

# 天津市药品供应保障大数据监测分析预警系统架构设计与实施

王博飞, 顾嘉, 柳涛

(天津市卫生健康委员会综合服务中心, 天津 300074)

**摘要:**近年来, 药品供应保障问题受到多方关注, 药品监测预警成为了热点。建设药品供应保障综合管理平台和短缺药品监测预警信息系统, 综合分析多维度药品供应保障大数据, 将大幅提升药品供应保障质量, 满足老百姓的用药需求。完善药品供应保障体系是“健康中国”国家战略重点任务之一, 是推进健康中国建设的重要基础。药品供应保障体系与人民身体健康、国家长治久安密切相关。基于此, 设计了天津市药品供应保障大数据监测分析预警系统, 重点监控药品的多维数据并加以分析, 其中包括药品生产、流通、采购、使用四个维度的数据, 辅助各医疗机构、医保部门、医保部门更好地保障药品的供应。

**关键词:** 药品供应保障; 大数据分析; 监测预警; 系统架构

**中图分类号:** R95

**文献标识码:** A

**DOI:** 10.12230/j.issn.2095-6657.2022.33.038

早在 2007 年, 我国就提出要建设覆盖城乡居民的药品供应保障体系, 2016 年《“健康中国 2030”规划纲要》则将完善药品供应保障体系上升到了国家战略的高度。药品供应保障的方面是与百姓的福祉息息相关的, 是民生保障的重要内容<sup>[1]</sup>。

## 1 系统建设目的与设计

### 1.1 建设目标

构思之初首先明确: 天津市药品供应保障大数据监测分析预警系统, 要在数据质控、监测分析、安全保障、精准预警、报告输出、综合查询及辅助决策支持等方面实现便捷数字化。

### 1.2 设计思路

建立起一套能反映药品供应保障情况的指标体系。要统筹考虑, 系统可重用、可对接, 不是简单的一次实施, 更多地为搭建一个完整的数字化体系架构而进行设计和优化, 以及实施这个体系架构的一套规范、方法和工具。

## 2 数据收集和分析

基础分析数据来自多个渠道, 将不同渠道的数据串联起来使用, 以及每个药品都对应到单一的 YPID 是大数据分析 with 监测的必要条件。目前国家药品基本数据库中收录了 21 万余条赋有 YPID 码的药品信息, 基础分析数据范围拓宽并关联到药品出入库、费用、诊疗用药等数据后, 数量级在千亿级别<sup>[2]</sup>。

### 2.1 多维度药品相关数据收集

大量数据的挖掘是监测分析预警的基石。系统数据源分别来自全市的公立医疗机构上报支持、市医保局结算中心支持、国家药品使用监测平台数据支持、国家短缺药品直报信息平台数据支持、药品生产和经营企业相关数据支持、互联网公开数据检索、文献检索、医疗机构专家支持等渠道。数据范围涵盖了药品信息、医疗机构信息、生产流通信息、处方诊断信息等。

整体上数据来源广、数据覆盖面全、数据体量大。

### 2.2 多维度药品数据质控装载

数据源不同为后期的数据分析带来了巨大的质控难度。首先, 数据来源的主体不同。需要依靠 YPID、机构 ID、医保 ID 等具有相关性的信息进行搭桥链接, 将不同的信息库关联在一起, 最大化利用数据信息, 其次, 数据来源的格式载体不同。数据在处理前, 有 DOC、XLS、CSV、DMP 等格式的文件, 也有需要人工处理电子化的纸质文献材料, 在解析不同格式文件时, 经常遇到字符串行、数据字段不规范或者缺少关键关联信息等问题, 解决起来非常棘手。最后, 数据源更新的时间不同。在数据源主体、数据源载体都不同的基础上, 数据源的更新时间也不固定, 这给数据库的更新方式也带来了难度<sup>[3]</sup>。

### 2.3 多维度药品数据分析探索

分析主体包括政策标签级、行政区域级、医疗机构级、用药群体级、生产企业级、经营企业级、药品属性级、时间级等要素。政策标签级包括国家基本药物目录、国家医保药品目录、国家谈判药物、国家重点监控合理用药目录、国家短缺药品和临床易短缺药品重点监测清单目录、罕见病药品目录、儿童药品目录等各类维度的标签目录; 行政区域级包含了天津市 16 个行政区的分类; 医疗机构级包含了天津市所有公立医疗机构, 按照医疗机构等级、类型、所属行政区、管辖主体划分维度; 用药群体级包含了年龄、性别、区域几个方向; 生产企业级分为国有企业、合资企业、外资企业, 也可分为地产企业和外省企业; 经营企业级则囊括了天津主要的流通企业; 药品属性级细分药品通用名、商品名、剂型、规格、药理功效、YPID、批准文号、医保 ID、ATC 分类、ICD 分类、DDD 值、单价等多方面信息; 时间级包括药品采购时间、订单时间、送货时间、入库时间、出库时间等时间维度。

分析指标含有增长率、年均复合增长率、增长量、年均增

长量、占比、同比、环比等。所有药品基础数据由人工导入系统后，多维数据分析与可视化界面相结合，经系统自动处理后出图、表、文字报告分析结果。

### 3 系统架构设计

设计药品基础数据由人工导入系统后，系统自动处理并得出图、表、文字分析结果。系统在整个工作流程中自动完成基础数据的统计分析工作，减少大量的人工作业，提高工作效率，确保工作质量。

#### 3.1 多维度药品数据分析引擎构成

本系统采用以 SOA（面向服务的架构）软件开发模型设计的中心业务组件搭建按需定制的应用药品数据管理信息系统，以 UKey 密钥为核心建立安全机制，后台以 SqlServer 数据库存储中心药品数据，并以 XML 开放信息交互标准为系统标准信息交互模型，实现未来与其他系统的无缝集成。

相对于 SaaS（Software as a Service 软件即服务），SOA 即 Service Oriented Architecture（面向服务的架构）更加符合系统开发需求。从技术方面来看，SaaS 是简单的部署，不需要购买任何硬件，只需要简单注册即可，但是安全性方面不符合需求，因为核心数据不能由第三方来负责。而 SOA 是以服务作为软件的基础构造块来组装软件，它将系统的不同功能单元（称为服务）通过这些服务之间定义良好的接口和契约联系起来。从网络架构演变的角度看，SOA 从根本上突破了客户机 / 服务器模式。SOA 是一种框架的方法，而 SaaS 是一种传递的模式。但是 SOA 在大数据的处理和渲染展示方面对于硬件系统的要求非常苛刻，而本系统的基础分析数据涉及各个医疗机构多年的药品数据，庞大的基础数据库和海量的运算需求，并不适用于单独的 SOA 框架。所以结合对于大数据处理的需求，设计借鉴了 SaaS 和 SOA 各自的优点，以 SOA 为基础结合 SaaS 方式设计了系统方案。

Web Service 是当前最为成熟的一种 SOA 技术。它将网站的功能封装为可复用的软件服务，提供给其他 Web 应用程序调用。Web Service 技术，能使运行在不同机器上的不同应用无须借助附加的、专门的第三方软件或硬件，就可以相互交换数据或集成。依据 Web Service 规范实施的应用之间，无论它们所使用的语言、平台或内容协议是什么都可以相互交换数据。WSDL（描述文档）是用于描述 Web Service 的一种 XML 格式的语言，用来说明服务端接口、方法、参数和返回值，通知其他的 Web 应用如何相互调用。所以本系统设计的开放信息交互标准也是 XML（Extensible Markup Language）扩展型可标记语言。面向短期的临时数据处理、面向万维网络，是 Soap 的基础。XML 本身是一种格式规范，是一种包含了数据以及数据说明的文本格式规范，它只是定义数据的标机语言，并没有定义如何传输，可以把他放在 http get 的 response 里面，也可

以直接拷贝一个 xml 文件到 U 盘里。其扩展 xsd 定义了 xml 的 schema，而 xslt 又定义了 xml 的转换方式，这些使得 xml 更加强大，更加适合在多系统、多平台之间相互传递。

#### 3.2 系统的安全使用机制

药品数据属于保密数据，因此系统将在身份认证、数据安全存储、软件安全运行三个方面设计安全机制。

首先，系统设计采用 USB Key 方式实现用户身份认证。Ukey 是通过 USB 直接与 PC 连接、具有密码验证功能、可靠高速的小型存储设备，具有极高的安全性和可靠性，并且携带方便、使用灵活。

其次，由于系统软件的基础数据采用外部人工导入本地的方式，所以数据要求安全存储。系统设计将存储设备进行加密分区，对于分区内数据的访问要求密钥确认，通过密钥确认后用户可以拷贝和删除数据文件。

最后，在系统软件设计方面，主要通过身份认证与标识鉴别、资源利用、访问控制、安全审计、密码支持、UKey 加密方案、加密算法、操作日志等要素确保软件系统使用安全。为了防止系统软件被外部病毒或其他恶意软件攻击，系统将在独立的空间运行，可以有效地阻止病毒的侵蚀和软件的攻击，保证系统软件的安全性。

#### 3.3 系统架构设计标准

##### (1) 标准化

系统开发遵循卫生部《（2003-2010 年）全国卫生信息化发展规划纲要》、卫生部 2002 年版《医院信息系统基本功能规范》以及国际、国内相关标准如 ICD-10、ASTM、DICOM3.0、HL7、IHE 等。

##### (2) 平台化

按照 SOA 的软件开发思想，将功能划分成独立的组件，整个信息系统则像搭积木一样由这些组件拼接而成，从而使按照工作实际需要定制信息系统成为可能。

系统组件采用完全开放的接口，后期也可以自己开发或委托其他厂商开发兼容的接口，从而实现最大程度的灵活性。

##### (3) 一致化

系统采用统一的数据窗口服务，保证数据采集、存储、整理、分析、提取、应用的一致性。

### 4 多维度药品数据分析架构层级构成

#### 4.1 前端平台

主窗口界面是所有用户直接操作和使用系统的接口，根据系统用户角色（系统管理员、高级业务人员、普通业务人）的不同，系统软件为每个角色配置了不同的主窗体界面，每个窗体界面的内容和权限都不相同。界面的设计应该秉承操作优化为主，界面友好、直观为辅，坚持设计一个操作简便、直观易懂的界面操作流程。

#### 4.2 用户加密认证

用户插入 UKey 到 PC 的 USB 端口, 认证模块读取 UKey 信息。如无法获取 UKey 信息用户将无法使用系统软件。认证模块将获取的 UKey 信息提交后台系统进行认证, 认证后模块将返回系统的用户信息, 包括用户名称、用户类型及用户权限。

#### 4.3 基础数据接入

用户按照模板 excel 文件汇总数据, 模块将指定 excel 文件数据导入到指定系统数据库中, 系统将自动检测 excel 文件内容是否与指定数据库匹配, 如不匹配系统将报错, 并退出数据导入过程。数据导入共分药品库数据、基础分析数据、监测药品清单数据等六类数据, 基础数据需确定数据唯一标识用以保证数据不会重复插入数据, 字典数据每次导入将覆盖或更新以往数据, 基础分析数据根据文件名进行合并。

#### 4.4 数据字典管理模块

此模块用于系统软件的数据字典维护, 前期由系统管理员完成基础数据库的建立和维护, 其中包括建立管理员基本资料并为其分配相应权限、导入药品目录、导入医院目录等。在基础数据库建设完成后, 此模块也负责基础数据的维护工作(增、删、改), 包括药品字典、药理/功效、酸根盐基、药品剂型、药品类别, 剂型规格。由于历史数据过于庞大, 人为录入工作量很大, 所以设计 excel 导入功能以减轻初期建立数据库的工作量。

#### 4.5 数据自动模糊匹配

此模块根据数据匹配字典, 自动将非标准数据匹配为标准数据, 匹配内容包括药品名称、药品单位、药品剂量等。

#### 4.6 数据分析

在数据分析模块, 系统基于多维药品供应保障大数据, 开发实施了目录查询分析、药品采购和使用分析、处方诊断分析、医联体药品配备分析、数据分析报告输出、药品供应保障预警六大分析模块。

#### 4.7 看板系统

在系统的首页方案设计看板展示页面, 利用看板展示方式的直观效果, 利用可视化管理的力量来帮助团队最大化时间和效率。在首页看板展示中建立信息共享机制和平台, 实现现有资源及药品信息的数字化监测与控制。在首页的可视化设计中, 系统采用了 Highcharts 可视化工具, 这是一个用纯 JavaScript 编写的图表库, 能够很简单便捷地在 Web 网站或是 Web 应用程序添加有交互性的图表。它界面美观, 由于使用 JavaScript 编写, 所以不需要像 Flash 和 Java 那样需要插件才可以运行, 而且运行速度快。另外, HighCharts 还有很好的兼容性, 能够完美支持当前大多数浏览器, 相较于其他工具, HighCharts 具有以下特点。

(1) 兼容性。HighCharts 采用纯 JavaScript 编写, 兼容当今大部分的浏览器, 包括 Chrome、Safari、IE 等。

(2) 图表类型。HighCharts 支持图表类型, 包括曲线图、

区域图、综合图表等。

(3) 不受语言约束。HighCharts 可以在大多数的 WEB 开发中使用, 支持 ASP、PHP、Java 等多种语言。

(4) 提示功能。HighCharts 生成的图表中, 可以设置在数据点上显示提示效果, 即将鼠标移动到某个数据点上, 可以显示该点的详细数据, 并且可以对显示效果进行设置。

(5) 放大功能。HighCharts 可以将大量数据集中显示, 并且可以放大某一部分的图形, 将图表的精度增大, 进行详细显示, 可以选择横向或者纵向放大。

(6) 外部数据。从服务器载入动态数据。

### 5 系统实施后的运行状态

该系统是药品采购、使用、流通等数据进行数字化集中管理的平台。通过使用本系统可以标准化不同来源的非标数据、明确工作流程、简化大量的人工数据统计工作, 大幅提高工作效率, 杜绝统计和计算错误。系统可以根据工作要求按照药品、采购单位和生产企业等要素自由输出统计报表、文字报告和图表, 按照预先设定的清单规则, 根据基础数据分析药品的使用情况, 做出准确的分析预测。

### 6 结语

天津市药品供应保障大数据监测分析预警系统的建设是助力改革完善药品供应保障机制的进一步探索。建成后的系统既可以服务医疗机构有效抵御药品供应保障的风险, 也可以作为临床用药或科研的有效工具。通过大数据多维度分析来了解重点专科疾病的用药趋势, 还可以作为临床综合评价的有力助手, 更可以作为医政、医保部门调整政策的辅助, 从而促进医疗卫生事业的进步, 更好地满足人民健康需求。总的来说, 借助天津市药品供应保障大数据监测分析预警系统这一平台, 充分挖掘大数据的潜在价值, 以落实建立健全短缺药品供应保障体系和机制的责任为核心, 为全面推进健康中国建设提供有力支撑。

#### 参考文献:

- [1] 陆建成, 申远, 王晶, 等. 国家药品使用监测平台中的多维数据分析设计与实现[J]. 中国数字医学, 2020, 15(06): 25-28.
- [2] 王强, 李俊杰, 陈小军, 等. 大数据分析平台建设与应用综述[J]. 集成技术, 2017, 5(02): 2-18.
- [3] 陆建成, 申远, 王晶, 等. 国家药品使用监测平台体系架构设计与实施[J]. 中国卫生信息管理杂志, 2020, 17(02): 163-167.

作者简介: 王博飞(1994-), 男, 安徽亳州人, 大学本科, 助理馆员, 主要从事药品供应保障研究。