2</sup>

(<sup>1</sup> 卫宁健康科技集团股份有限公司 上海 200072 <sup>2</sup> 宁夏医科大学 银川 750004)

**[摘要]** 目的/意义 探索健康画像的数据模型设计与智能化应用。方法/过程 分析健康画像应用现状及问题,规范健康画像数据采集、立法与治理,系统性建立健康画像数据模型,探索健康画像潜在应用场景。结果/结论 设计形成健康画像数据模型,增强了健康画像的全面性、针对性;建立数据模型管理系统,提升了模型规范化管理效率;总结“健康管理和评估”“精确诊断和个性化诊疗”两大健康画像应用场景,帮助提高医疗质量和效果,满足个体健康需求。

**[关键词]** 健康画像;数据模型设计;数据模型管理;智能化应用;知识图谱

**[中图分类号]** R-058 **[文献标识码]** A **[DOI]** 10.3969/j.issn.1673-6036.2024.06.003

## The Data Model Design and Intelligent Applications of Health Portrait

FAN Chun<sup>1</sup>, MA Jie<sup>1</sup>, XU Anqi<sup>1</sup>, XU Yihan<sup>1</sup>, SHI Jinlong<sup>2</sup>

<sup>1</sup>Winning Health Technology Group Co. Ltd., Shanghai 200072, China; <sup>2</sup>Ningxia Medical University, Yinchuan 750004, China

**[Abstract]** **Purpose/Significance** To explore the data model design and intelligent applications of health portrait. **Method/Process** The paper analyzes the current situation and problems of health portrait applications, standardizes the data collection, legislation and governance of health portrait, systematically establishes the data model of health portrait, and explores potential application scenarios for health portrait. **Result/Conclusion** The data model of health portrait is designed to enhance the comprehensiveness and pertinence of health portrait. The data model management system is established to improve the standardized management efficiency of the model. Two application scenarios of health portrait are summarized, namely “health management and assessment” and “accurate diagnosis and personalized diagnosis and treatment”, which help to improve the quality and effect of medical treatment, and meet individual health needs.

**[Keywords]** health portrait; data model design; data model management; intelligent applications; knowledge graph

## 1 引言

“健康中国 2030”战略下,医防融合等内涵建

设加速了我国医疗卫生服务从只关注治疗到更关注预防保健的理念转变<sup>[1]</sup>。2022 年 4 月发布的《“十四五”国民健康规划》指出形成“病前主动防,病后科学管,跟踪服务不间断”的一体化健康管理服务<sup>[2]</sup>。在公立医院高质量发展、医疗数字化转型等驱动下,智慧医院建设由以功能为导向逐步转变为以数据赋能为导向,注重基于真实世界数据和真实世界证据的医疗健康服务,健康画像<sup>[3]</sup>成为为患者提供个性化、针对性服务的重要载体。

健康画像通过提供能够识别或反映特定自然人生

**[修回日期]** 2024-01-17

**[作者简介]** 范春,高级工程师,发表论文 13 篇;通信作者:史金龙。

**[基金项目]** 宁夏回族自治区重点研发计划(项目编号:2022BEG01001)。

理或心理健康的相关电子数据的量化或图形化表达,直观反映个体健康状况,是疾病防治、个性化诊疗和精准医疗发展的必然趋势<sup>[4]</sup>。针对健康画像的数据模型设计、管理以及基于健康画像的智能化应用场景,是数字化医疗健康领域关注度较高的主题。

## 2 现状分析

健康画像的临床应用价值逐渐得到研究认可。Facio F M 等<sup>[5]</sup>早在 2010 年证实家庭健康画像工具准确评估遗传性家族史的能力,但半自动化的数据采集限制使其并未完全替代传统的遗传风险评估。近年来伴随电子病历、健康管理 App、可穿戴设备相继兴起,国内利用统计学、数据挖掘、机器学习等多方法推进健康画像在疾病诊疗、健康管理等领域的研究与实践逐渐增多。刘海鸥等<sup>[6]</sup>研究表明基于深度学习的健康画像模型能够准确识别用户的抑郁情感和抑郁状态,从而辅助治疗抑郁症。王佳琪等<sup>[7]</sup>利用健康画像开展慢性心力衰竭患者健康管理、加强心血管疾病患者服药管理、建立心血管疾病风险预测模型。刘莉等<sup>[8]</sup>以基于知识的标签提取技术结合电子病历数据挖掘证实健康画像能够为慢性阻塞性肺疾病的诊断、评估、病情监测、疾病预后、管理等提供数据支撑。

然而,健康画像在医防融合、精准医疗等领域的积极作用尚未得到充分发挥。一是健康画像的数字化表征方法、层次和内涵受限,数据模型泛化能力不足。这是由于缺乏系统性设计,画像组成要素缺失,画像与支撑数据之间的相关性、因果性不明确,健康画像的全面性、针对性不足,难以适应多场景应用。二是存在的数据质量问题对健康画像的质量、效率产生负面影响。现有数据治理方法更关注对历史数据的清洗、转换,忽略数据立法工作,针对数据元的定义和管理机制缺失使数据的一致性、逻辑性和完整性难以保证,导致健康画像出现偏差和误判。三是对健康画像模型管理的智能化手段不足,画像从数据元到数据模型的映射机制尚未建立,数据定义过程烦琐而低效,模型管理的规范性和效率难以保证。

## 3 健康画像的系统性建立

### 3.1 目标

聚焦健康画像的数字化表征方法,系统性建立健康画像数据模型并实现模型的智能化管理,通过全面梳理相关数据指标、强化数据立法与治理、分析数据变化对健康影响的关联度和因果性,解决现有健康画像组成要素缺失、数据质量不稳定、针对性及可持续性不足等问题,以期健康画像应用场景的拓展奠定技术基础;在数据即服务、模型即服务的机制下,分析挖掘健康画像的潜在智能化应用场景,为开展医疗健康各项业务拓展基于真实世界数据的可适配、可交互应用服务,以期强化健康画像在个人健康管理、精准医疗等领域的价值作用,提高数字化医疗服务水平。

### 3.2 健康画像数据采集、立法与治理

健康画像的数据涵盖预防、保健、体检、治疗、康复等领域,来自不同的平台或系统,涉及多系统间数据传输通信。不同类别数据更新的频度也不同,如个人身份信息、既往病史、家族史等为相对静态的信息;主诉、症状、生理指标、医疗服务信息等为动态信息,往往需要 T+1 (即数据产生后的一天内更新)或实时更新。应制定数据采集标准技术规范,明确各类数据的采集字段、采集来源、接口规范、采集频率等要求。采用抽取、转换和加载(extraction-transformation-loading, ETL)模式从核心业务抽取所需的数据字段至 ETL 前置机,或者采用数据接口模式从信息集成平台直接读取所需数据字段,实现真实世界数据采集。

引用《卫生健康信息数据元目录》《卫生健康信息数据元值域代码》《电子病历基本数据集》《城乡居民健康档案基本数据集 WS 365—2011》《基本信息基本数据集个人信息 WS 371—2012》《疾病管理基本数据集》等标准,通过数据立法,制定健康画像数据字典,对健康画像数据元进行标准化定义,详细规范数据元的名称、数据类型、标识格式,并统一定义数据元编码标识符号,从源头保证

数据的一致性、准确性和完整性。

依据数据治理制度的要求，建立健康画像数据治理体系。从数据目录管理、数据标准管理、数据采集交换、数据质量控制、数据质量评估等多维度对健康画像数据开展治理，实现大规模多源异构数据的标准化处理和全过程数据质量控制，从而实现数据标准落地、数据管理受控、数据质量追踪、数据可视可用，有力促进健康画像数据的标准化管理、汇集融合和开放共享。

### 3.3 健康画像的数据模型设计

3.3.1 健康画像数据的三维关系 健康画像可由“事件/活动/医疗健康服务”“个人健康状况/医疗服务结果”“时间”3个维度的数据构成，见图1。“事件/活动/医疗健康服务”是发生在个人身上的与健康相关的事件、活动或医疗健康服务记录，如

一次骨折、摔倒、心梗、卒中、抢救等事件，一次具体的运动、饮食构成、社交沟通等活动，一次手术、检查、治疗、康复、心理干预等医疗健康服务，是对过程的记录；“个人健康状况/医疗服务结果”是对个人当前各类健康指标、疾病特异性标志、治疗结果、健康评估和预测警示等信息的记录，是对健康结果或评估预测信息的记录；“时间”是围绕个体全生命周期，对过程和结果数据打上时间标签，形成健康画像在时间维度上的快照。在该三维关系中，每个健康结果或评估预测都有相应的时间属性和过程记录与之关联，形成数据支撑。通过对健康结果、时间属性和过程记录的分析处理，可以提取个体在不同时间点的健康状况、健康变化趋势和健康事件之间的相关性等信息，对健康画像与数据之间的相关性和因果性进行数字化表征，形成全面、准确、动态更新的个体健康画像。

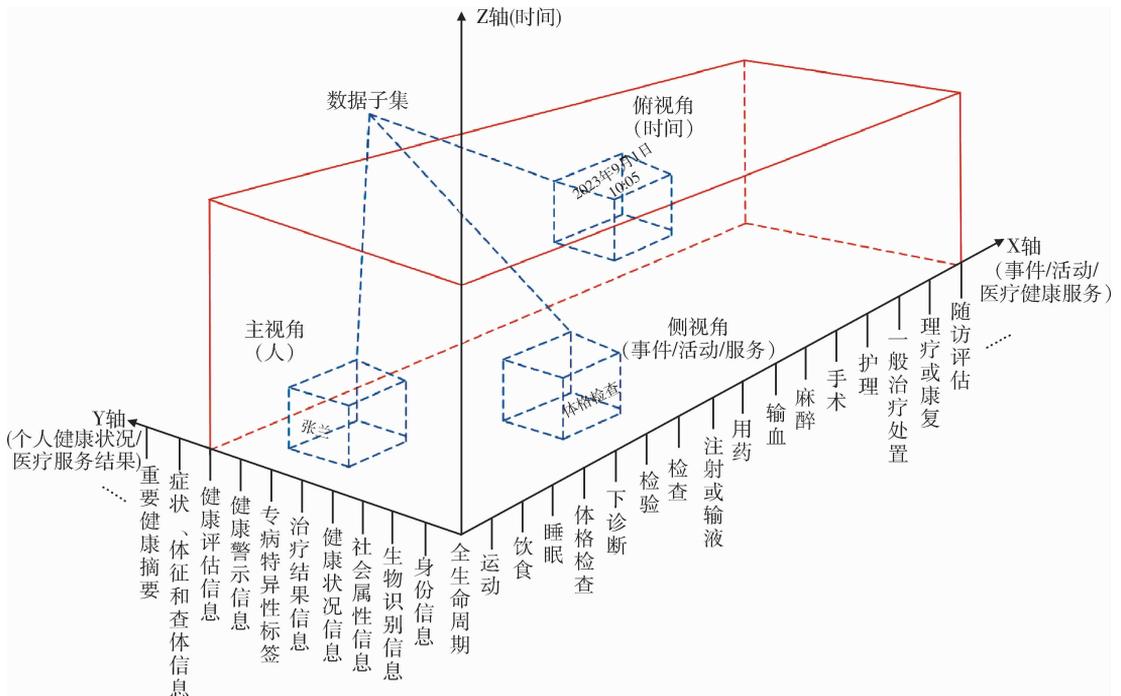


图1 健康画像数据的三维关系

3.3.2 健康画像的数据模型 为保证健康画像的系统性、全面性，以患者为中心梳理个体健康指征和健康影响因子数据，构建健康画像数据模型，分

析健康现状、健康影响的相关性及因果性，预测健康趋势。数据模型包括患者身份信息、生物识别信息、社会属性信息等，见表1。

表 1 健康画像数据模型

数据类别	内容
身份信息	姓名、身份证号、医保卡号、联系电话等
生物识别信息	基因、指纹、声纹、虹膜、面部特征等
社会属性信息	出生日期、性别、民族、国籍、职业、住址、家庭成员信息、联系人信息、婚姻状态等
一般信息	性别、体重、身高、年龄、血型等
健康状况信息	主诉、现病史、现存主要健康问题、既往病史、体格检查（体征）、家族史、过敏史、症状、生理指标数据、遗传咨询数据、残疾类型、残疾等级、残疾部位等
生活方式信息	运动、饮食、营养、睡眠、社交、嗜好、居住环境等描述
职业暴露危险因素	粉尘、放射性物质、其他有毒物质等信息
医疗服务信息	门（急）诊病历、住院医嘱、检查检验报告、用药信息、病程记录、手术记录、麻醉记录、输血记录、护理记录、入院记录、出院小结、转诊（院）记录、知情告知信息、康复训练信息等
公共卫生服务信息	免疫接种、健康体检、疾病筛查、随访信息等
疾病特异性标签	检验检查数据、基因测序、转录产物测序、蛋白质分析测定，代谢小分子检测、人体微生物检测等

在健康画像整体数据模型框架下，还可以针对某些特定的主题如手术、护理、康复、躯体功能、营养、衰弱、睡眠、精神、多病共存、风险、多重用药等设计相应的模型子集，用以对个人健康状况进行多维度、专业化的描述与评估，警示和干预健康风险。如多重用药存在一定的安全隐患，临床上可根据模型视图对多重用药进行动态管理和审查，根据常见病、基础病当前所用的所有药物，对药物的副作用、药物之间可能产生反应或代谢过程中的相互作用进行分析，最大限度地减少不必要用药，保证用药安全，优化治疗方案，同时，提高患者的认知能力，减少对药物的依赖性。

不同模型之间还可建立关联或因果关系，见图 2。有脑梗基础疾病的患者做了一次髌部骨折手术，髌部骨折手术者为静脉血栓栓塞症高危人群，考虑术后 1~3 天易发生术后静脉血栓栓塞症，可通过药物提升血流速度，该脑梗患者平时也服用疏血通改善血液循环，但上述药物会增加术后出血风险，因此用药需综合权衡利弊。考虑多种措施联合使用，术后 1~3 天以足踝运动为主，加强康复运动进行物理预防，3 天后改用药物预防和物理预防相结合，这样既预防了术后静脉血栓栓塞症，也降低

了术后出血的风险。上述实例需在手术、多病共存、多重用药、风险和康复 5 个模型之间建立相互关联关系，通过模型可视化使临床医生更直观地诊疗和决策。

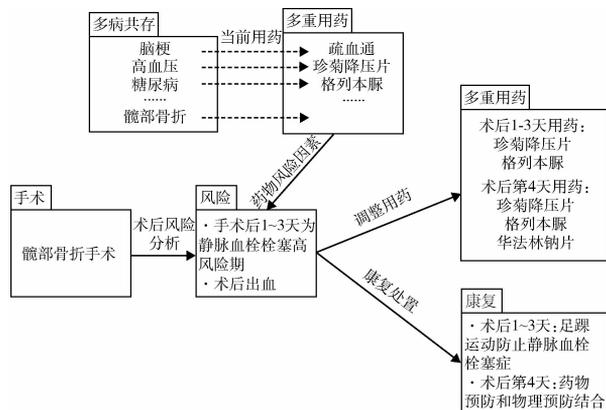


图 2 不同模型之间的关联关系

3.3.3 数据模型的智能化管理 建立标准化、可视化、智能化的数据模型管理系统，对健康画像数据模型进行统筹管理，提升模型规范化管理的效率，通过数据中台和微服务实现数据即服务、模型即服务。数据模型的智能化管理包括建模预览、模型管理和版本管理，见图 3。

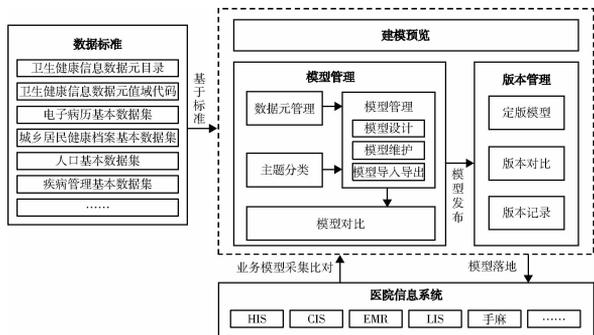


图 3 数据模型的智能化管理

其中，模型管理是数据模型智能管理系统的核心，借助信息化、智能化手段解决普遍存在的数据血缘关系和模型之间关联关系复杂、模型适应性差和扩展性弱、维护工作量大等突出问题。模型设计基于数据标准，通过数据元管理支撑数据模型设计与维护。主题分类是对临床业务标准的落地，通过创建分类并执行分类新增、删除、编辑、状态变更等操作，可以对目标画像的数据模型进行自定义设计，实现对健康画像数据模型的扩展，增强健康画像的完整性、适应性、针对性和精准性。版本管理中模型对比通过自动化比对定版模型与业务库中采集的模型，实现对模型数据的查漏补缺或基于真实世界数据的模型修正，快速稽核模型，提升模型的符合度和管控效率。

## 4 健康画像的应用场景分析

### 4.1 概述

基于数据中台实现数据即服务、模型即服务，健康画像可呈现个体健康的总体视图，包括疾病筛查、疾病管理、诊疗记录、体征、用药、健康评估、健康管理等板块。同时，借助微服务方式调用，健康画像可与健康相关的其他模型（如运动、饮食、营养、检验、检查等）融合建立满足不同需求场景的细分版块及细分视图，实现健康画像的可拆可合，丰富了画像内涵，增强了画像的适用性和完整性。

### 4.2 健康管理和健康评估

健康视图可以针对某个模型版块详细展开。健

康管理板块可针对当前主要疾病、疾病预测、体征指标、体征达标率、体征趋势、健康危险因素、健康实施率等主题展示和分析。基于对个人健康状态的监测，健康画像还可以帮助慢性病、亚健康、老年人等群体针对个体健康状况制定相应的健康管理计划，包括饮食、运动、用药、睡眠、心理疏导等，变被动的健康维护为主动的健康管理，更加有效地控制疾病风险，改善患者健康状况。针对高血压、糖尿病、脑卒中等老年慢性病，健康画像可融合生成式人工智能技术进行个性化居民健康评估，精准标识异常部位，并基于异常指标项提供针对性健康评价，包括健康风险提示、评估结论等。

### 4.3 精确诊断和个性化治疗

健康画像能够辅助医生更深入地分析多模态医学数据、更全面地掌握患者的健康情况，基于健康画像可进一步开发更智能的诊断、建议治疗策略和预后预测等工具。基于多维度临床数据和疾病知识图谱，将全量基础数据集成、治理和开发，转化为大数据宽表，通过机器学习、特征分析等方法，构建面向不同医疗诊断场景的疾病辅助诊断模型，实现临床诊疗关键信息的医学逻辑判定和数字医学循证，得到基于真实世界数据的临床辅助诊断，包括疾病诊断、分型、阶段和确诊疾病危险度分层等。按照筛查、诊断、治疗、康复、随访的全诊疗过程，针对个案提供分阶段、个体化、连续性的智能诊疗方案推荐，包括诊断方案、手术方案、药物方案、康复方案等，进行诊疗效果预测、诊疗风险提示等。

## 5 结语

基于健康画像的三维关系，以患者为中心系统性建立健康画像数据模型，实现对健康画像数据指标的结构化、规范性描述和智能化管理，解决当前健康画像组成要素缺失、数据质量不稳定、针对性及可持续性不足等痛点，有助于提升健康画像数据模型的生命力，保障健康画像的系统性、全面性和针对性。通过健康画像的智能化应用，可以实现患

者健康状态的可视化监测、健康风险预测,同时也为患者健康管理、健康干预与临床诊疗提供决策支持,提高医疗质量和效率。

然而,本研究聚焦于健康画像的泛在模型设计与场景应用,面对不同的疾病专科数据,后续可针对某一专科或病种对健康画像进行更为精细的描绘,从而更好地支撑临床辅助决策及个性化健康管理;同时,伴随数据治理、湖仓一体、批流一体、人工智能等技术的发展,健康画像的数字化表征方法将不局限于单一技术的选择,多技术融合满足健康画像对大规模、多模态、高质量数据的要求,提高健康画像的精准度和应用价值。

**作者贡献:** 范春负责论文整体框架设计和内容撰写;马洁负责数据治理方法设计;徐安琪负责数据模型设计;徐一涵、史金龙负责应用场景分析。

**利益声明:** 所有作者均声明不存在利益冲突。

## 参考文献

1 “健康中国 2030”规划纲要 [EB/OL]. [2023 - 10 - 25]. [https://www.gov.cn/gongbao/content/2016/content\\_5133024.htm](https://www.gov.cn/gongbao/content/2016/content_5133024.htm).

- “十四五”国民健康规划 [EB/OL]. [2023 - 04 - 27]. [https://www.gov.cn/gongbao/content/2022/content\\_5695039.htm](https://www.gov.cn/gongbao/content/2022/content_5695039.htm).
- 文庭孝,刘晓琦. 国内外健康画像研究综述 [J]. 医学信息学杂志, 2023, 44 (3): 2 - 9.
- 张冬云,刘晓琦,吴旭生等. 基于人体特征信息标签的个体健康画像研究 [J]. 医学信息学杂志, 2023, 44 (3): 10 - 17.
- FACIO F M, FEERO W G, LINN A, et al. Validation of my family health portrait for six common heritable conditions [J]. Genetics in medicine: official journal of the American college of medical genetics, 2010, 12 (6): 370 - 375.
- 刘海鸥,姚苏梅,何旭涛,等. 基于深度学习的在线健康社区抑郁症用户画像研究 [J]. 小型微型计算机系统, 2021, 42 (3): 572 - 577.
- 王佳琪,王一焱,李沪生,等. 用户画像技术在心血管疾病健康管理的应用现状 [J]. 中国慢性病预防与控制, 2022, 30 (11): 868 - 872.
- 刘莉,朱勋梅,李俊,等. 健康画像在慢阻肺个性化健康管理中的应用研究 [J]. 中国医学物理学杂志, 2020, 37 (7): 918 - 926.

## 《医学信息学杂志》编辑部严正声明

近期,有不法人员冒充《医学信息学杂志》编辑部工作人员,以核对收录信息、审核数据、发送录用通知等名义,微信要求添加好友或发邮件,进而收取稿件处理费等。以上行为严重侵害了广大作者、读者及本刊的权益,编辑部保留追究其法律责任的权利。本刊特此严正声明:(1)《医学信息学杂志》暂不通过微信进行正式沟通,未委托任何个人或机构代理收稿、征稿等业务,唯一投稿渠道为杂志官网 <http://www.yxxxx.ac.cn> 在线投稿,点击“作者投稿”按钮注册后即可投稿。杂志官方邮箱为 [yxxxx@imicams.ac.cn](mailto:yxxxx@imicams.ac.cn) 和 [yxxxxzz01@163.com](mailto:yxxxxzz01@163.com),不会通过其他邮箱发送通知或对外联系。(2)《医学信息学杂志》是中国知网、万方数据、维普网全文收录期刊。期刊出版后 1 个月左右可在上述数据库查阅论文。(3)《医学信息学杂志》编辑部目前收取版面费仅有对公转账一种方式。开户行:中国建设银行北京雅宝路支行;开户名称:中国医学科学院医学信息研究所;银行账号:11001028400059856368。

敬请广大作者、读者提高警惕,请勿向任何个人账户支付任何费用,以免造成不必要的损失。必要时建议通过报警等方式维护正当权益。遇到任何问题可拨打联系电话 010 - 52328672/86/87 与编辑部取得联系。

《医学信息学杂志》编辑部