

基于数据整合模式的基层医疗卫生机构智能报表填写模型构建路径研究*

吴卓存 张重阳 胡红濮 万艳丽 赵姗姗 曾庆嘉

(中国医学科学院/北京协和医学院医学信息研究所 北京 100020)

[摘要] **目的/意义** 探索面向基层医疗卫生机构的智能自动化模型构建路径, 解决报表重复上报问题。**方法/过程** 运用现场调研和专家咨询等方法, 在北京市东城区基层医疗卫生机构开展实地调研, 通过信息资源规划方法、数据整合与映射技术等, 梳理基层医疗卫生机构业务交互与信息流动情况, 研究自动化报表模型的构建路径。**结果/结论** 建模梳理业务流、数据流, 整理报表相关指标库, 构建多源数据自动映射模型与规则, 为实现基层医疗卫生机构智能报表提供具有可行性的参考路径。

[关键词] 基层医疗卫生; 报表; 数据整合; 自动映射

[中图分类号] R-058 **[文献标识码]** A **[DOI]** 10.3969/j.issn.1673-6036.2024.05.006

Study on the Construction Path of an Intelligent Reporting Model for Primary Healthcare Institutions Based on Data Integration Patterns

WU Zhuocun, ZHANG Chongyang, HU Hongpu, WAN Yanli, ZHAO Shanshan, ZENG Qingjia

Institute of Medical Information, Chinese Academy of Medical Sciences & Peking Union Medical College, Beijing 100020, China

[Abstract] **Purpose/Significance** To explore the construction pathway of an intelligent automation model tailored for primary healthcare institutions, aiming to address the issue of repetitive reporting. **Method/Process** Through methods such as on-site investigations and expert consultations, a field study is conducted in primary healthcare institutions in Dongcheng District, Beijing. Utilizing information resource planning methods and data integration and mapping technologies, the business interactions and information flow within these institutions are analyzed to investigate the construction pathway of an automated reporting model. **Result/Conclusion** The business flow and data flow are mapped out by the modeling process, a repository of relevant reporting indicators is organized, and a multi-source data automatic mapping model and rules are developed. The study provides a feasible reference pathway for realizing intelligent reporting in primary healthcare institutions.

[Keywords] primary health care; report form; data integration; automatic mapping

[修回日期] 2024-03-28

[作者简介] 吴卓存, 硕士研究生; 通信作者: 万艳丽, 博士, 副研究员。

[基金项目] 中国医学科学院医学与健康科技创新工程经费资助项目 (项目编号: 2022-I2M-1-019); 国家自然科学基金重大项目 (项目编号: 22&ZD141)。

1 引言

基层医疗卫生机构作为我国分布最广、服务人群最多的卫生健康机构,承担着基本公共卫生、基本医疗服务、常态化疫情防控等诸多重任^[1-2]。自全民健康信息化建设推动以来,信息化手段在基层医疗卫生机构的应用提高了运行效率,方便了临床诊疗及公共卫生等工作的开展,管理工作也走向精准化^[3-4]。

基层医疗卫生主管部门运用报表收集和分析数据,实现精准管理^[5]。然而垂直系统间数据难以共享,基层报表重复填报问题明显^[6]。同时信息系统自动统计功能难以满足精准管理的个性化需求,需人工统计并上报^[7]。报表数量众多,来源、用途各异,统计加工复杂,在基层形成较大工作压力^[8-9]。要解决报表繁多、信息重复报送等非业务工作强度大的问题,亟待打通系统间数据壁垒,实现智能填报。

本研究立足基层实际,探索能普遍适用的填报优化路径。选取一地作为研究对象,在充分进行文献与实地调研的基础上对基层报表情况建模分析,并以该地具体系统与数据为例构建契合需要的数据上报模型,解决该地区基层医疗卫生机构自动化报表问题。

2 现状调研与需求分析

采取现场调研与专家咨询相结合的方式开展现状研究,一是邀请基层医疗卫生机构、属地社区卫生服务管理中心信息化工作人员以及属地卫生健康主管部门人员进行专家座谈,深入了解调研地区整体和当地情况、数据上报和信息互联互通需求,明确现场调研重点与方向;二是以驻场方式在基层医疗卫生机构开展现场调研,参与机构报表填报工作,梳理目前使用的信息系统,明确业务归属、系统级别、数据流向及各系统支撑的具体业务;最后从信息系统、纸质报表、台账档案中收集实际业务报表,整理报表条目,并脱敏处理涉及个人隐私和机构信息的部分。

2.1 基层医疗卫生机构信息系统现状分析

选取北京市东城区为调研地区。北京市服务于基层医疗信息业务的信息系统在北京市辖区内并不统一,其中除经济技术开发区外16个市辖区中11个区使用市级统建信息系统,5个区使用区自建信息系统。东城区为自建系统区,相比统建区而言核心业务系统增补与功能调整更为灵活。

经调研(截至2023年12月),北京市东城区基层医疗卫生机构业务工作依靠“东城区社区卫生精准化健康服务系统”及各细分业务区段的数十个信息系统完成诊疗服务、管理服务以及国家规范的12类公共卫生服务^[10]。若按一线工作人员实际交互操作感知情况将系统划分为主要系统(有较多人机信息交互)和外围支持系统(后台运行或极少人机信息交互),则可分为16个主要系统、33个外围系统。其中主要系统包括国家级系统2个、市级系统11个、区级系统3个,见表1;外围系统包括区域实验室信息管理系统、区域影像存储与传输系统、分诊叫号系统、药品前置审核系统等。

表1 北京市东城区基层医疗卫生机构主要信息系统

系统级别	系统名称	业务类别
国家	中国疾病预防控制中心信息系统	传染病防控
国家	“云上妇幼”远程医疗平台	妇幼保健
北京市	“易督导”患者管理系统	传染病防控
北京市	北京市精神卫生管理系统	精神卫生防治
北京市	北京市预防接种信息系统	预防接种
北京市	北京市新冠病毒疫苗统一管理平台	预防接种
北京市	北京市儿童预防接种管理系统	预防接种
北京市	北京市食品安全综合信息平台	食品卫生
北京市	北京市人口死亡登记报告信息系统	生命统计
北京市	北京市妇幼保健网络信息系统	妇幼保健
北京市	北京市医保系统	医疗保险
北京市	信用医疗	医疗保险
北京市	院前院内医疗急救信息衔接平台	卫生应急
东城区	东城区区域人口健康信息平台	机构运行管理
东城区	东城区社区卫生精准化健康服务系统	一般性业务工作
东城区	北京市东城区社区卫生服务系统	药品管理

其中“东城区社区卫生精准化健康服务系统”

为核心业务系统，覆盖医护工作站、健康档案、家庭医生签约等业务模块；预防接种业务的预约数据分散在“首都疫苗”应用、“声智健康”小程序等应用内，尚未与信息系统对接。还有极少部分业务工作未通过信息系统实现，利用本地表单等方式完成。

2.2 基层医疗卫生机构报表填报需求分析

在北京市东城区卫生健康委员会、东城区社区卫生服务管理中心等单位开展报表填报专题座谈。整体上看，东城区在垂直系统与业务系统互联互通方面存在短板，自动统计分析 with 考评方面尚有改进空间，预约挂号、健康档案、划价收费等部分业务内容互联互通受限。基于东城区宏观情况，在北京市东城区建国门社区卫生服务中心及其下辖站保健科等核心业务科室进行报表指标专题调研，共收集 167 种在用报表。报表类别与数量分布，见表 2。基层医疗卫生机构报表填报工作涉及十余个方面，主要集中在预防接种、妇幼保健、传染病、人口学信息收集方面，4 项占比达总量的 73.65%。

表 2 东城区基层医疗卫生机构报表填报情况

数据类别	数量 (种)	占比 (%)	累计占比 (%)
预防接种	37	22.16	22.16
妇幼保健	34	20.36	42.51
传染病	32	19.16	61.68
人口学	20	11.98	73.65
卫生监督协管	2	1.20	74.85
精神卫生防治	12	7.19	82.04
绩效考核	5	2.99	85.03
护士工作站	5	2.99	88.02
家庭医生签约	3	1.80	89.82
其他	17	10.18	100

从社区一线收集的报表填报情况与其使用信息系统情况对比可知，预防接种报表涉及 3 个北京市属垂直系统，妇幼保健与传染病报表各涉及 1 个市属和 1 个国家级垂直系统，其他各项除人口学报表涉及卫生健康主管部门外数据交换，均需从各垂直系统获取数据。因此，构建可用的报表自动化填报模型，实现分散在不同垂直系统数据的整合、统一存储与精准分发，是解决报表繁多和重复填报等基

层问题的关键。

3 模型构建路径

在对业务、系统、指标建模的基础上整合跨系统跨业务指标数据，完成从综合业务数据到具体报表数据的映射，实现智能上报。首先分析基层医疗卫生机构业务所属条线，明确各业务在机构间协同情况；其次分析基层工作中使用的信息系统，结合业务分析，构建信息共享交换模型；最后分解现场调研获得的基层医疗卫生机构实际业务报表，完成数据元、数据集以及数据的统计、加工和指标映射，实现从报表条目到指标库的整合转化。业务协同模型将用于刻画基层医疗卫生机构与各部门的业务交互关系；共享交换模型将作为数据整合与分发的规则；指标库将作为数据整合的内容。

3.1 基层医疗卫生机构业务协同模型构建

参考相关国家标准^[11-12]，结合实地调研情况，对基层医疗卫生机构与其他业务部门的协同关系进行分析与建模，明晰协同业务间信息需求。业务协同模型网络，见图 1，其中以字母标识的业务协同具体类目，见表 3。基层医疗卫生机构与各级疾病预防控制中心、二级以上医院间的单一业务协同类目较多；基层医疗卫生机构与急救中心、血液中心、卫生健康管理部门间的重复业务协同类目较多。

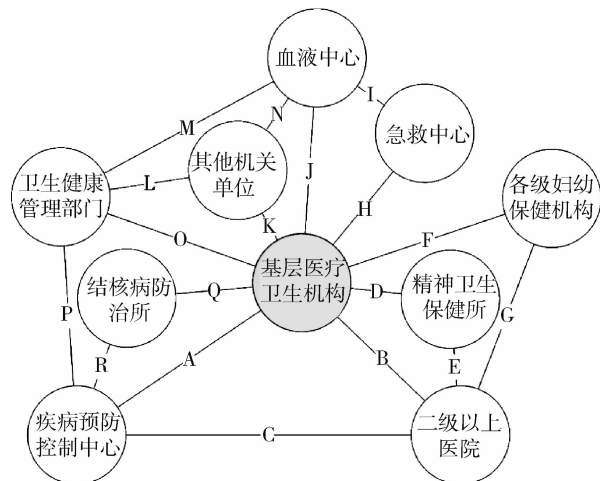


图 1 基层医疗卫生机构业务协同模型网络

表3 业务协同模型网络协同类目

编号	内容
A	健康教育、营养健康服务管理、环境卫生管理、免疫规划、食品安全风险监测、卫生监督协管、传染病防控、寄生虫病防控、慢性病防控、地方病防控、老年人健康服务管理、伤害防控、卫生应急
B	营养健康服务管理、环境卫生管理、传染病防控、寄生虫病防控、慢性病防控、地方病防控、精神卫生防治、癫痫防治、老年人健康服务管理、妇幼健康服务管理、伤害防控、康复服务、食品安全风险监测、医疗协同
C	慢性病防控、地方病防控、伤害防控、营养健康服务管理、环境卫生管理、食品安全风险监测、传染病防控
D、E	精神卫生防治
F	妇幼健康服务管理、计划生育
G	妇幼健康服务管理
H、I	卫生应急
J、K、L、P	卫生应急、机构运行管理
M、N	机构运行管理
O	绩效管理、机构运行管理、妇幼健康服务管理、计划生育、卫生应急
Q、R	传染病防控

3.2 基层医疗卫生信息共享交换模型构建

3.2.1 现况模型构建 以业务协同模型为基础，广泛调研北京市东城区信息系统功能应用情况，参考相关国家标准^[13]，以基本业务包含的信息流动方

向为核心构建信息共享交换模型。根据调研结果和国家标准，对东城区信息共享交换现况建模分析，模型示例，见图2。不同业务系统中的数据在疾控中心、医院、妇幼保健机构去向方面有较多汇集。

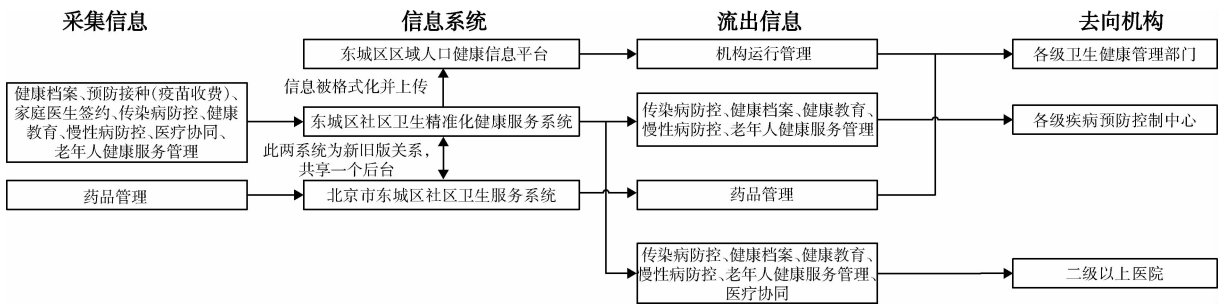


图2 北京市东城区基层医疗卫生信息共享交换模型（核心业务系统）

3.2.2 需求模型构建 在现况建模基础上，充分了解一线工作人员的信息互联互通需求，将数据融

合与数据共享需求建模分析，模型示例，见图3。

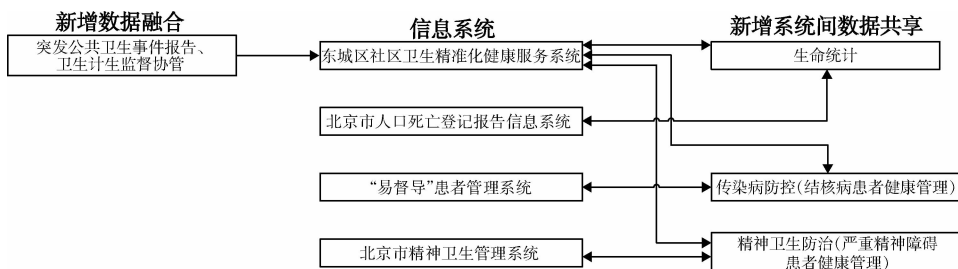


图3 待新增共享交换模型（统计、结核病防控、精神卫生防治）

可以发现，东城区主要业务系统（“东城区社区卫生精准化健康服务系统”）功能集成较为完备，多数日常诊疗业务功能均可在该平台实现并存储数据；健康档案、家医签约作为系统核心业务模块，与基层医疗机构内部各系统间，以及协同部门间横向纵向的妇幼保健信息等大量业务数据存在调阅和共享；传染病防控、预防接种、妇幼保健、医疗保险等垂直系统数据尚未完成与主要业务系统的数据

整合。

3.2.3 数据整合模式构建 结合业务协同模型所展示的部门间业务划分与交互情况，自动化填报应当依托“东城区社区卫生精准化健康服务系统”，以其健康档案信息为核心，结合指标库从各类信息系统中抽取数据，并完成数据的映射与整合。按业务类型细化的信息整合模式（以妇幼健康为例），见图 4。

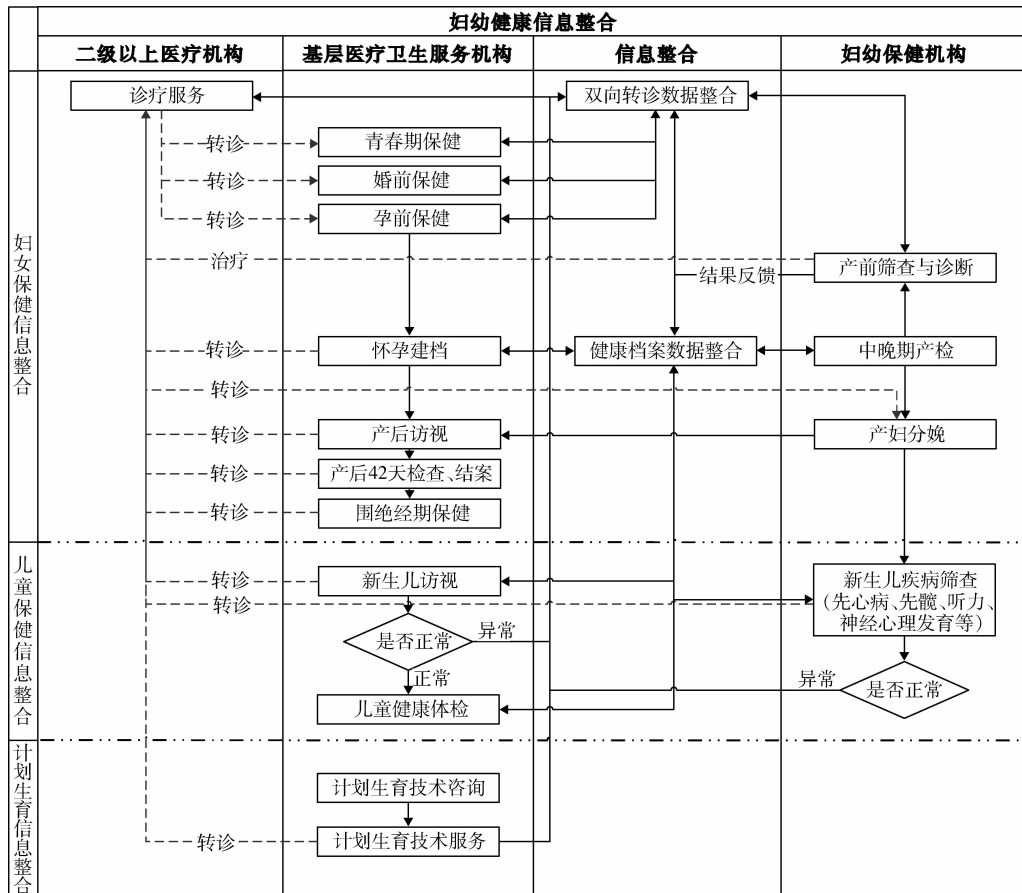


图 4 信息整合模式（妇幼健康）

3.3 基层医疗卫生机构报表指标库构建

梳理现场调研获得的 167 种报表，共分解合并出 3 476 项指标，按业务类别对指标数量归类统计，见表 4。将分离的指标统合并整理为指标库，其中每项指标按照具体需求，包含指标的类别、来源、去向、格式、被包含的表单、解释、人群限制及统计限制等补充描述，见表 5。

表 4 东城区基层医疗卫生机构报表指标情况

数据类别	数量（项）	占比（%）	累计占比（%）
预防接种	595	17.12	17.12
妇幼保健	1 091	31.39	48.50
传染病	635	18.27	66.77
人口学	310	8.92	75.69
卫生监督协管	69	1.99	77.68
精神卫生防治	233	6.70	84.38
绩效考核	143	4.11	88.49
护士工作站	84	2.42	90.91
家医签约	102	2.93	93.84
其他	214	6.16	100

表5 北京市东城区基层医疗卫生机构报表指标库(示例)

指标类别	指标来源	指标名	指标解释	统计限制	指标格式	指标所对表单	上报去向	上报周期
妇幼保健	市妇幼	首筛-应筛高危儿人数	首次神经心理发育筛查符合筛查标准的儿童总数	年龄 ≥ 6个月; 高危因素=1	整数	13. 年高危儿筛查情况统计(神经心理发育)	妇幼保健院	年报
妇幼保健	市妇幼	首筛-实筛高危儿人数	首次神经心理发育筛查实际筛查的儿童总数	年龄 ≥ 6个月; 高危因素=1	整数	13. 年高危儿筛查情况统计(神经心理发育)	妇幼保健院	年报
妇幼保健	市妇幼	首筛-筛查率	(实筛人数/应筛人数) × 100%	—	百分数	13. 年高危儿筛查情况统计(神经心理发育)	妇幼保健院	年报
.....								
传染病	“易督导”	正在管理肺结核患者人数	社区中心正在随访的总人数	—	整数	43. 肺结核患者社区督导管理半年报表	结核病防治所	半年报
传染病	“易督导”	半年内社区接收管理通知单人数	1月1日—6月30日或7月1日—12月31日, 接到通知单的总人数	—	整数	43. 肺结核患者社区督导管理半年报表	结核病防治所	半年报
传染病	纸质台账	半年内实际开展督导管理肺结核患者人数	1月1日—6月30日或7月1日—12月31日, 接到通知单且有面访记录的总人数	—	整数	43. 肺结核患者社区督导管理半年报表	结核病防治所	半年报

3.4 基于数据整合的基层医疗卫生多源数据自动上报模型构建

数据整合是通过数据集成平台统一处理多个来源的数据, 去除数据间差异, 实现统一表示、存储和管理的方法^[14-16]。医疗卫生数据由于各平台建设时间差异、承建方差异等因素, 很难遵循统一的标准规范, 导致数据之间存在语义异构问题。以东城区为例, 其核心业务信息系统“东城区社区卫生精准化健康服务系统”包含挂号、诊疗、健康档案、家医签约等主要基础数据, 另有妇幼保健、精神卫生防治、预防接种等数据分散在其他信息系统, 此类数据的利用需数据整合。本研究构建的自动上报模型包含数据整合、数据存储、数据分发3个环节。模型先对不同入口、不同格式的数据整合、存储, 再将整合的数据与报表指标建立映射关系, 完成信息分发。

3.4.1 数据整合 数据整合环节解决跨系统多格式数据标准化问题。首先从机构运行的实际业务数据、在用的信息系统数据库和相关国家标准^[13, 17]中构建动态更新的元数据集; 其次在各数据集表结构中筛选出可用于数据集对准的共有信息, 通过数据接口提取报表基础数据, 并根据基础数据中的身份信息与核心业务系统存储的健康档案信息对齐, 提取档案信息; 最后基于指标库完成对提取数据的统计加工。针对半结构化和非结构化数据, 可运用资源描述框架模型整合。如运用“D2RQ”工具完成结构化数据整合、运用扩展样式表语言转换技术完成半结构化数据整合^[18]。利用自动抽取方法构建本体, 运用基于“WordNet”的双向自动本体对准方法^[19-21]实现本体实时更新, 基于本体进行自然语言处理, 实现非结构化数据结构化, 从而获取用资源描述框架模型统一格式的元数据集。

3.4.2 数据存储与分发 将整合后的数据按核心

业务系统数据库结构存储，数据分发环节将存储的整合数据按指标库标定的映射规则向具体表单映射，并按报表上报要求分发。根据实地调研情况，报表上报分 3 种途径：一是通过电子邮件、通信软件、纸质文档等形式将表单传递至对应单位；二是通过开放信息接口的信息系统上报；三是通过未开放信息接口的信息系统上报。针对第 1 种途径，信

息分发模块将直接导出表单文档，供信息上报使用；针对第 2 种途径，信息分发模块将按对应信息系统的数据库格式要求输出数据至相应系统；针对第 3 种途径，信息分发模块可使用界面自动化技术^[22]输出数据。东城区报表所涉数据绝大多数来源于结构化数据库，以预防接种业务为例，结构化数据跨信息系统整合与上报模型，见图 5。

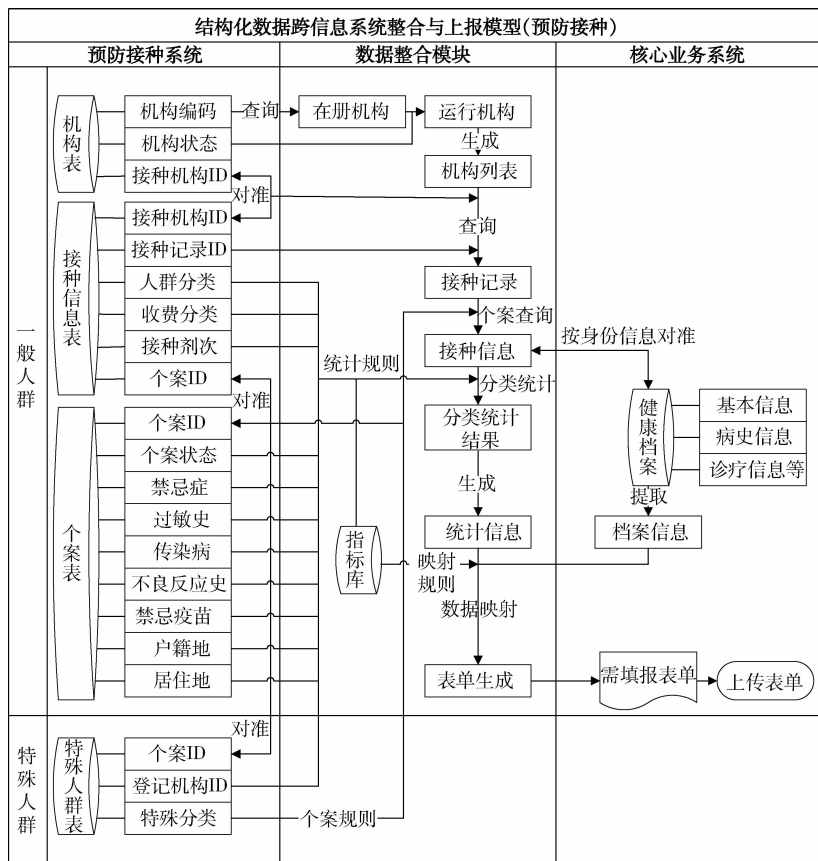


图 5 结构化数据跨信息系统数据整合与上报模型（预防接种）

数据整合模块从垂直系统中获取报表所需的数据集，从业务系统中获取健康档案，并在表间建立主键关联。指标库提供统计规则统计分析，提供映射规则完成表单生成，获得目标表单。其中，“特殊人群”为需专项统计的人群，如外来务工人员、学生等，每个人群对应单独的数据集记录及其专项统计相关信息。无需专项统计的人群为“一般人群”。“特殊人群表”中“特殊分类”信息根据表对应人群不同而不同，如对应外来务工人员时，该信息为“单位类型”，包含参照国民经济行业分类（GB/T 4754—2017）并按本地情况划分粒度制定的用工单位类型。

4 结语

本研究尝试探索能普遍适用的报表优化路径。第 1 步：选取一地进行实际调查研究，具体梳理该地区基层医疗卫生机构业务开展情况、报表情况、信息系统应用情况，收集一线业务人员信息整合需求，构建指标库。第 2 步：根据调研实际，查阅相关资料、联系相关部门，明确以基层医疗卫生机构为核心、涉及信息交互的机构间业务协同关系，构建基层医疗卫生机构业务协同模型。第 3 步：根据

调研获得的系统设置与信息交互情况, 结合一线需要, 构建基层医疗卫生机构信息共享交换模型。第 4 步: 综合调研结果、建模分析, 构建多源数据自动上报模型, 解决不同业务、不同系统间语义异构问题, 实现数据的自动整合、存储与分发。

本研究提供的模型构建路径具有较好的可推广性, 能因地制宜构建报表自动化模型, 但仍存在一些不足。一是模型仅能实现结构化数据库的信息整合, 仍存在少量非结构化数据无法整合。其他地区结构化数据来源可能较少, 难以满足自动化报表需求。二是数据存储和数据整合模块内, 数据承载力有限, 长期运行稳定性受到限制。要解决以上问题, 可在本研究基础上, 探索运用自然语言处理等方法处理非结构化数据、尝试通过个人识别码和数据中间件实现统一存储, 在国家标准^[10]规定的健康档案内容基础上扩展经数据融合的数据集, 以增量获取方式更新并存储, 形成元数据库^[23-24]。本研究设计的自动化模型可成为开发基层医疗卫生管理自动填报系统的核心, 还可以支撑区域医疗卫生资源中心的建设, 为基层医疗卫生业务减负与智能化改造提供基础。

利益声明: 所有作者均声明不存在利益冲突。

参考文献

- 1 中华人民共和国基本医疗卫生与健康促进法 [EB/OL]. [2023 - 12 - 10]. http://www.npc.gov.cn/npc/c2/c30834/201912/t20191231_304414.html.
- 2 谭秋成. 基层医疗卫生机构的性质、行为及发展问题 [J]. 学术界, 2021 (8): 195 - 209.
- 3 郝晓宁, 马骋宇, 刘志业, 等. 中国基层卫生信息化改革的成效及问题研究 [J]. 卫生经济研究, 2020, 37 (7): 3 - 5, 9.
- 4 单既桢, 郑攀, 琚文胜. 基层医疗与公共卫生管理服务信息系统设计与实现 [J]. 中国卫生信息管理杂志, 2021, 18 (3): 346 - 50.
- 5 胡红濮, 秦盼盼, 雷行云, 等. 我国全民健康信息化发展历程及展望 [J]. 医学信息学杂志, 2019, 40 (7): 2 - 6.
- 6 邢春国, 夏迎秋, 吴丹云, 等. 江苏省基层医疗卫生机构信息化建设与使用情况研究 [J]. 中国全科医学, 2022, 25 (16): 2008 - 2013.
- 7 张岩岩, 刘英华, 梁士平. 基层公共卫生统计数据质量改进研究 [J]. 统计与管理, 2022, 37 (7): 122 - 128.
- 8 夏英华, 洪紫慧, 曹蓉, 等. 社区卫生服务中心信息化建设实践及政策建议 [J]. 中国全科医学, 2019, 22 (25): 3068 - 3075.
- 9 彭博, 李亚子, 郑见立, 等. 基层医疗卫生机构卫生健康类报表填报现状调查 [J]. 中国初级卫生保健, 2021, 35 (8): 9 - 12.
- 10 国家基本公共卫生服务规范 (第三版) [EB/OL]. [2023 - 12 - 10]. <http://www.nhc.gov.cn/jws/s3578/201703/d20c37e23e1f4c7db7b8e25f34473e1b.shtml>.
- 11 全国基层医疗卫生机构信息化建设标准与规范 (试行) [EB/OL]. [2023 - 12 - 10]. https://www.gov.cn/zhengce/zhengceku/2019-09/16/content_5430265.htm?eqid=b894bf92000e34040000006645f87c1.
- 12 全国公共卫生信息化建设标准与规范 (试行) [EB/OL]. [2023 - 12 - 10]. https://www.gov.cn/zhengce/zhengceku/2020-12/12/content_5569035.htm.
- 13 基层医疗卫生信息系统基本功能规范 [EB/OL]. [2024 - 03 - 01]. <http://www.nhc.gov.cn/wjw/s9497/202305/b5bf8f1e707f40da9066651359f17ca2.shtml>.
- 14 李亢, 李新明, 刘东. 多源异构装备数据集集成研究综述 [J]. 中国电子科学研究院学报, 2015, 10 (2): 7.
- 15 毛俊华. 基于模糊聚类的智慧医院多源异构数据整合系统 [J]. 电子设计工程, 2022, 30 (7): 120 - 124.
- 16 柳原. 多源异构数据整合系统在医疗大数据中的研究 [J]. 电子制作, 2019 (14): 64 - 65.
- 17 卫生健康信息数据元标准化规则 [EB/OL]. [2023 - 12 - 10]. <http://wsbz.nhc.gov.cn/wsbzw/upload/news/05299749dd284db498e52efe55f589e.pdf>.
- 18 刘俊, 李亚辉, 王立林, 等. 基于本体的战场多源信息整合技术研究 [C]. 北京: 第五届中国指挥控制大会, 2017.
- 19 DJEDDI W E, KHADIR T. A dynamic multistrategy ontology alignment framework based on semantic relationship using WordNet [EB/OL]. [2023 - 12 - 10]. <https://api.semanticscholar.org/CorpusID:7006468>.
- 20 潘有能, 刘朝霞. 基于 WordNet 的关联数据本体映射研究 [J]. 情报杂志, 2013, 32 (2): 4.
- 21 赵天忠, 苗壮, 张亚非, 等. 基于 WordNet 重用的领域本体构建方法 [J]. 系统仿真学报, 2007, 19 (19): 5.
- 22 张博刚, 张威, 陈月宁, 等. 基于运行监测的图形用户界面自动化测试模型 [J]. 计算机应用, 2010, 30 (10): 2749 - 2753.
- 23 INGER JOHANNE B, ARIANSENAMS, GUN PEGGYK, et al. The Norwegian patient registry and the Norwegian registry for primary health care: research potential of two nationwide health - care registries [J]. Scandinavian journal of public health, 2020, 48 (1): 49 - 55.
- 24 曾汪旺, 谢颖夫, 胡光阔. 医院多源异构医疗数据整合方法研究 [J]. 中国卫生信息管理杂志, 2017, 14 (2): 197 - 200, 204.