

省级健康医疗大数据平台设计与应用实践^{*}

路 杰 姚进文 闫宣辰 刘红亮 高 歆

(甘肃省卫生计生统计信息中心 兰州 730030)

〔摘要〕 阐述甘肃省健康医疗大数据平台建设目标、业务需求、数据来源、业务内容、系统架构及部署,介绍平台应用实践,包括跨地区患者流向分析、构建人工智能决策引擎等方面,对进一步完善平台、建立健康医疗大数据模型等进行探讨。

〔关键词〕 健康医疗; 大数据平台; 设计与实现

〔中图分类号〕 R-056 〔文献标识码〕 A 〔DOI〕 10.3969/j.issn.1673-6036.2018.08.002

Design and Application Practice of Provincial Health Care Big Data Platform LU Jie, YAO Jin-wen, YAN Xuan-chen, LIU Hong-liang, GAO Xin, Statistical Information Center, Health and Family Planning Commission of Gansu, Lanzhou 730030, China

〔Abstract〕 The paper dilates on the building target, business requirement, source of data, content of business, system architecture and deployment of the health care big data platform of Gansu province, introduces the application practice of the platform, including analysis of trans-regional flow of patients, building of the Artificial Intelligence (AI) decision-making engine and other aspects, as well as discusses the further perfection of the platform and the establishment of the health care big data model, etc.

〔Keywords〕 Health care; Big data platform; Design and realization

1 引言

健康医疗大数据泛指人类在社会生活中因健康活动产生的相关数据,包含整个健康活动过程及其产生的结果,贯穿整个生命周期^[1]。健康医疗大数据除具备维克多·迈尔舍恩伯格提出的大量、多样、高速、价值 4 大特征外,还具备真实、实时、冗余、隐私等特点。健康医疗大数据的挖掘、应用

在人类健康、医疗卫生管理领域展示出广阔的前景。精准医疗、健康管理、辅助科研、临床决策支持、医疗保障监管、医药研发等都是健康大数据应用的探索方向^[2]。2016 年 6 月国务院办公厅印发《关于促进和规范健康医疗大数据应用发展的指导意见》,将健康医疗大数据应用发展纳入国家大数据战略布局,部署通过“互联网+健康医疗”探索服务新模式、培育发展新业态,努力建设人民满意的医疗卫生事业,为打造健康中国提供有力支撑^[3]。甘肃省随即开展健康医疗大数据应用研究工作,建成甘肃省健康医疗大数据应用平台。

〔修回日期〕 2018-04-21

〔作者简介〕 路杰,高级工程师,发表论文 10 余篇。

〔作者简介〕 甘肃省科技厅软科学项目“甘肃省规范和促进健康医疗大数据应用发展研究”(项目编号:17CX1ZA008)。

2 健康医疗大数据平台设计

2.1 建设目标

通过建设健康医疗大数据平台,汇集公共卫生、医疗服务、医疗保障、药品管理、计划生育、综合管理、临床研究、基因组学、转换医学等方面涵盖个人全生命周期的健康医疗大数据,推动医疗健康业务与服务模式转型升级,进一步深化医疗卫生体制改革,提升健康医疗服务效率和质量,不断满足人们多层次、多样化的健康需求,培育新的业态和经济增长点。

2.2 业务需求

甘肃省健康医疗大数据平台主要基于已建成的省、市、县 3 级全民健康信息平台、基础数据库(电子病历、健康档案、全员人口数据库)和已与全民健康信息平台互联互通的各类业务信息系统进行设计。通过模型再造进一步提升医疗服务效率和质量、变革医疗模式、完善统计制度,加强统计分析能力,实现现代科学化控费和精算管理,提高服务能力和管理水平,推进健康医疗行业治理、临床和科研、公共卫生大数据的应用,加强健康医疗数据安全保障和患者隐私保护,发展智慧健康,提供医疗便民服务。

2.3 数据来源及分组

按照健康医疗数据产生的来源,将构建以下几个分类的大数据中心:(1)临床大数据。包括电子病历(Electronic Medical Records, EMR),医学影像,患者终生就医、住院、用药记录,标准化临床路径。(2)健康大数据。包括电子健康档案(Electronic Health Records, EHR),监测个人体征,个人偏好,康复医疗,健康知识数据等。(3)生物大数据。包括不同组学的的数据,如基因组学、代谢组学、蛋白组学等。(4)经营运营大数据。包括成本核算,医药、耗材、器械采购与管理,不同病种治疗成本与报销,药物研发,消费者消费,产品流

通,第 3 方支付数据等。

2.4 业务内容

2.4.1 辅助临床决策 运用海量临床病历数据在短时间内分析可能的病因及结果,提供患者互动、临床护理、诊断、研究、数据可视化等服务,依据与疗效相关的临床、病理等特征为医生提出规范化临床路径及个性化治疗建议。

2.4.2 辅助监管决策 通过整合全员人口、基层医疗机构信息、药品、医疗行为、公共卫生绩效考核、重点工作监管、医疗综合数据、综合业务数据等多类业务应用数据信息,为医疗、医药、医保监管工作提供精确的决策支持,使决策者有针对性地制定措施,更加有效率地解决问题,分析不同级别医疗机构中的医疗需求,合理配置医疗资源,为医疗机构、医务人员、医疗行为的新“三医”监管提供数据支撑。

2.4.3 精准医疗 基于个人健康档案记录的全生命周期的健康、医疗数据,利用体征指标偏离以及临床诊疗数据向用户提供更精准科学的健康保养及疾病预防建议,为医生提供更精确的病因、发病历程以及诊疗后药效、康复数据等,向个人提供个性化、精准、有效的健康管理计划和医疗方案。

2.4.4 实现分级诊疗效果科学评估 通过大数据从全局掌控分级诊疗全省异地就诊患者流向,直观揭示患者就诊现状,将患者的多个特征进行比对,获取信息的差异点并按差异性进行排序分类,分析出区域患者全维度的特征,从而客观地反映分级诊疗政策实施后的效果,为医疗资源优化配置、分级诊疗工作的决策与评估、分级诊疗政策的进一步完善提供科学依据。

2.5 系统架构

甘肃省健康医疗大数据平台以 2 级以上医院信息系统、公共卫生业务系统、妇幼系统、疾控系统和其他平台的数据为资源池,由数据交易、大数据管理、基础支撑、智慧医疗分析 4 大平台共同构建全省医疗大数据平台,系统架构,见图 1。

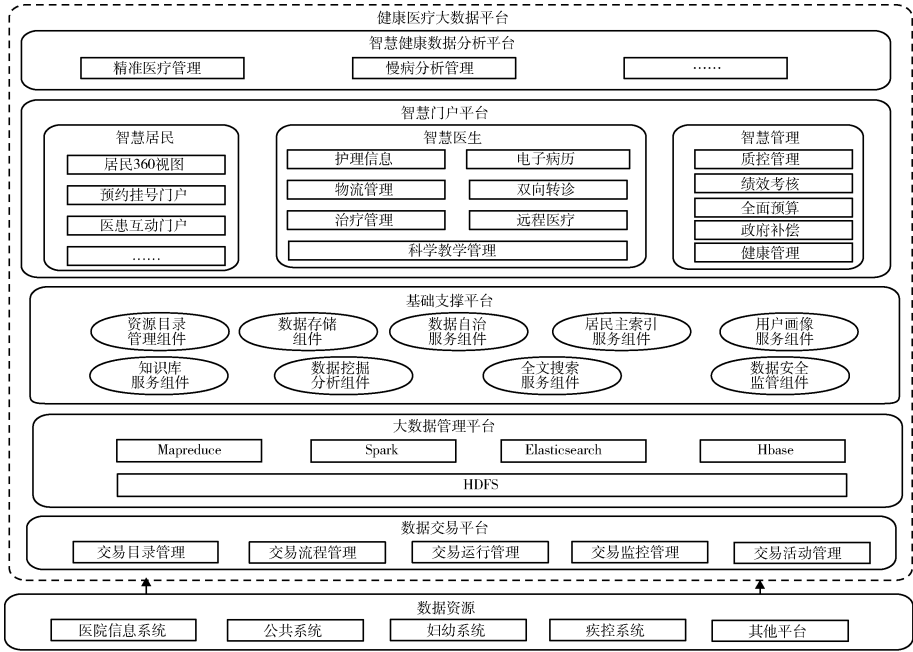


图 1 健康医疗大数据平台系统架构

2.6 系统部署

大数据平台系统部署是面向服务架构（Service-oriented Architecture, SOA），基于“3+2”的云计算数据中心。“3”指基础设施层、平台层和软件服务层 3 个层次，“2”指安全支撑体系、运维管理体系。“3”信息化基础资源共享服务体系，每层既对上一层提供服务支撑，同时又具有独立的面向业务支撑的应用服务体系，体现基础设施即服务、平台即服务和软件应用即服务的云计算系统部署，见图 2。

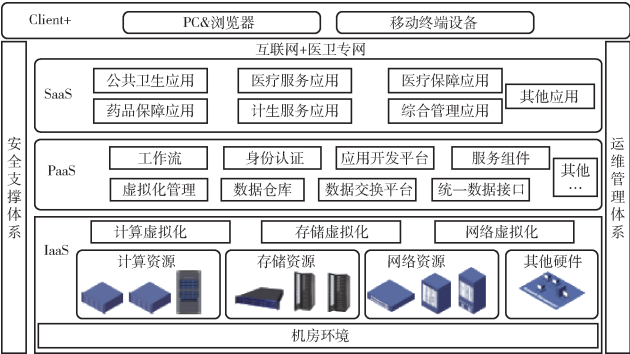


图 2 健康医疗大数据平台系统部署

3 应用实践

3.1 跨地区患者流向分析

利用大数据为支撑，以用户身份证号作为信息采集依据，分析患者在诊疗过程中发生的跨地区就诊活动的具体原因及流向地等，利用这些数据为合理安排医疗资源提供数据支撑，跨地区患者源向分析，见图 3。通过选择年份、地区、流向等可动态显示选择地区患者流向的相关数据，并且支持各州市数据下钻，为医疗监管部门提供最清晰直观的数据分析。

3.2 超级健康档案生成及应用

目前大部分医院信息系统已实现与健康档案系统对接，医生在问诊时调阅患者的健康档案信息以帮助诊疗。但在实际工作中由于医生用于每位患者的门诊诊疗时间非常有限（约 3 分钟），往往没有时间调阅患者的健康档案信息，即便调阅也很难在短时间内对杂乱无章的健康档案信息做出有效判断。超级健康档案生成是依托健康医疗大数据平台，以患者身份证号、当次医院就诊卡号和门诊或住院信息作为数据采集和整合的依据，将患者健康

档案信息中与本次就诊相关的数据进行自动抽取和分析,当门诊患者到达医生诊室,医生刷取该名患者就诊卡号后,系统自动生成与该患者本次就诊相关的超级健康档案信息(包括既往疾病史、家族疾病史、本次就诊可能的疾病分析)推送给医生,帮助其快速、准确地进行诊疗。

3.3 构建医联体人工智能决策引擎,助推分级诊疗实施

数据统计显示现今约有 10% 的患者需要到大型医院就诊,大部分慢性病、常见病应在社区医院完成治疗。目前患者就医数据能够被完整准确地记录,有利于构建治疗和预防、康复和自我保健管理一体化的电子云服务。患者可随时在社区就诊,也可在就近的医疗机构接受治疗,疑难重症则可通过远程方式解决,以此实现分级诊疗。该引擎主要通过智能疾病诊断相关分组(Diagnosis Related Groups, DRGs)大数据驱动分级诊疗,有效解决医疗资源不足的问题。通过医联体内数据共享、决策优化提升运营利润的同时,提高医院资源使用率。数据显示“HIS + DRGs + 大数据 + 人工智能 + 商业保险”的完整闭环引擎能提高医院设备使用率 25% 以上,药物成本降低 20% 以上,节省医院 18% 的投资和运维成本,降低 27% 以上的人力成本。

3.4 脑卒中发病预警建模

脑卒中是我国人群主要的死亡原因之一。中国 2013 年死于脑血管疾病的城镇居民约为 83.58 万人,农村居民约为 101.2 万人^[4]。美国弗雷明汉心脏研究所(Framingham Heart Study, FHS)建立和改良的脑卒中风险评估模型(Framingham Stroke Risk Profile, FSRP)是目前国外应用最广泛的脑卒中风险评估工具^[5-6],也是被美国心脏协会和美国脑卒中协会的脑卒中预防指南推荐使用的脑卒中风险评估手段^[7]。甘肃省健康医疗大数据平台的脑卒中发病预警风险评估采用改良后的 Framingham 模型。即用年龄、收缩压(治疗或未经治疗)、有无

糖尿病、是否吸烟、有无心血管病史、有无房颤及左心室肥大作为估计变量。利用前沿的机器学习算法,基于全省两万人卒中大数据进行强化训练,构建新型 Gs - Framingham - DLR 模型,提升脑卒中适应性和预测准确性,该模型达到国际领先水平。

3.5 基于深度学习的 ICU 急重症转移预警建模

深度学习在大数据应用方面备受关注,模型能够通过组合底层的特征形成高层的数据抽象特征进而模拟人类大脑认知机制,且自动发现隐藏在数据中的模式^[8]。该模型主要基于回归分析 ICU 急重症病历,对 2 级以上医院 3 年的临床队列数据进行深度学习模型训练,采集住院患者心率、体温、血压、血氧等 37 项生命体征数据进行 48 小时滑动窗口的实时监测,对患者的 ICU 重症及死亡风险进行预警。该模型系统得到临床数据的支持,大数据分析干预可以显著减少患者 ICU 重症转移率,经过智能决策预警将重症转移率从 27% 降到 18%,对重症预防及死亡率的降低有显著效果。

3.6 健康扶贫大数据开发及应用

基于全民健康信息平台的健康扶贫大数据分析主要由健康扶贫综合分析、综合疾病分析、综合评价、聚类分析和“3 个一批”分析 5 大部分组成。其中健康扶贫综合分析主要动态展示因病致贫人口分布、病种排名、疾病治疗费用构成比例、治疗机构分类情况等内容,健康扶贫大数据综合展示,见图 4。综合疾病分析主要展示因病致贫患病病种、先天性疾病占比情况以及因病致贫县区分布、年龄分布、费用构成情况等,见图 5。精准扶贫、精准治病是健康扶贫工作的基础和重点。在健康扶贫工作中利用健康扶贫大数据平台,以大数据为引领,积极探索用数据甄别、决策、管理、考核的精准扶贫方式开展健康扶贫工作,使扶贫措施精准到位,同时运用大数据分析全面、动态地掌控医保基金在健康扶贫中发挥的作用和支付比例,保障健康扶贫精准到位、最大程度发挥医保资金的使用效益。



图3 跨地区患者流向分析



图4 健康扶贫大数据综合展示



图5 健康扶贫因病致贫疾病情况大数据展示

4 讨论及建议

4.1 提高数据质量

数据质量关系到后续数据的分析、利用和展

示。全民健康信息平台与相关的各类业务系统对接，数据标准、接口规范还有待进一步完善，部分数据质量不高，下一步需通过考评机制监督各相关系统开发商和医院信息系统厂商完善接口规范、优化数据上传流程，向全民健康信息平台持续不断地

上传所需数据,逐步提高平台数据质量。

4.2 探索建立符合我国国情的健康医疗大数据模型

目前国家、省、市、县 4 级全民健康信息平台相继建成,海量的健康医疗大数据储备、快速的数据生产和共享、数据结构化能力的提升以及效果显著的人工智能算法使得更广泛、更深入的健康医疗大数据应用成为可能。以大数据分析为基础,物联网服务运营平台为依托,实现个性化健康管理将成为未来健康产业的发展趋势和突破口。我国应积极探索建立符合国情的健康医疗大数据模型,进一步提高医疗效率,改进服务质量,降低安全风险,减少医疗成本。

4.3 加强健康医疗大数据安全

在做好健康医疗大数据应用的同时,要按照严谨细致的工作要求,严格落实主体责任制,强化防范管理、监测预警和应急处置等工作措施,全力确保信息系统不受篡改、入侵、窃密等攻击,坚决防止重大网络安全事件的发生。

4.4 培养健康医疗大数据人才和开展跨学科研究

人才是提高核心竞争力的关键因素,健康医疗大数据分析应用需要具备计算机、数据统计分析、医学、管理等多方面知识的人才,这样的复合型人才在我国较少,与快速发展的健康医疗大数据应用不相适应,因此在高校应建立大数据分析相关专业,以此加强大数据理论与基础建设,培养高端的大数据分析人才。此外大数据分析是一项多学科、跨专业的研究工作,需要不同领域的专业技术人员相互协作,充分利用各自的专业知识来共同挖掘海

量数据的潜力,完成共同目标。

5 结语

健康医疗大数据正迅速地改变人们的生活方式,通过对海量医疗大数据的分析处理、挖掘潜在价值,推进医疗卫生机构治理模式创新,降低医疗服务成本,减少医疗资源浪费,实现医疗卫生服务精准化、高效化。持续推进“互联网+医疗”,使就医更便捷,不断提升健康医疗和公共卫生服务能力,进而提高公众生命质量。

参考文献

- 1 徐志祥,王莹.我国医疗行业大数据应用现状及政策建议[J].中国卫生信息管理杂志,2017,14(6):822-823.
- 2 孟群.大健康大数据大融合大发展[J].中国卫生信息管理杂志,2017,14(3):379.
- 3 国务院办公厅.关于促进和规范健康医疗大数据应用发展的指导意见[EB/OL].[2017-06-21]:http://www.gov.cn/zhengce/content/2016-06/24/content_5085091.htm.
- 4 国家心血管病中心.中国心血管病报告 2014[M].北京:中国大百科全书出版社,2015.
- 5 Wolf PA, D, Agostino RB, Belanger AJ, et al. Probability of Stroke: a risk profile from the Framingham Study[J]. Stroke, 1991, 22(3): 312-318.
- 6 D'Agostino RB, Wolf PA, belanger AJ, et al. Stroke Risk Profile: adjustment for antihypertensive medication. The Framingham Study[J]. Stroke, 1994, 25(1): 40-43.
- 7 Goldstein LB, Bushnell CD, Adams RJ, et al. Guidelines for the Primary Prevention of Stroke: a guideline for health-care professionals from the American Heart Association / American Stroke Association[J]. Stroke, 2011, 42(2): 517-584.
- 8 张娟,汪西莉,杨建功.基于尝试学习的形成建模方法[J].计算机学报,2018,41(1):134-135.