

气象敏感性疾病综合数据平台设计与实现*

郑环 陈强 张睿 孟郁洁 赵嘉 金丽珠 杜雪杰 王松旺

(中国疾病预防控制中心 北京 102206)

[摘要] **目的/意义** 设计综合数据平台, 以满足气象敏感性疾病调查数据收集和共享的需要, 提升我国对气象敏感性疾病的监测能力。**方法/过程** 通过数据抽取、数据交换、数据导入、定制采集等多种方式, 将疾病数据、气象数据、环境数据、诊疗数据等加工处理, 纳入平台统一管理。**结果/结论** 该平台实现了数据采集、汇交、可视化展现、数据共享功能, 可为全国各基地的科研人员更好地管理和利用气象敏感性疾病调查数据提供支撑。

[关键词] 气象敏感性疾病; 数据定制; 数据质量评估; 数据可视化

[中图分类号] R-058 **[文献标识码]** A **[DOI]** 10.3969/j.issn.1673-6036.2024.06.014

Design and Implementation of a Comprehensive Data Platform for Meteorological Sensitive Diseases

ZHENG Huan, CHEN Qiang, ZHANG Rui, MENG Yujie, ZHAO Jia, JIN Lizhu, DU Xuejie, WANG Songwang

Chinese Center for Disease Control and Prevention, Beijing 102206, China

[Abstract] **Purpose/Significance** To design a comprehensive data platform to meet the needs of collecting and sharing survey data on meteorological sensitive diseases, so as to enhance monitoring capabilities for meteorological sensitive diseases in China. **Method/Process** Through various methods such as data extraction, data exchange, data import and customized collection, disease data, meteorological data, environmental data and diagnosis and treatment data are processed and integrated into the platform for unified management. **Result/Conclusion** This platform realizes the functions of data collection, aggregation, visualization display and data sharing, which can provide support for scientific researchers in various bases across the country to better manage and utilize meteorological sensitive disease survey data.

[Keywords] meteorological sensitive diseases; data customization; data quality assessment; data visualization

1 引言

随着全球气候变化、城镇化和人类对环境资源

的开发和利用, 气象条件变化、极端天气事件、气象灾害等气象因素给传染性疾病、慢性非传染性疾病的防控带来极大挑战。中国是全球气候变化的敏感区^[1], 受气象变化的影响较大, 同时, 人口老龄化加剧, 气象敏感性疾病(即与气象条件密切相关的疾病, 包括心脑血管疾病、呼吸系统疾病、传染病、精神疾病等)呈现加重之势^[2]。虽然近年来针对气象敏感性疾病的研究陆续开展^[3-8], 但覆盖不同气象地理区划, 不同类型疾病的科学调查研究较少。考虑到敏感人群在不同地区所受影响存在差

[修回日期] 2024-03-14

[作者简介] 郑环, 助理研究员, 发表论文 9 篇; 通信作者: 王松旺, 研究员。

[基金项目] 科学技术部科技基础资源调查专项(项目编号: 2017FY101207)。

异, 针对目标人群应制定个性化的防护措施^[9]。因此, 开展跨领域跨区域人群气象敏感性疾病的科学调查, 产出权威系统的气象敏感性疾病科学数据、调查报告、科技资料并实现开放共享对于维护公共卫生安全和推动国家战略发展具有至关重要的意义。

为作好各类数据的统一集中管理和共享, 更高效地开展气象敏感性疾病调查的数据收集、数据质量管理、数据分析和共享, 迫切需要建立一套覆盖气象、环境和疾病领域调查、汇交、分析、共享的多功能业务系统。通过设计开发气象敏感性疾病综合数据平台, 实现多领域数据采集、数据汇交、数据可视化、数据共享和辅助决策等功能, 可减少重复采集、促进各数据提供方沟通合作, 从而提高我国气象敏感性疾病的预测预警水平, 为相关疾病的预防控制政策制定提供数据支撑^[10]。

2 需求分析

2.1 服务对象

气象敏感性疾病综合数据平台面向政府、科研和公众 3 大群体提供服务。面向政府人员需支持调查数据的报表查询、可视化展示、数据管理与共享; 加强数据监测监管, 提供决策支持。面向科研人员需满足科研人员以及各调查基地对数据汇交、数据标准管理、数据上报、数据共享和数据分析等方面的需要。面向公众需满足人群干预服务需要和调查基地地区气候预测预警服务需要。平台的用户类型设计有系统管理员、数据管理员、数据质控人员、项目人员和普通注册用户、公众(未注册用户)。不同用户的职责和权限不同, 见表 1。

表 1 用户类型分析

用户类型	用户	需求	系统使用
系统管理员	项目管理人员	整个项目组的管理, 包括人员、成果、时间进度等	实现对系统用户的管理、用户角色的分配, 以及对用户的数据资源权限分配、维护基础数据等
数据管理员	项目和课题组相关负责人	项目或课题组的人员、进度、成果管理; 负责权限分配审批和数据申请审批	实现对气象敏感疾病数据管理、展示、统计、基于主题的多维数据分析、监测预警预测、数据申请审核等操作、门户网站内容维护等
数据质控人员	项目和课题组相关人员	审核基地上报的数据以确保数据质量	实现对基地上报的本课题数据的审核, 审核通过后的数据作为正式数据
项目人员	基地工作人员, 各级疾控中心人员	管理与自己课题相关的数据	实现数据的查询、浏览、下载; 实现数据上报功能
公众用户(注册/非注册)	公众	浏览门户网站已公布的数据资源及申请数据资源	实现对公开发布的数据资源的浏览

2.2 业务流程

平台通过数据抽取、数据交换、数据导入、定制采集等多种方式, 集成疾病数据、气象数据、环境数据、诊疗数据等。获得数据后存入数据交换中间库, 根据已定义的标准数据结构进行数据交换,

根据定义的数据质量校核规则进行数据校核, 对不合规的数据标记和反馈, 校验合规的数据存储到业务数据库中。同时, 系统界面提供数据文件导入功能, 导入文件格式支持 Excel、XML、JSON 等, 导入的数据同样需要校验, 并标记校验不通过的数据。总体流程, 见图 1。

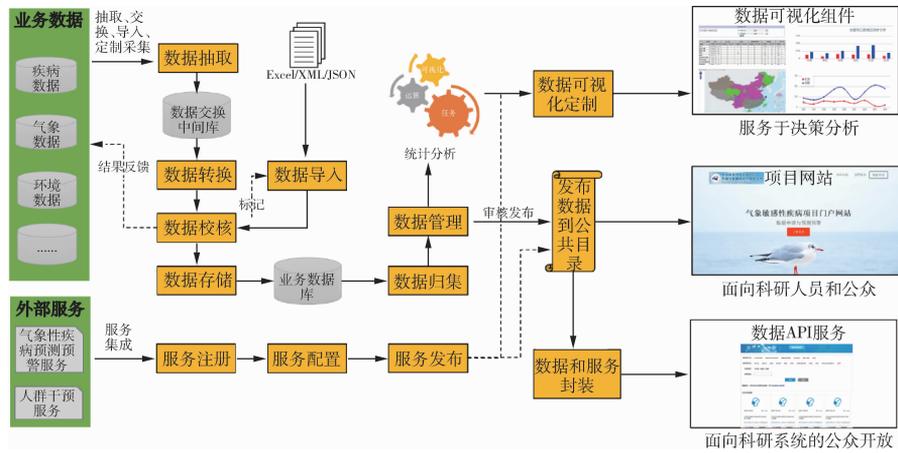


图 1 总体流程

平台管理的数据有 3 个流向：一是将数据发布到公共目录上，供有权限用户查看使用，或者面向科研人员和公众在项目门户网站上开放数据共享；二是将数据进行统计分析运算，结合平台已集成的预测预警与干预服务，产出供数据统计分析和数据可视化使用的各类统计分析报表，用户通过自定义数据可视化组件的形式，使用平台已有数据；三是将数据封装成一个个数据服务接口，通过平台统一对接口封装，演变为平台可对外共享和发布的服

务。通过服务集成，将气象性疾病预测预警服务、人群干预服务等外部服务集成注册至平台，经过服务配置和服务发布，转变为可供平台再次利用的服务。这类服务可与平台已有数据整合，为统计分析和数据可视化服务，也可以经过接口的二次封装，演变为平台可对外共享和发布的服务。

3 系统设计与实现

3.1 架构设计 (图 2)

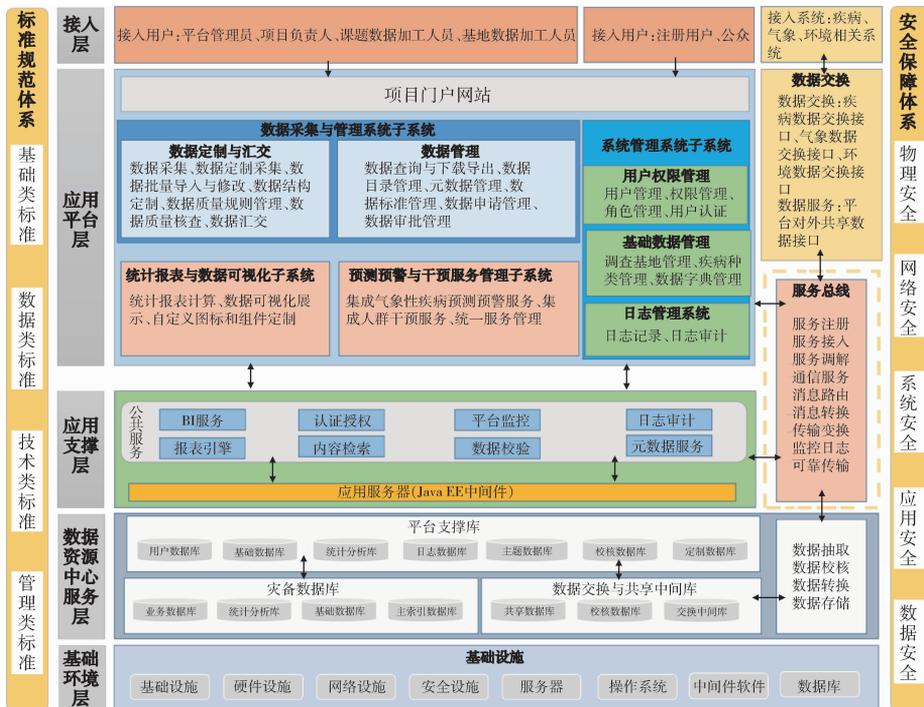


图 2 总体架构

从横向上看项目由 3 部分组成：标准规范体系、业务应用体系、安全保障体系。从纵向上看，业务应用体系分层次组织，包括 5 层：基础环境层、数据资源中心服务层、应用支撑层、应用平台层、接入层。

3.2 数据库设计

为保证数据的完整性、一致性和跨平台使用，业务数据库采用关系型数据库 MySQL，部署为集群数据库模式，支持数据读写分离、数据分区、并行查询，引入 Redis 缓存^[11]、FastDFS 文件系统等技术，提高系统性能和可伸缩性。

采用清晰的 3 层架构：表现层、业务逻辑层和持久层。每层专注于特定功能，实现高度的职责分离，保证数据模型的扩展性和适应性，当添加新功能或改进某个功能时，只需要在相应的部分进行修改，不影响其他部分。系统的业务处理功能和数据统计分析功能应设计为相对独立的系统，分别使用独立的云服务器、各自的数据库/数据存储系统，避免互相干扰。业务处理系统产生的数据，根据具体要求不同，通过分布式消息队列，以及抽取、转换和加载（extraction - transformation - loading, ETL）工具同步或批量导入数据统计分析系统。

平台包含数据采集与管理、统计报表和数据可视化、预测预警与干预服务管理、系统管理和气象敏感性疾病网站共 5 个子系统。各子系统主要功能，见图 3。

3.3 功能设计与实现

平台包含数据采集与管理、统计报表和数据可视化、预测预警与干预服务管理、系统管理和气象敏感性疾病网站共 5 个子系统。各子系统主要功能，见图 3。

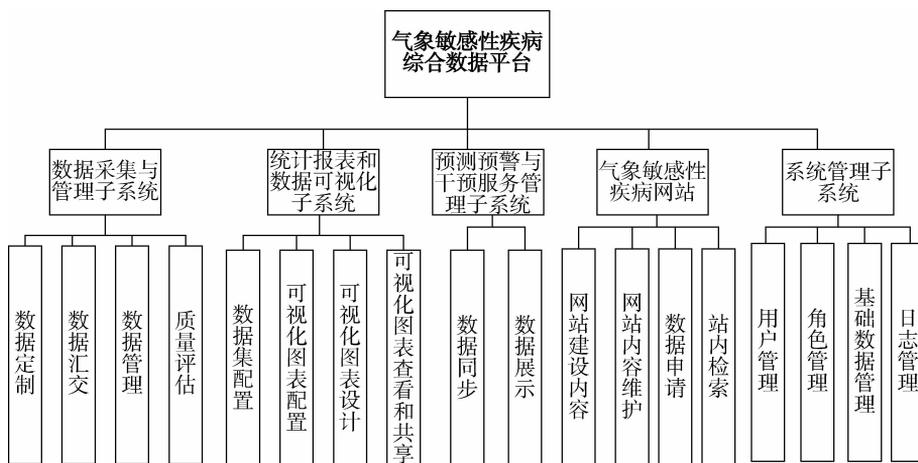


图 3 系统功能结构

3.3.1 数据采集与管理子系统 是平台的核心子系统，通过定制数据表单结构、数据值域字典和数据质量规则，将调查数据上传至平台，并对导入的数据进行质量评估。该子系统由数据定制、数据汇交、数据管理、质量评估 4 个模块组成。其中数据定制模块包括定制管理、校验规则管理、数据表标准定制功能，可批量将调查数据导入数据综合平台，为数据汇交和数据管理提供基础。数据汇交模块提供面向调查基地用户的操作界面，支持数据文件上传和批量导入，或者通过集成流行病学动态数据采集平台传输数据^[12]。数据管理模块可统一管理平台上的数据资源，实现数据目录管理、元数据管

理、数据审批管理，如各项增、删、改、查，提供数据的分类管理功能，还可对门户网站的数据申请进行受理和审批。数据质量评估模块主要根据定义的数据质量校验规则统计上报数据的质量，通过数据质量评估指标权重进行计算评分，并生成评估报告。通过该子系统可实现采集数据的动态定制、批量上传、质量评估和统一管理，提高数据管理效率。

3.3.2 统计报表与数据可视化子系统 主要用于可视化页面的定制，对数据采集与管理子系统采集的数据进行统计分析和可视化展示，包括数据集配置、可视化图表配置、可视化图表设计、可视化图

表查看和共享4个模块。其中数据集配置模块包括数据源管理和数据集管理。可视化图表配置模块通过数据库查询语句获取字段数据,根据获得的字段数据配置维度列、指标列、可选表达式、可选过滤等信息。可视化图表设计模块对生成的各类图表进行管理,支持表格、饼图、直方图、线图、条形图和专题地图等图表类型。可视化图表查看和共享模块提供可视化看板的设计、管理和共享功能。通过该子系统可实现统计报表和可视化图表的动态定制,使数据资源更加直观地展现和比较,便于用户理解和分析数据。

3.3.3 预测预警与干预服务管理子系统 实现预测预警相关数据的动态获取、同步展示,主要形式包括健康宣传材料、宣传视频、极端天气公众提醒等。各类健康宣传材料可供在线查询与浏览,并具备增、删、改功能。该子系统可实现疾病预测预警服务和人群干预服务的统一管理和维护,支撑平台向公众和干预目标人群提供相关服务。

3.3.4 气象敏感性疾病网站 是发布和分享项目相关信息和数据的重要渠道,分前台门户网站和门户网站后台管理系统两大部分,其中前台实现用户注册登录、数据资源信息公开、疾病预测预警服务及在线数据申请、互动交流;后台实现各类内容信息发布管理、文件管理以及数据申请审批管理等功能。门户网站栏目有平台简介、热点追踪、工作动态、基地简介、数据申请、元数据、数据下载、科普资料、预测预警、人群干预、成果展示等。通过门户网站,可开展项目信息公开、调查数据共享、重点人群干预和项目成果推广等工作。

3.3.5 系统管理子系统 包括用户权限管理、基础数据管理和日志管理等。用户权限管理是对综合数据平台各类用户进行统一管理、权限分配,保障数据访问控制安全。基础数据管理是整个平台的支撑系统,是保障业务系统正常运行的基础,主要功能包括:调查基地管理、疾病种类管理、数据字典管理等。日志管理分为用户操作日志及系统运行日志管理,通过日志的采集进行各类日志审计工作。

3.4 安全保障

平台存储大量的疾病监测数据,安全性要求较

高。为保证系统运行和数据内容的安全性,考虑如下措施:一是系统用户身份认证采用双因子认证;二是敏感数据采用国密SM2算法加密存储;三是采用https协议进行数据传输;四是对不同用户控制不同的访问权限;五是对用户的所有操作记录进行审计日志存储,通过审计日志找出非法存取数据的用户、时间和内容。

4 应用效果

气象敏感性疾病综合数据平台于2019年上线并应用于我国科技基础资源调查专项“我国区域人群气象敏感性疾病科学调查”项目,对44个调查基地的4488万余条气象、环境和气象敏感性疾病的监测数据进行统一采集和管理,在数据采集方面通过系统导入功能提高了数据输入的效率,节省了用户的时间和精力;汇交过程中实现了数据质量控制,通过多重校验,保证数据的一致性和准确性;针对调查区域和疾病建立专题可视化分析,有助于课题人员更直观地理解数据,发现数据间的关联关系;为公众提供气象敏感性疾病科普宣传、人群干预服务和相关疾病的预测预警服务,在维护公众健康、降低疾病发生率、提高应对能力等方面发挥了重要作用。通过该平台的部署应用,能够满足业务人员对多源数据的汇交、质量评估、可视化展示和综合分析等功能需求,为制定更为精准有效的疾病预防控制政策提供坚实的数据基础。同时,该系统具有较高的稳定性和可靠性,可推广应用于更多地区。

5 结语

气象敏感性疾病综合数据平台集数据定制、采集、汇交、可视化、探索性分析、数据共享和辅助决策功能于一体,不仅能满足当前数据管理需要,而且能满足以后扩大监测范围和监测内容的需求,使我国气象敏感性疾病的监测能力得到显著提升,为制定精准有效的疾病预防控制策略提供坚实的数据基础。但在实际应用过程也有一定的局限性,如用户使用手机访问时界面友好性不强,在数据深加

工、个性化分析和创新服务方面仍有较大提升空间。另一方面,气象敏感性疾病错综复杂,要实现全面的气象敏感性疾病的预警和风险评估需深入探究气候变暖对疾病发病规律及其病理的影响机理,构建气象部门与卫生、环保等多部门间的协同合作体系,共同推进相关综合研究。

作者贡献: 郑环负责论文初稿撰写和修改;陈强负责论文润色;张睿、孟郁洁负责文献调研;赵嘉、金丽珠负责文献整理;杜雪杰负责技术支持;王松旺负责论文修改与审核。

利益声明: 所有作者均声明不存在利益冲突。

参考文献

- 1 胡利娟.《中国气候变化蓝皮书(2022)》发布中国升温速率高于全球平均水平[J].中国科技财富,2022(8):48.
- 2 王嘉鑫,石彦军,卢山,等.我国华东与西南县域主要气象敏感性疾病变化特征及其医疗费用研究[J].沙漠与绿洲气象,2019,13(6):133-140.
- 3 肖冰霜,马玉霞,郑晓东,等.南京市最低气温对不同人群循环系统疾病死亡人数的影响[J].兰州大学学报(自然科学版),2017,53(3):401-406.
- 4 董继元,陈永聪,张本忠,等.兰州市气温对脑卒中发

病的滞后效应研究[J].气候变化研究进展,2017,13(4):366-374.

- 5 张莹,王式功,贾旭伟,等.气温与PM2.5协同作用对疾病急诊就诊人数的影响[J].中国环境科学,2017,37(8):3175-3182.
- 6 杨丝絮,彭丽,叶晓芳,等.气温与PM2.5协同作用对心脑血管疾病死亡的影响研究[J].上海预防医学,2023,35(7):660-666.
- 7 ZHANG R, MENG Y, SONG H, et al. The modification effect of temperature on the relationship between air pollutants and daily incidence of influenza in Ningbo, China [J]. Respiratory research, 2021, 22(1): 153.
- 8 李传玺,刘起勇,马伟.广州市极端降水事件对不同特征人群登革热发病的影响[J].山东大学学报:医学版,2021,59(12):151-157.
- 9 马盼,王式功,尚可政,等.气象舒适条件对呼吸系统疾病的影响[J].中国环境科学,2018,38(1):374-382.
- 10 徐飞龙,吕佳,吴佳荣,等.中医药科学数据管理平台设计与实现[J].医学信息学杂志,2023,44(12):78-82.
- 11 张俊.基于Redis实现关系型数据库内存化研究[D].成都:四川师范大学,2021.
- 12 孟郁洁,戚晓鹏,马睿,等.流行病学动态数据采集平台在公共卫生调查中的应用[J].医学信息学杂志,2013,34(6):18-22.

(上接第78页)

参考文献

- 1 徐婷,鲍勇,王韬.中国公共卫生应急管理体系的变迁与效果分析[J].中国公共卫生,2020,36(12):1704-1706.
- 2 于佳会,刘佳静,郑建明.多源多维数据融合研究态势:理论、方法与应用[J].情报杂志,2022,41(5):133-138,207.
- 3 邓建高,吴灵铭,齐佳音,等.突发公共卫生网络舆情信息传播博弈分析[J].现代情报,2021,41(5):139-148.
- 4 刘樑,金瑞丰,魏华.突发公共卫生事件研究的趋势和特征——基于CiteSpace的知识图谱分析[J].电子科技大学学报(社会科学版),2021,23(4):46-52.
- 5 STEINBERG A N, BOWMAN C L, WHITE F E. Revisions to the JDL data fusion model [C]. Orlando: Sensor Fusion: Ar-

chitectures, Algorithms, and Applications III, 1999.

- 6 苏雪梅,赵自雄,赵嘉.疾病预防控制信息化建设、应用与规划[J].医学信息学杂志,2023,44(5):1-6.
- 7 宋磊.EMPI在医院信息系统中的研究与构建[J].医学信息学杂志,2016,37(4):63-65,82.
- 8 郭青,赵自雄,苏雪梅,等.基于电子疾病档案的疾病动态监测全周期管理模型研究[J].中国卫生信息管理,2020,17(4):411-415.
- 9 国家卫生健康委办公厅关于进一步加强疫情防控期间医务人员防护工作的通知[EB/OL].[2023-09-26].<http://www.nhc.gov.cn/zylg/s7659/202002/75c6e88ecbeb42a9a26acb538383e2fc.shtml>.
- 10 WANG F, AN Z, RODEWALD L, et al. Guangdong's study of the effectiveness of China's inactivated vaccines against the SARS-CoV-2 B.1.617.2 (Delta) variant [J]. China CDC weekly, 2021, 3(34):728-730.