

免疫学数据库和分析门户 ImmPort 建设经验与启示*

胡拯涌¹ 周伟² 王安然¹ 段一凡¹ 胡万飞¹ 吴思竹¹

(¹ 中国医学科学院/北京协和医学院医学信息研究所 北京 100020

² 中国医学科学院国家人口健康科学数据中心 北京 100005)

[摘要] **目的/意义** 总结免疫学数据库和分析门户 ImmPort 的建设经验, 为我国建设大型免疫学数据库提供经验和参考。**方法/过程** 全面分析 ImmPort 的整体架构, 讨论各模块功能, 同时梳理其数据收集、组织、共享和分析复用的标准化数据流, 总结 ImmPort 平台实践成果。**结果/结论** 我国在大型免疫学数据库建设过程中应重视数据标准化组织和建模; 积极引用标准化术语提供语义支持; 在分类分级管理中共享数据, 同时开发配套工具体系, 保障免疫学数据安全有效地共享和利用。

[关键词] 数据共享; 数据标准化; FAIR 原则; 免疫学数据库

[中图分类号] R-058 **[文献标识码]** A **[DOI]** 10.3969/j.issn.1673-6036.2024.08.004

Construction and Implications of the Immunology Database and Analysis Portal: ImmPort

HU Zhengyong¹, ZHOU Wei², WANG Anran¹, DUAN Yifan¹, HU Wanfei¹, WU Sizhu¹

¹ Institute of Medical Information, Chinese Academy of Medical Sciences & Peking Union Medical College, Beijing 100020, China; ² National Population Health Data Center, Chinese Academy of Medical Sciences, Beijing 100005, China

[Abstract] **Purpose/Significance** By summarizing the construction experience of the immunology database and analysis portal (ImmPort), the study aims to provide insights and references for the development of a large-scale immunology database in China.

Method/Process It comprehensively analyzes the architecture of ImmPort and the functionalities of its various modules, and delineates the data flow for data collection, organization, sharing, and analysis within the database. Finally, it summarizes the practical achievements of the ImmPort platform. **Result/Conclusion** In the process of constructing the immunology database, our country should prioritize the standardization of data organization and modeling. Standardized terminology should be actively adopted to provide semantic support, and data sharing should be carried out in the management of classification and grading, while developing a supporting tool system to ensure the safe and effective sharing and utilization of immunology data.

[Keywords] data sharing; data Standardization; FAIR principles; immunology database

[修回日期] 2024-06-11

[作者简介] 胡拯涌, 硕士研究生, 发表论文 1 篇; 通信作者: 吴思竹。

[基金项目] 中国医学科学院医学与健康科技创新工程项目 (项目编号: 2021-I2M-1-057)。

1 引言

免疫力是全民健康的核心问题之一, 实现《“健康中国 2030”规划纲要》等战略迫切需要评

估、干预和重建个体和群体免疫力,开展基于免疫力干预的健康维护及疾病防治新策略。随着高通量技术的发展和免疫系统研究的深入,各类免疫学数据量不断增长^[1-2],基于免疫学大数据,免疫力机制解码和免疫力数字化呈现逐渐成为可能。然而目前我国缺少高质量、标准化的免疫学数据,对免疫学数据的整合、共享水平以及挖掘利用程度无法支持免疫力数字化研究的进一步发展。因此亟须建设拥有多来源数据整合、共享和分析能力的大型免疫学数据库。

目前国际上免疫学数据库建设更多面向免疫学特定研究领域,并且主要收集整理公共数据,基于用户汇交数据的数据库较少。如免疫表位数据库(the immune epitope database, IEDB)^[3],由美国国家过敏和传染病研究所(The National Institute of Allergy and Infectious Diseases, NIAID)资助建设,主要提供抗原表位、抗原识别和抗原抗体相互作用等相关数据,还提供在线分析工具以支持预测新抗原表位和相关性序列分析等,其数据来源于对PubMed等数据库中相关文献图片、文本和表格的收集整理。我国也有类似免疫学特定领域数据库的建设研究,如Chen S Y等^[4]建设的人类T细胞受体序列数据库,从美国国家生物技术信息中心的公共T细胞受体测序数据集和其他T细胞受体序列数据库收集整理相关数据;Liu Z等^[5]建设的中医肿瘤免疫数据库,主要收集文献中的肿瘤免疫靶点、中药和复方制剂以及对应化学成分数据,支持探索中医调节癌症免疫微环境的分子机制。

免疫学数据库和分析门户(the immunology database and analysis portal, ImmPort)是整合临床试验、免疫表型和基因组等多模态免疫学数据的综合性数据库,不仅免疫学数据类型多样,涵盖多个免疫学研究领域,并且有完善的数据汇交、管理和共享复用的标准化流程,有效推动了高质量免疫学数据的传播和复用,为免疫学研究发展提供数据支持,是多个研究机构和出版社在免疫学领域的推荐存储数据库。探索ImmPort平台的建设经验,对解决我国免疫学数据面临的问题以及建设国家级综合免疫学数据库有重要意义。本文全面分析ImmPort

的技术架构、主要功能及数据流程,总结其建设成果和经验,以供我国免疫学及其他特色医学数据库参考。

2 ImmPort 概述

ImmPort是2004年由NIAID资助建设的数据库,最初用于共享其内部项目的临床和基础数据,后逐渐发展为收集、存储和共享多来源、多模态免疫学数据的学科级开放数据库^[6]。目前ImmPort为NIAID等美国国立卫生研究院下属研究机构以及私人资助研究项目提供长期、可持续的免疫学数据存储服务,并对研究社区开放免疫学数据资源共享和分析服务。为推动可发现(findable)、可获取(accessible)、可互操作(interoperable)和可复用(reusable)的FAIR原则^[7],ImmPort构建了一套数据收集、组织、共享复用的标准化流程,并进行了较好实践。经过近20年的建设积累,ImmPort现有1257项注册研究,截至DR 50.2发布(2024年2月),已共享其中890个研究数据集。共享数据资源中包含近10万例受试者,分别为8万例人类受试者和其他非人类受试者,覆盖流感、肾移植等160种疾病和病症。按数据类型划分,主要为研究元数据和结果数据,包含3000余种试验和相应700万余份试验结果,主要为流式细胞术、酶联免疫吸附试验、RNA测序等方法的结果数据。

3 平台技术架构

目前ImmPort平台根据数据安全性,将私有数据和共享数据分别部署在NIAID基础设施环境和亚马逊云环境,整体项目架构,见图1。在亚马逊云(Amazon web service, AWS)环境ImmPort实例化了两个AWS Aurora MySQL数据库,分别用于存储只读的共享数据和云端应用的读写记录,并托管共享数据和数据浏览器等面向公众的应用程序。共享数据是基于Java Spring的Web应用,用于支持共享数据集检索查询,在表示层使用Java Script AJAX框架的JQuery和Angular以及Spring MVC实现,在数

据持久层通过 Hibernate 结合 Spring JDBC 实现与底层 Aurora MySQL 数据库的交互，数据浏览器使用类似技术框架实现，但主要与 NIAID 环境中的 Aspera 服务器交互，通过 Aspera connect 工具实现共享数

据的安全快速传输。亚马逊云环境中托管多个应用程序编程接口（application programming interface, API），如 Data Query API 用于对 Shared Data 应用实现编程访问，Data API 用于浏览和下载共享数据等。

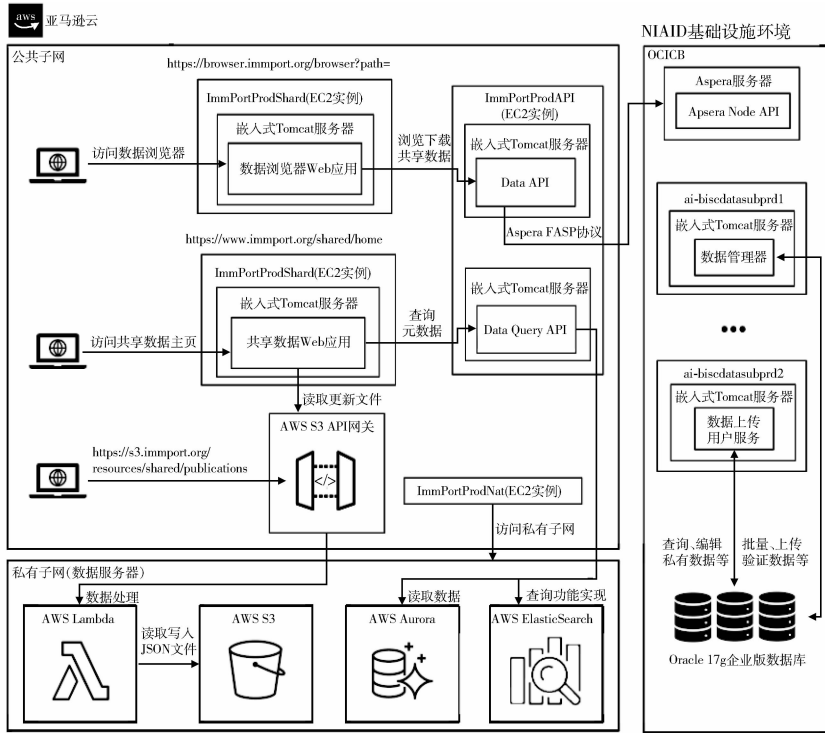


图 1 ImmPort 整体技术架构

NIAID 基础设施环境由网络基础设施和计算生物学办公室（Office of Cyber Infrastructure and Computational Biology, OCICB）管理，使用 Oracle 17g RAC 集群存储和管理私有数据和读写操作，并托管数据提交、批量更新、用户管理和 Core Query API 等应用。OCICB 数据库依靠实时应用集群（real application clusters, RAC）环境提供系统负载分配和负载均衡以及故障转移等功能，核心数据库 CORE_IMPORT 记录所有读/写操作，并拥有相应镜像数据库。

4 主要功能及数据流程

ImmPort 生态系统由 4 个子模块构成：私有数据、共享数据、数据分析和其他资源，数据收集、组织、共享和分析复用的流程贯穿 4 个子模块，依靠各模块功能实现，见图 2。

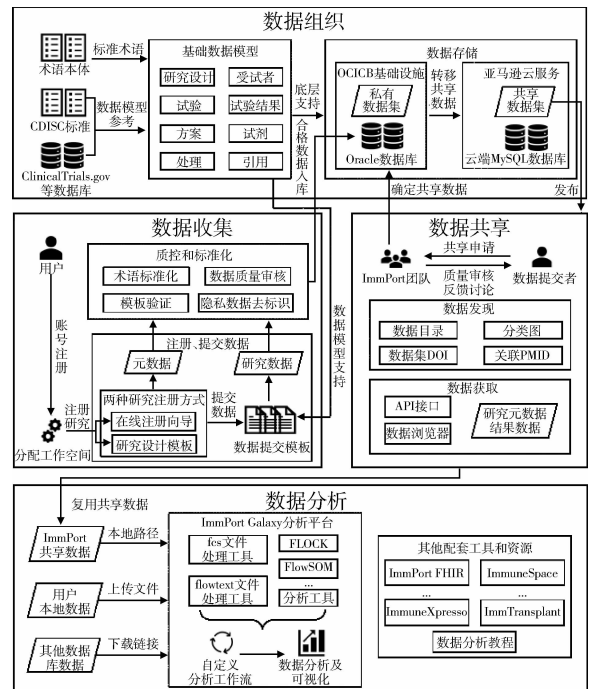


图 2 ImmPort 数据流程

4.1 数据收集

私有数据模块提供研究注册、数据上传、模板验证、工作空间管理、注册研究检索和浏览等功能，用于实现 ImmPort 数据流中数据收集、标准化和质控。ImmPort 数据收集包括 4 个步骤：创建账户、获取工作空间、注册研究和上传数据。用户完成账户注册并申请获得私有工作空间后即可开始注册研究，ImmPort 为注册研究提供两种方式：研究注册向导和研究设计模板。研究注册向导是 ImmPort 开发的在线注册工具，由 11 个连续页面组成，详细收集基本研究元数据，用户根据页面导航填写元数据并验证即可完成注册。也可以下载 basic_study_design 和 protocol 两个研究设计模板后本地填写，其收集的信息与在线工具相同，模板提供 xlsx 和 txt 两种格式，上传时需统一使用 txt 格式，填写并经过在线验证工具完成模板验证后即可上传完成注册。已注册研究保存在注册用户的私有工作空间中，只有该用户拥有管理权限。私有数据模块支持浏览所有注册研究的基本信息，如资助机构、主要研究者等。

完成研究注册后用户即可提交研究数据，研究数据必须使用数据提交模板完成提交，除两个研究设计模板外，ImmPort 提供 66 个其他模板用于收集研究数据，不同模板依据收集数据类型包含不同字段，但都包含 Schema Version 字段进行版本控制，该字段提示模板引用的数据库和标准术语的版本信息，帮助在线模板验证工具等识别和匹配。因模板间存在依赖关系，数据模板需参考 ImmPort 数据模型按特定顺序上传，典型顺序为研究、受试者、干预、不良反应、评估和实验室测试，上传时大于 512 Mb 的文件需要使用 Aspera Connect 工具。若研究数据已在其他数据库共享，ImmPort 支持注册研究时仅填写数据集 URL 以避免数据冗余。

为保护受试者的个人隐私，数据提交模板中会避免收集个人信息，且提交到 ImmPort 中的数据将依据健康保险流通与责任法案^[8]被去标识化，包括受试者的年龄、地址和电话等人口学信息，以及受试者在来源系统中的唯一标识符，ImmPort 工作人员将对隐私数据模糊处理或直接删除。

4.2 数据组织

为提高研究数据可用性，保证每项研究拥有统一、规范且丰富的描述信息，ImmPort 参考临床数据交换标准协会标准、通用生物模型数据库和基因综合表达数据库等设计数据模型，见图 3。按照描述内容的类别将数据模型分为研究设计、受试者等 8 个部分，见表 1，其中引用部分为其他数据模型内容提供术语支持，该部分参考细胞本体^[9]、疫苗本体^[10]和蛋白质本体^[11]等本体构建受控词汇表。

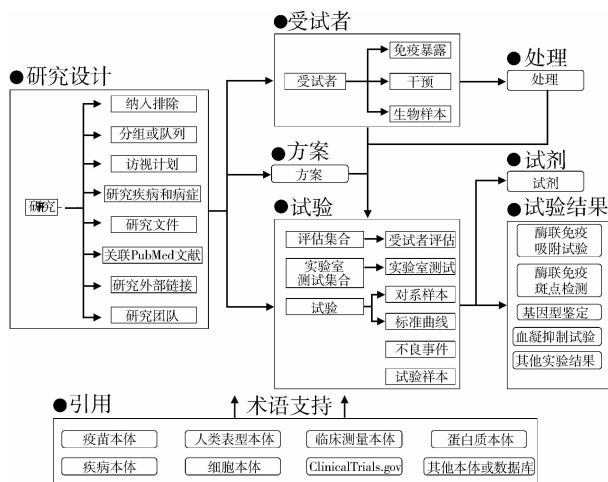


图 3 ImmPort 数据模型结构

表 1 ImmPort 数据模型分类和功能

数据模型分类	功能
研究设计	描述研究的上下文信息，以及分组、纳入排除、来源链接等
受试者	描述人类和动物受试者信息，包括干预、免疫暴露和生物样本等
试验	描述研究的实验室测试、不良事件和对照样本等信息
试验结果	描述流式细胞术等试验的结果
方案	描述研究的方法、步骤等
试剂	描述试验中使用的试剂信息
处理	描述试剂（如化合物）作用于样品的持续时间等信息
引用	存储各类本体或数据库的术语信息，形成内部受控词汇表

数据模型是 ImmPort 底层关系数据库的基本框架，数据模型各部分包含相应的关系表，以及链接不同关系表的中间表。数据提交模板也基于该数据

模型设计, ImmPort 积极与标准机构和领域专家合作, 及时更新和扩展数据模板, 以满足数据提交者的需要并准确收集元数据^[12]。私有数据完成验证后将上传到 OCICB 管理的 Oracle 17g 数据库中, OCICB 为私有数据集提供最高等级保护措施, 严格控制数据访问。只有当研究数据完成共享流程并决定数据发布时, 数据才会通过亚马逊云服务向研究社区开放, 转为共享数据集。

4.3 数据共享

共享数据模块用于实现数据库中共享数据的发现和获取, 集成共享数据的检索、浏览和获取等功能。ImmPort 数据资源覆盖范围广, 为提高可发现性, 从研究领域、研究对象、疾病和病症等维度统计数据, 使用矩阵、柱状图和散点图等多种方式展示数据资源, 同时通过检索框和数据目录实现对数据资源的高效灵活检索。

当用户准备共享私有数据时需向 ImmPort 团队提出申请, 团队会再次检查数据质量, 并与用户确定共享时间。ImmPort 每季度公开一次主要数据, 其他月份发布小型数据, 发布时 ImmPort 工作人员将使用 Datacite 平台为共享数据集创建数字对象唯一标识符。共享数据模块提供 Aspera Connect 下载工具和 API 两种方式获取数据, 可获取文件包括研究元数据文件和结果文件, 元数据主要以 txt 格式下载, 结果文件则以 fcs 等特定格式下载。用户可在 archive 文件路径下获取研究历史版本的元数据, 提高 ImmPort 中数据的可获取性。ImmPort 没有数据集层面的个性化数据共享协议, 而是发布了统一的数据使用协议, 同时参照美国国立卫生研究院数据管理和共享计划指南, ImmPort 也发布了数据管理和共享计划文件, 以提供规范。

4.4 数据分析

为辅助免疫学数据的分析和挖掘, ImmPort 开发了配套工具、教学文档等多种资源, 集成在数据分析模块和其他资源模块。数据分析模块即 ImmPort Galaxy 工具, 其他资源模块保存了 ImmPort 中的数据分析教程、辅助文档和部分特殊研究项目

信息, 如流式细胞数据深度学习的代码讲解, 也可在该模块浏览引用 ImmPort 的相关出版物。

ImmPort Galaxy 是使用 Galaxy 框架开发的流式细胞术数据分析工具, Galaxy 是基于 Web 的公共数据分析平台, 使用简易的图形界面操作代替复杂的命令行指令, 旨在让编程技能较弱的生物学家也能够在线计算分析^[13]。Galaxy 框架有高度的可定制性、可扩展性, 支持集成各种分析工具、软件。ImmPort Galaxy 继承了 Galaxy 的框架优点, 在其基础上集成了流式细胞术分析软件和工具, 如 FLOCK^[14] 和 FlowSOW^[15] 等。该模块支持上传本地数据或 URL 两种方式导入外部数据, 而 ImmPort 库中共享数据只需复制其文件路径即可直接导入。FCS File Tools 和 Flow Text File Tools 板块可对 fcs 文件进行基本处理, Flow Analysis Tools 和 Flow Visualization Tools 板块提供分析以及结果可视化工具。WorkFlows 工作流系统允许用户对以上处理和分析工具自由组合, 实现批量化分析, 每步均会被历史记录功能保存, 保证研究可复现。

5 实践成效及待完善之处

ImmPort 在多方面取得较好实践成果, 被多个期刊和机构推荐使用, 包括 Nature Scientific Data、PLOS、Elsevier 和 Springer Nature 等。2017 年 ImmPort 通过 CorerTrustSeal 国际认证, 同时也被 Re3data 和 FAIRsharing 数据仓储目录收录, 代表其数据资源收集、管理和长期保存等能力得到国际认证。ImmPort 的关联出版物已有 1 374 篇, 其中 835 篇直接与具体研究数据集相关联, 有效推动了免疫学数据传播和新研究课题的提出。ImmPort 开发的一系列配套工具也被广泛应用, 按工具类型可划分为术语工具 (如 Cell Ontology Browser)、知识库工具 (如 Immune Xpresso 和 Cytokine Registry)、特殊资源集工具 (如 10k Immunome 和 ImmPort FHIR) 和辅助分析工具 (如 MetaCyto R 包), 从多个维度为数据分析和挖掘提供支持。

在取得较多建设成果的同时, ImmPort 还存在可提升之处。首先在数据收集方面, 用户友好性有

待提高,目前数据提交模板的选择和提交顺序需要用户参考数据模型自行判断,可以参考在线研究注册工具,开发结果数据提交的步骤引导工具,辅助用户准确快速完成结果数据提交。其次在数据共享方面,数据集的辅助信息不足且缺少推荐服务。建议为研究数据集提供浏览量、下载量、被引用量等信息,结合质量审核结果实现高质量、高热度数据集的排行和推荐服务。最后对于数据分析,ImmPort 数据分析工具类型局限于流式细胞相关分析,而作为综合性的免疫学数据库,ImmPort 包含临床试验、基因表达和免疫细胞表型等多种类型数据,因此应利用 Galaxy 框架的较好扩展性,整合差异表达分析、表型分析等更多类型数据分析工具,辅助用户更全面地利用库中多类型免疫学数据。

6 对我国免疫学及特色数据库建设的启示

6.1 重视数据标准化组织和语义建模

ImmPort 在免疫学数据资源组织方面有较好实践,通过数据模型和标准化模板实现多源免疫学数据的规范化收集和管理。首先 ImmPort 基本数据模型参考了临床数据交换标准协会、基因综合表达数据库等其他数据模型,能够较好地覆盖免疫学研究相关元数据。其次基于该数据模型,ImmPort 团队积极同领域专家合作,开发出一套标准数据模板并不断更新、改善,通过强制使用数据模板准确收集数据模型中的元数据和结果数据,实现数据标准化。最后 ImmPort 在多个方面参考 FAIR 原则组织数据,包括为数据集创建唯一标识符、存档历史版本元数据等,在标准化基础上进一步提高数据的可发现、可获取、可互操作和可复用性。我国免疫学数据资源建设中也面临数据多源异构、质量参差、难以统一的问题,因此应特别重视数据的标准化收集存储和组织管理,加强数据的标准化,结合免疫学数据特点设计好数据资源的组织结构,并与研究人员、免疫学和数据科学专家等及时深入沟通,做好数据模型及数据模板的开发和持续更新,并在数据收集、组织和管理过程中严格执行,从而保障收集到高质量的标准化免疫学数据。

6.2 复用及开发标准术语增强数据语义

在处理多来源数据时,不同研究间使用的术语含义并不完全相同,数据间缺乏可比性,无法直接集成分析^[16]。为解决语义模糊问题,ImmPort 引用多个医学领域本体和主题词表,通过在底层数据库中构建内部受控词表实现语义规范。受控词表存储术语名称、术语描述和术语来源链接等内容,通过将数据表和受控词表映射实现对数据资源的语义标准化。除引入外部术语资源外,ImmPort 还参考其数据资源开发了抗体本体,用于支持免疫学研究中单克隆抗体的标准化表示^[17]。我国大型免疫学数据库建设时也应考虑语义标准化问题,通过构建内部受控词表等方式注释和扩展,但领域内中文标准术语资源相对匮乏,可以引用国外成熟的细胞本体、疾病本体等进行本地化应用。同时发挥数据资源量大和类型丰富的优势,根据免疫学数据特点尝试构建特定领域本体、数据标准等,一方面推动我国免疫学数据标准制定,促进免疫学数据的标准化发展;另一方面丰富中文标准术语资源。

6.3 分级分类实现免疫学数据安全共享

ImmPort 为私有数据和共享数据提供了不同的管理方案,私有数据托管在 OCICB 管理的 Oracle 数据库中,对数据严格访问控制并记录所有读写操作,隐私数据和敏感数据会被模糊或删除处理,这些措施为私有数据提供了最高级别的安全保护,当私有数据集确定共享时才会被转移到亚马逊云端的 MySQL 数据库中共享,通过只读 API 访问,这种隔离系统设计保证了 ImmPort 对各类数据的最大程度控制。我国越来越重视人类遗传资源、隐私和敏感数据安全,已发布《个人信息保护法》《人类遗传资源管理条例》等法律法规,因此在组织和共享免疫学数据时,不仅要结合最新的政策和法律法规要求,也应对数据进行细粒度的分类分级管理,安全管控包含敏感和隐私信息的数据,区分私有数据和共享数据,并制定数据全生命周期流程和管控的制度和方法,建立有效的数据共享、使用规范和协议,保障免疫学数据安全可控地利用和共享。

6.4 开发配套工具推动数据分析利用

ImmPort 另一个特点和优势是其自主开发了多个免疫学数据的配套工具,形成数据资源结合分析工具的综合平台。ImmPort 的工具体系以 ImmPort Galaxy 分析平台为核心,通过使用公开框架 Galaxy 大大降低开发和维护工作量,使团队可以专注于对 ImmPort Galaxy 扩展,开发结果可视化等其他功能。围绕该分析平台,ImmPort 工具体系还包括一系列多元工具辅助用户,如术语工具、知识库工具和辅助分析工具等,从多个维度为 ImmPort 数据分析和挖掘提供支持。目前我国除了免疫学数据资源整合和共享存在问题外,对免疫学数据的利用也不够充分,分析工具和其他多元工具主要依赖于国外平台,数据利用能力的发展受制于国外。因此,从数据利用的便捷性、安全性考虑,应当开展我国免疫学数据分析挖掘关键技术和工具系统开发,并将其与数据库整合、关联,从而提高数据利用率,推动免疫学科研人员的科技创新。

6.5 建立健全特色数据库建设体制机制

我国在免疫学等医学专业数据库建设实践方面经验还不足,缺乏具有国际影响力的特色数据库,在医学细分领域的数据库汇交、质量控制和数据管理等方面缺少统一标准和完善机制。以免疫学为例,首先应参考国外成熟的免疫学数据库,充分考虑各类免疫学研究和数据特点,开发免疫学基本数据模型并不断更新完善。此外,应建设一套系统、完善的数据汇交、管理流程,完善隐私数据保护法律法规体系,健全免疫学数据分类分级共享机制,鼓励相关机构在建设过程中总结经验形成内部标准、团体标准和行业标准等,推动免疫学数据集标准化建设。最后应建立我国的公共数据库评级评价机制和注册目录,鼓励数据库不断提高数据治理水平,同时提高国内公共数据库的可发现性,避免出现已有国内相关数据库而科研人员不知道或找不到的情况,逐渐减少对国外数据库的依赖。

7 结语

ImmPort 平台依靠数据标准化组织和建模以及内部受控词表的语义支持,在数据格式、内容和语义等层面实现了元数据和研究数据的标准化收集、整合和共享,同时开发了配套工具鼓励用户进行免疫学数据复用,取得了较好的建设成果。相比之下我国缺乏综合性免疫学数据库以及免疫学数据的配套分析利用工具,无法提供高质量免疫学数据和分析服务。因此本研究细粒度分析 ImmPort 平台的整体架构、数据流以及实践成效,针对现存问题为国内免疫学及特色数据库建设提出建议。通过综合性免疫学数据库建设,响应国家《“数据要素 x”三年行动计划(2024—2026 年)》等规划,推动免疫学数据有序开放共享,促进数据要素价值释放,以数据库和高质量科学数据集为依托,实现免疫学研究创新发展。

作者贡献: 胡拯涌负责论文撰写、研究结果可视化;周伟负责提供指导;王安然、段一凡、胡万飞负责资料收集;吴思竹负责提供指导、论文修订。

利益声明: 所有作者均声明不存在利益冲突。

参考文献

- 1 GEORGIU G, IPPOLITO G C, BEAUSANG J, et al. The promise and challenge of high-throughput sequencing of the antibody repertoire [J]. *Nature biotechnology*, 2014, 32 (2): 158-168.
- 2 WARREN E H, MATSEN F A T, CHOU J. High-throughput sequencing of B- and T-lymphocyte antigen receptors in hematology [J]. *Blood*, 2013, 122 (1): 19-22.
- 3 VITA R, MAHAJAN S, OVERTON J A, et al. The immune epitope database (IEDB): 2018 update [J]. *Nucleic acids research*, 2019, 47 (D1): D339-D343.
- 4 CHEN S Y, YUE T, LEI Q, et al. TCRdb: a comprehensive database for T-cell receptor sequences with powerful search function [J]. *Nucleic acids research*, 2021, 49 (D1): D468-D474.
- 5 LIU Z, CAI C, DU J, et al. TCMIO: a comprehensive database of traditional Chinese medicine on immuno-oncology

- [EB/OL]. [2024-03-24]. <https://doi.org/10.3389/fphar.2020.00439>.
- 6 BHATTACHARYA S, ANDORF S, GOMES L, et al. ImmPort: disseminating data to the public for the future of immunology [J]. Immunologic research, 2014, 58 (2/3): 234-239.
- 7 WILKINSON M D, DUMONTIER M, AALBERSBERG I J, et al. Comment: the FAIR guiding principles for scientific data management and stewardship [J]. Scientific data, 2016, 3 (1): 160018.
- 8 ATCHINSON B K, FOX D M. From the field: the politics of the health insurance portability and accountability act [J]. Health affairs, 1997, 16 (3): 146-150.
- 9 MEEHAN T F, MASCI A M, ABDULLA A, et al. Logical development of the cell ontology [J]. BMC bioinformatics, 2011, 12 (1): 1-12.
- 10 LIN Y, HE Y. Ontology representation and analysis of vaccine formulation and administration and their effects on vaccine immune responses [J]. Journal of biomedical semantics, 2012, 3 (1): 17.
- 11 NATALE D A, ARIGHI C N, BLAKE J A, et al. Protein ontology: a controlled structured network of protein entities [J]. Nucleic acids research, 2014, 42 (D1): D415-D421.
- 12 DUNN P J, THOMSON E, CAMPBELL J, et al. ImmPort: shared research data for bioinformatics and immunology [C]. Washington: The 2015 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), 2015.
- 13 COMMUNITY T G. The Galaxy platform for accessible, reproducible and collaborative biomedical analyses; 2022 update [J]. Nucleic acids research, 2022, 50 (W1): W345-W351.
- 14 QIAN Y, WEI C, EUN-HYUNG LEE F, et al. Elucidation of seventeen human peripheral blood B-cell subsets and quantification of the tetanus response using a density-based method for the automated identification of cell populations in multidimensional flow cytometry data [J]. Cytometry part B: clinical cytometry, 2010, 78 (S1): S69-S82.
- 15 VAN GASSEN S, CALLEBAUT B, VAN HELDEN M J, et al. FlowSOM: using self-organizing maps for visualization and interpretation of cytometry data [J]. Cytometry A, 2015, 87 (7): 636-645.
- 16 DUGAS M, HEGSELMANN S, RIEPENHAUSEN S, et al. Compatible data models at design stage of medical information systems: leveraging related data elements from the MDM portal [EB/OL]. [2024-03-24]. <https://doi.org/10.3233/shti190194>.
- 17 BHATTACHARYA S, DUNN P, THOMAS C G, et al. ImmPort, toward repurposing of open access immunological assay data for translational and clinical research [J]. Scientific data, 2018, 5 (1): 180015.

2025 年《医学信息学杂志》征订启事

《医学信息学杂志》是国内医学信息领域创刊最早的国家级期刊。主管单位：国家卫生健康委员会；主办单位：中国医学科学院。中国科技核心期刊（中国科技论文统计源期刊），美国《化学文摘》、《乌利希期刊指南》及 WHO 西太区医学索引（WPRIM）收录，并收录于国内 3 大数据库。主要栏目：专论，医学信息研究，医学信息技术，医学信息资源管理与利用，医学信息教育等。读者对象：医学信息领域专家学者、管理者、实践者，高等院校相关专业的师生及广大医教研究人员。

《医学信息学杂志》国内外公开发行人，每册定价：20 元（月刊），全年 240 元。邮发代号：2-664，全国各地邮局均可订阅。也可到编辑部订购：北京市朝阳区雅宝路 3 号（100020）中国医学科学院医学信息研究所《医学信息学杂志》编辑部；电话：010-52328686，52328687。

《医学信息学杂志》编辑部