

多学科诊疗模式下医疗信息数据融合应用发展探析

曹磊

宁夏医科大学总医院 宁夏 银川 750001

摘要：多学科诊疗已经逐渐成为处理疑难杂症的重要手段，在此过程中需要大量来自电子病历、检验、检查、影像、病理、基因组等多种来源且类型多样化的医疗信息进行有效汇集并加以合理应用。但是由于目前医疗数据在体系建设、语义标准化、访问接口以及质量管理等方面都存在问题，极大地阻碍了MDT由“经验驱动”发展到“数据驱动”。基于信息科学与管理科学，文章分析了多学科诊疗中医疗信息数据融合的概念、难点以及发展趋势，提出了包括数据整合层、语义对齐层、知识表示层和决策服务层四层融合方案，并讨论了隐私计算、可解释人工智能和联邦学习对于保证数据安全性及模型可靠性的作用。本文指出，数据融合不仅仅是一项技术任务，更是一种需要改变整个医疗机构工作方式、制定统一的标准和建立新的治理结构的工作。

关键词：多学科诊疗；医疗信息；数据融合；数据治理；决策支持

引言

多学科诊疗模式（MDT）由多领域专家组成固定团队，为患者制定个性化治疗方案，其核心优势在于整合跨学科知识。然而，该优势的发挥依赖于信息的完整性与及时性——所有成员需迅速掌握患者的病史、检验、检查、影像、病理、基因及既往治疗等全面数据。传统纸质或初级电子化系统中，信息分散于“孤岛”，查阅耗时且易遗漏关键细节，影响MDT效率与决策质量。当前，随着医院信息化建设推进，电子病历系统、影像归档和通信系统、实验室信息管理系统、基因检测分析系统等提供了丰富的大数据资源。但因各系统标识、术语、时间轴不统一及数据质量差异，数据难以有效融合。因此，如何在尊重各专业信息表达习惯的前提下，构建可互操作、可计算的一致性融合数据环境，成为医学信息学研究热点。

1 多学科诊疗模式对医疗信息数据融合的内在需求

1.1 MDT决策范式的信息特征与数据需求

MDT决策是一种典型的“信息密集型”合作方式，它对信息的需求具有多维性、时序性和关联性的特点。多维性是指信息包括了解剖学层面（影像学数据）、分子层面（如病理及基因组学信息）、功能层面（如实验室检验、生理监测及生命体征参数）以及社会心理层面（如患者心理状态与社会支持因素）等；时序性是指需要以时间为轴来反映疾病的进展及疗效，从而进行分析评价；关联性在于不同类型的资料之间存在因果或者相关的关系，例如某种基因改变会导致什么样的药物代谢表型等。传统的单个科室诊疗只需要查看本专科的诊疗指南即可，但是MDT则需要在一个共同的认识基础上将

这些多方面的信息综合在一起，这就对查阅诊疗数据的速度、数据表现形式（比如同时显示一张影像和对应的病理切片）以及诊疗方案推导能力都提出了硬性要求。

1.2 融合深度与MDT成熟度等级的对应关系

根据医疗信息学会提出的临床决策支持成熟度模型，MDT数据融合程度可以分为五级：L0级是纸质文档在现场分发；L1级是各个系统分别访问，不能跨系统之间进行交互；L2级是有一个患者主索引，在这个基础上可以同时打开多个系统的窗口；L3级是有疾病的主题的数据视图，根据MDT讨论的内容预先准备好一系列的信息组合方式；L4级是在语义上进行了融合，可以进行跨模态的检索以及推导，而且还可以及时地生成量化的总结报告^[1]。目前我国大多数MDT小组处于从L1到L2的过程之中，而一小部分先进机构已经达到了L3级。融合的程度越高，对于工作效率的影响越大，同时也对决策的质量有重要影响。

1.3 从数据共享到知识协同的演进逻辑

早期MDT信息化建设以数据共享为主，使各个学科的医生能够“看到”对方的数据。但是共享并不等同于融合——融合是在此基础上实现“互操作+可计算”。例如，影像组学特征以及基因组变异数据仅仅并列显示时，仍然需要人工观察它们之间的联系；但融合之后的知识协作可以通过关联规则发现“这种影像学表现，在具有相似基因型的人群中对于免疫治疗的效果”。所以，数据融合的意义是产生新的知识，即将不同类型的数据结合起来之后得到单个数据所不能提供的有价值的医疗信息。这也意味着融合系统必须有知识提取以及推理的能力而不是简单的数据传输。

2 医疗信息数据融合的关键技术瓶颈与突破方向

2.1 多源异构数据的语义异构与互操作障碍

医疗数据异构性体现在语法层面（格式、编码、传输协议）以及语义层面（术语含义、上下文解释）。语法上的异构可以通过一些标准接口来解决一部分问题（例如HL7 FHIR, DICOM等），但是语义上仍然有很大的困扰——比如“肿瘤大小”，在影像学检查结果中是长径，在病理检查报告中是镜下浸润程度，由于它们所表示的意义不同而无法进行比较。解决方法就是制定一个统一的医学本体（例如SNOMED CT, LOINC等）并用语义网络的方式把各个数据元素之间的关系连接起来。近几年以图数据库为基础的语义映射工具逐渐发展成熟，可以自动发现大量同义词的不同表述方式，准确率可以达到90%以上，在中文环境中还有很大的改进空间。

2.2 数据质量非标准化对融合可靠性的制约

融合质量取决于源数据质量，在MDT所使用数据来源中，普遍存在一些问题，例如：重要信息缺失（如病理报告中无分子分型）、时间错误（如检验时间仅到天）、不同单位混用（如肌酐有 $\mu\text{mol/L}$ 和 mg/dL 两种表示形式）、检查序列参数不一致等。这些问题在一个系统内可能不会造成较大影响，但是融合之后互相比对则会加大误差，甚至带来错误的相关性^[2]。因此有必要制定MDT的数据质量准入制度，在数据进入融合池之前进行完整度、准确性以及及时性的检测，不符合要求的数据返回去重新填写或者打上可靠程度标签。

2.3 隐私保护与数据可用性的平衡难题

MDT融合应用需要大量临床数据支持，但是患者隐私保护法律法规（例如《个人信息保护法》以及医疗数据安全相关规定）对于数据可利用范围、去标识化手段和授权机制均有限制。传统的脱敏方法（如删除标识符）会使得数据之间联系被破坏从而降低融合效果。差分隐私、k-匿名等方法可以在一定程度上保护个人隐私同时保留统计数据，但是MDT个性化诊疗需要准确的数据，过多地保护隐私和临床的实际需求相悖。而联邦学习提供了一个折中的方案——只交换模型参数而不是原始数据，在多方协作下完成整个模型训练过程，在保证各个参与方病人隐私的前提下完成跨院MDT数据融合。

3 面向MDT的医疗信息数据融合应用架构设计

3.1 四层融合架构的逻辑划分与功能定位

本文提出面向MDT的数据融合应用架构，从下至上共四个层次。最底层是数据汇聚层，在各个业务系统中利用ETL（提取-转换-加载）的方式获取原始数据，并以患者主索引（EMPI）为基础进行时间和空间上的对齐后

形成标准的数据湖。中间一层是语义对齐层，利用医学自然语言处理的方法对自由文本报告（例如影像所见、病程记录等）进行结构化抽取并映射到一个共同的术语集中，同时也建立了影像-病理-基因组等多种类型的数据之间的语义桥梁。再往上是一层的知识表示层，搭建了MDT专病的知识图谱，将患者的个人信息以及疾病的数据库、指南规则以及治疗方法相结合得到多种特征组成的向量。最上面一层是决策服务层，提供了个性化的MDT视图、可视化的分析工具、智能化的警报和建议等功能，用“仪表盘”的形式展示患者的全貌以及建议作出决定。

3.2 数据中台作为融合枢纽的技术实现路径

数据中台是四层架构最底层的核心载体，主要包含以下技术组件：①统一数据模型，用FHIR资源模型将各种临床数据统一标准化，使不同系统之间具有相同的语义；②主数据管理，维护一份可靠的病人信息以及就诊事件列表，在多个系统中解决病人ID重复问题；③时序数据库，快速地读取和查询以时间为顺序的数据，如检查结果、监控数据等，便于时间轴滚动比较；④特征工程流水线，对影像组学、基因组学等高维数据进行降维以及特征提取，得到可以被后续使用数值向量^[3]。数据中台是基于微服务体系结构，每个模块都可以单独伸缩并且可以通过公开API（应用程序接口）为MDT前端应用提供数据服务，同时也能够方便地接入外部分析工具。

3.3 知识图谱驱动的跨模态语义融合机制

传统融合只是做到“数据同框”，而知识图谱可以做到“语义互联”。比如肿瘤MDT，建立一个包括“疾病-基因-药物-影像征象-病理类型-预后因子”等实体以及它们之间的关系的知识图谱，然后把患者的个案数据作为实例填充到这个图谱中，形成一个“患者-表现-诊断-治疗”的实例化的子图。当MDT成员需要查询某个指标时，系统不仅给出这个指标的结果，还会利用图谱进行推导，给出这个指标之前的因素以及之后的影响。如查询“PD-L1表达”，同时给出该病理标本的质量、检测手段、以及和之前免疫治疗的关系等信息。这种方式已经远远超出了单纯的结构化查询，是基于语义的认知辅助，大大减少了跨专业的理解难度。

4 融合应用的关键实施维度与治理策略

4.1 标准化体系建设：术语、数据元与交换协议

标准化是融合的基础，需从三方面入手，术语标准化方面，在线使用国家发布的医学名词标准并制作本地术语映射关系，将不符合规范命名转换为符合要求名称；数据元标准化方面，依据国家卫生健康委发布的电

子病历数据元集,制定MDT核心数据集,规定每个数据元含义、数据类型、取值范围以及单位等;交换协议标准化方面,全部采用FHIR R4标准进行系统之间信息传递,摒弃原有自定义接口,减少不同系统之间连接难度^[4]。标准化工作是一项长期任务,应成立专门的数据标准化工作组,不定期检查及修订对应关系。

4.2 数据治理框架:质量闭环与责任划分

数据治理贯穿整个数据产生到使用的过程,在此过程中实行“谁生产、谁负责”的原则,各个源头部门对自己所生产的数据负责;信息中心进行技术上的监控以及报警;MDT数据管理员进行综合的数据检查以及问题追查。而治理工作是实行PDCA模式:每天自动产生一个数据质量报表,标注出缺失、超出范围以及不符合要求的数据;每周开一次质量提升会,找出问题所在并且提出解决方案;每个月公布一次融合数据可用性指标(例如完整性、准确性、及时性),作为科室考核的一部分。同时还需要有数据血缘关系管理系统,保存每一条融合数据的数据源、处理方式以及变化情况等信息,以便回溯问题或者重现结果。

4.3 组织文化与人员能力建设

技术架构的应用需要有组织上的配合。在MDT的数据融合应用中需要临床医生、信息工程师以及数据分析师之间有良好的合作氛围。一方面,医生应参与到融合的需求提出以及最终的验收工作中来,使生成的结果符合临床上的习惯;另一方面,信息人员需要了解一些基本的医学知识,理解临床的专业术语以及治疗的过程等。可以设置一个“临床信息沟通官”的角色,作为连接临床和信息技术的纽带,进行需求的翻译、相关的培训以及反馈的收集等工作。此外,在MDT会议上增加对

数据的解读部分,在信息人员或者是数据分析师简要介绍数据融合的结果可信度以及局限性之后,提高大家对于基于数据做决策的认可程度。

5 结语

多学科诊疗模式广泛开展对医疗信息数据融合应用也提出从“辅助工具”到“核心基础设施”的要求,在分析MDT对数据融合的需求后总结出语义异构、数据质量以及隐私保护三方面问题并给出了一种包括汇聚、对齐、表示和服务四层次的方法论,并且从标准、管理和文化三方面阐述了解决方法。我们认为,医疗信息数据融合是技术、管理和人三位一体的变化过程,其成功不是单纯依靠技术创新和技术平台搭建,还需要改变现有的工作方式以及建立良好的数据治理机制才能实现。未来,随着联邦学习、可解释人工智能和数字孪生等新技术的发展,将会使数据融合的应用更加深入广泛,但是无论如何都离不开一个中心思想:融合是为了让MDT团队看得更清楚、判得更准确、决策更好而不是为了技术而技术。

参考文献

- [1]中华医学会医学信息学分会.中国医疗健康信息数据融合与互操作白皮书.中华医学信息导报,2024,39(5):3-10.
- [2]李建中,刘杰,王浩然.基于语义网的异构医疗信息融合技术研究进展.计算机学报,2023,46(7):1423-1441.
- [3]张鑫,陈晓峰,赵立强.医疗数据中台架构在临床决策支持系统中的应用.中国数字医学,2024,19(3):45-51.
- [4]周晓华,马宁,孙德伟.多学科诊疗信息化平台建设指南及实践探索.中国医院管理,2023,43(11):56-60.