

重症肺炎 (severe pneumonia, SP) 是由肺组织(细支气管、肺泡、间质等)炎症发展到一定疾病阶段恶化加重形成, 常引起器官功能障碍甚至危及生命^[1], 导致严重的医疗负担和经济负担, 提高 SP 的诊治水平迫在眉睫。当前随着大数据、人工智能算法迅速发展, SP 领域正处于快速发展的关键时期, 深刻认识数据管理和样本库构建的重要性, 对于提高 SP 治愈率、实现 SP 全程管控具有重要作用; 同时, 由于国内外重症肺炎病原谱差异大^[2]、检测技术水平参差不一、多种技术平行检测率低、规范化数据样本库欠缺等原因, 我国 SP 领域亟待完善数据管理体系及样本库搭建体系。基于此, 本文拟通过梳理生物样本库构建流程, 以及列举评析其在呼吸感染精准医学研究中的基础支撑案例, 进而为国内 SP 数据管理及样本库搭建提供一定的参考与启示。

1 SP 数据库及生物样本库标准化平台

1.1 SP 数据库及生物样本库概述 医学标准化、高质量的生物样本库是重大疾病基础与临床研究、临床诊治技术研发、药物研发、健康研究与产业化, 即实现转化医学的最宝贵资源和最重要环节之一, 毫无疑问也是当今生命科学原创性研究、生物医药产业自主创新体系中至关重要的环节与保证^[3]。在医学领域, 问题的科学性探究往往离不开第一手临床数据, 而后续样本的处理和分析同样重要。生物样本数据库在临床诊疗到科学研究所这一过程中发挥关键作用, 即医学科学的发展需要一种能提供疾病相关的、具有普遍意义的、全局性的数据库^[4]。

SP 专病队列数据库信息平台通过接入医院系统, 可精准实现自动识别患者、添加分组和制定随访计划, 并邀请患者加入, 进行随访。同时可汇总院外随访数据, 实现动态匹配调整随访计划, 在平台形成样本临床信息库。进一步通过整合生物样本库数据, 将专病数据库与生物样本库对接整合、统一管理和检测, 实现重症肺炎专病队列数据库及生物样本库的“二库合一”。

1.2 SP 数据库及样本库标准操作流程

在 SP 样本库构建中, 标准化操作程序 (standard operating

【基金项目】国家四大慢病重大专项(编号:2023ZD0506200/2023ZD0506201)

【通讯作者简介】周敏,女,主任医师,教授,博士生导师。中华医学会呼吸病学分会委员、呼吸治疗学组副组长,中国医师协会呼吸分会政策委员会副主任委员,中华预防医学会呼吸专委会委员,上海预防医学会呼吸专委会副主任委员,上海医学会呼吸病分会委员、肺功能学组组长,上海女医师协会肺癌专委会副主任委员。主要研究方向:肺部感染、慢性阻塞性肺疾病。

procedure, SOP) 是确保工作高效、准确执行的基石。现就主要数据来源及其整合机制予以评述。

1.2.1 信息库 信息数据库即干库(data repository), 其整合了患者在医院的数据, 包括临床特征、病理学、影像学等多个维度的临床数据, 同时包含样本分析后的组学数据, 是从多模态数据到结构化数据一站式信息检索集成的表型数据体系。应在确保规范性、完整性、准确性基础上, 由各单位联合管理, 对临床表型数据、组学数据进行检验、质控、储备及整合。

1.2.2 实体库 实体样本库即湿库(biospecimen repository), 要在确保安全性、可靠性、可溯源性的同时, 实现生物样本采集、转运、处理、储存与应用的全生命周期监管体系。根据重症肺炎特点可初步确认样本库收集和应用的生物样本类型, 例如需全面收集痰液、支气管肺泡灌洗液、血清、胸腔积液、肺组织活检标本、尿液等标本, 样本采集和保藏时应遵循相应标准操作规程^[5], 并针对不同标本的特性分类管理。现将肺炎主要收集样本进行简述:①呼吸道标本: 呼吸道感染分为上呼吸道感染和下呼吸道感染。不同部位的感染病原菌差异较大, 上呼吸道感染多以病毒或定植菌为主, 下呼吸道感染病原菌多样, 因为标本易受口咽部菌群的污染, 选择合适的标本尤为重要。②深部痰标本: 痰标本仅用于下呼吸道感染, 主要是肺部感染的诊断, 但其不是诊断肺部感染的最佳标本。肺泡灌洗液或经气管吸取物等相对无菌的标本更为合适。但目前有研究表明深部痰标本取样具有较高灵敏度, 这可能与病毒载量、疾病严重程度有关^[6]; 同时, 研究显示其在儿童难治性肺炎病因诊断中可作为病因标本^[7]。因此, 其在重症肺炎应用过程中取样、储存、实验室鉴定等流程的标准化步骤仍需进一步探讨。③肺泡灌洗液标本: 采集肺泡灌洗液进行检测, 可减少口咽部菌群的污染, 提高检测结果的准确性。采集需注意患者和样本的情况。如支气管肺泡灌洗时支气管镜顶端要紧密嵌顿于段或亚段支气管开口, 以防止大气道分泌物混入或灌洗液外溢。后续再根据需要选择细胞学分析或微生物检测等分析相关的处理方式进行保存。④鼻咽拭子标本: 鼻咽拭子标本是用于检测呼吸道病原体(尤其是病毒, 如新型冠状病毒、流感病毒等)的常用临床样本, 例如使用鼻咽拭子专用采集管、拭子采集后短期内送检。

1.3 样本库质量控制体系 样本库质量控制包括采集标准、标准化管理、样本储存、设备维护以及成立专门的质量控制管理委员会等方面的内容。国

际上关于生物样本库建设与管理的政策文件和指南为生物样本的合规性、伦理性和质量管理提供了统一标准,这些国际指南共同构建了生物样本库管理的全球标准,确保样本在科学研究中的合法使用和安全保障。比如应参照国际规范标准,如国际生物与环境样本库协会的《生物样本库最佳实践》^[8]等建立重症肺炎生物样本采集的标准操作流程及质量控制体系。

1.4 数据库及样本库伦理规范 生物样本数据库建设全方位的生物信息隐私保护和多权限层次的样本信息访问系统极为重要。要对生物信息进行严格的保护,制定严格的授权程序,控制不同用户对数据的查询、录入、更改权限,以建立规范化的严格权限管理下的样本资源共享平台。同时,需进一步建立完善的医学科研伦理、医疗和科技安全审查机制,防范伦理和安全风险。例如按照有关法律法规和伦理准则,建立健全医学科研伦理、医疗和科技安全管理制度;加强伦理审查和过程监管,加强生物安全、信息安全等医疗和科技安全责任制等。

2 数据库推动 SP 精准研究的进展

生物样本及数据库可助力精准医疗的发展,主要是利用个体特征证据(如详细的临床数据、电子健康记录等)、生物分子学证据(如基因组学、转录组学、蛋白组学、代谢组学等数据)等内容,基于可靠证据证明特定标志物可有效预测或影响疾病发生发展。

目前对于 SP 领域的精准治疗研究,主要集中在发病机制、治疗靶点、个体化治疗方案等方面。通过重症肺炎队列进行病原体快速检测及耐药机制、宿主免疫与病原体互作、呼吸道微生态与感染研究;通过药物药动学-药物学分析合理使用抗生素等研究,能够帮助提高 SP 的病原体诊断水平、突破细菌耐药困境、寻找重症肺炎的干预靶点、提高对呼吸道急性传染病的预警能力。Schons 等指出 SP 大队列应包括全面的临床和影像诊断、生活质量评估、患者报告的结果和生物采样,以建立一个关于大队列人群呼吸道感染的标准化、高质量的数据和生物样本收集^[9]。这一观点即生物医学数据库驱动转化研究突破,目前国内外研究主要有以下几类:

2.1 疾病遗传架构解析 SP 临床表型差异很大,深入分析宿主遗传因素对 SP 的影响对于提高潜在疾病机制的认识非常重要。遗传信息与疾病发生、发展机制相关,例如 Degenhardt 等利用 3255 例重症患者与 1.2 万对照组的数据进行全基因组关联荟萃分析,发现 17q21.31 倒位多态性在 SP 的影响,并定位新易感位点 19q13.33^[10];通过对 1220 例 SARS-

CoV-2 感染者的全基因组测序,发现风险点位关联对疾病的影响^[11];同时,一项前瞻性研究发现 SARS-CoV-2 感染的急性期,人类白细胞抗原等位基因不会显著影响表型结果^[12]。通过对整个频谱中不同类型的变异进行联合分析,这将继续有助于阐明疾病病因。此外,遗传信息往往通过中心法则等途径导致诸多临床问题,例如细菌耐药性^[13]等。在铜绿假单胞菌方面,国内学者通过全基因组测序等技术利用多个队列进行分析,实现对铜绿假单胞菌碳青霉烯耐药性的直接预测,有助于特定耐药菌种的及时治疗和监测^[14]。

2.2 多组学揭示机制及新型标志物 干库、湿库可获取多种组学数据,以深入探究 SP 病理机制、挖掘新型生物标志物。在疾病发生、发展机理方面,干、湿库具有重要作用,例如对新冠肺炎相关肺曲霉病(COVID-19 associated pulmonary aspergillosis, CAPA)支气管肺泡冲洗液样本进行测序,发现单核细胞、巨噬细胞、B 细胞和 T 细胞紊乱以及中性粒细胞外陷阱在发病中发挥作用并且具有诊断价值^[15]。儿童多系统炎症综合征患者血液转录组分析发现自然杀伤细胞和 CD8⁺T 细胞衰竭与疾病发生发展机制相关^[16]。生物样本库在表观遗传研究层面同样具有重要作用,例如 Biñkowski 等利用甲基化数据、转录组学数据发现 SARS-CoV-2 感染的病理生理学中涉及的 DNA 甲基化变化是新冠患者所特有的,可作为疾病管理的生物标志物和潜在的治疗目标^[17]。此外,蛋白质组学、表型组学以及多组学整合分析同样是全面理解疾病复杂表型不可或缺的重要组成部分,例如利用多组学研究队列,系统描述新冠病毒灭活疫苗接种对奥密克戎感染免疫反应的影响,揭示疫苗在抗病毒作用中的潜在分子机制^[18]。通过对 986 例新冠患者血浆蛋白进行分析,筛选出含细胞因子通路与非免疫通路蛋白的蛋白组合以预测重症新冠肺炎严重程度^[19]。

2.3 临床表型与疾病模型搭建 随着分子时代和信息技术时代的迅速发展,精准医疗对实现个体化循证医学具有关键作用,通过科学分析临床表型等多种数据,可促进疾病个体化分子分型、生物标志物探索,推动精准医疗进一步发展。例如在严重社区获得性肺炎(severe community-acquired pneumonia, SCAP)领域,患者由于微生物、宿主反应等多种因素影响,临床治疗和预后结局往往差异性较大^[20]。国内利用多中心前瞻性研究队列,联合宏基因组、转录组筛选,鉴定出关键临床、组学特征,并开发 SCAP 综合预测模型,以达到推动改善 SCAP 治疗策略和预后的目的^[21]。在新冠病毒研究方面,数据库

在临床方面同样发挥重要作用。例如在密歇根基因组学倡议、英国生物库和美国国立卫生研究院 All of Us 生物库队列进行全表型组关联研究 (phenome-wide association studies, PheWAS)，发现多种临床表型是新冠肺炎重症的独立遗传风险因素^[22]；Willette 等利用新冠患者基线特征发现抗体滴度 (如巨细胞病毒血清学) 联合血脂指标具有良好预测效能，提出基于公共卫生和医疗机构标准的自我报告和生物医学数据，可以建立准确的风险概况^[23]。此外，在医院获得性肺炎 (hospital-acquired pneumonia, HAP) 方面，有学者利用 HAP 队列分析念珠菌属与 HAP 的关系，发现下呼吸道念珠菌属与死亡率、肺部基础疾病相关^[24]。

3 小结

当前 SP 生物数据库正逐渐得到重视、成为研究热点。但目前针对中国本土 SP 生物样本数据库建设尚不足，存在平行检测缺失、数据规范化程度不足以及数据库样本存在异质性等问题，限制了临床科学问题研究的深入推进，亟需在后续工作中重点解决。

【参考文献】

- [1] 中国急诊重症肺炎临床实践专家共识 [J]. 中国急救医学, 2016, 36(2): 97-107.
- [2] Martin II, Torres A, Nagavci B, et al. ERS/ESICM/ESCMID/ALAT guidelines for the management of severe community-acquired pneumonia [J]. Intensive Care Med, 2023, 49(6): 615-632.
- [3] 中国医药生物技术协会生物样本库标准(试行) [J]. 中国医药生物技术, 2011, 6(1): 71-79.
- [4] 董尔丹, 胡海, 俞文华. 生物样本库是生物医学研究的重要基础 [J]. 中国科学: 生命科学, 2015, 45(4): 359-370.
- [5] 丁永杰, 罗文婷, 李庆云, 等. 呼吸系统疾病生物样本采集保藏管理技术标准和规范 [J]. 中国医药导报, 2021, 18(12): 4-10.
- [6] Khiabani K, Amirzade-iranaq MH. Are saliva and deep throat sputum as reliable as common respiratory specimens for SARS-CoV-2 detection? A systematic review and meta-analysis [J]. Am J Infect Control, 2021, 49(9): 1165-1176.
- [7] Wang YN, Wu YT, Cao L, et al. Application of metagenomic next-generation sequencing in the etiological diagnosis of refractory pneumonia in children [J]. Front Microbiol, 2024, 15: 1357372.
- [8] Snapes E, Astrin JJ, Bertheussen KN, et al. Updating international society for biological and environmental repositories best practices, fifth edition: a new process for relevance in an evolving landscape [J]. Biopreserv Biobank, 2023, 21(6): 537-546.
- [9] Schons M, Pilgram L, Reese JP, et al. The German National Pandemic Cohort Network (NAPKON): rationale, study design and baseline characteristics [J]. Eur J Epidemiol, 2022, 37(8): 849-870.
- [10] Degenhardt F, Ellinghaus D, Juzenas S, et al. Detailed stratified GWAS analysis for severe COVID-19 in four European populations [J]. Hum Mol Genet, 2022, 31(23): 3945-3966.
- [11] Schmidt A, Casadei N, Brand F, et al. Systematic assessment of COVID-19 host genetics using whole genome sequencing data [J]. PLoS Pathog, 2024, 20(12): e1012786.
- [12] Marchal A, Cirulli ET, Neveux I, et al. Lack of association between classical HLA genes and asymptomatic SARS-CoV-2 infection [J]. HGG Adv, 2024, 5(3): 100300.
- [13] GBD 2021 Lower Respiratory Infections and Antimicrobial Resistance Collaborators. Global, regional, and national incidence and mortality burden of non-COVID-19 lower respiratory infections and aetiologies, 1990-2021: a systematic analysis from the Global Burden of Disease Study 2021 [J]. Lancet Infect Dis, 2024, 24(9): 974-1002.
- [14] Liu B, Gao J, Liu XF, et al. Direct prediction of carbapenem resistance in *Pseudomonas aeruginosa* by whole genome sequencing and metagenomic sequencing [J]. J Clin Microbiol, 2023, 61(11): e0061723.
- [15] Feys S, Vanmassenhove S, Kraisin S, et al. Lower respiratory tract single-cell RNA sequencing and neutrophil extracellular trap profiling of COVID-19-associated pulmonary aspergillosis: a single centre, retrospective, observational study [J]. Lancet Microbe, 2024, 5(3): e247-e260.
- [16] Beckmann ND, Comella PH, Cheng E, et al. Downregulation of exhausted cytotoxic T cells in gene expression networks of multisystem inflammatory syndrome in children [J]. Nat Commun, 2021, 12(1): 4854.
- [17] Bińkowski J, Taryma LO, Łuczkowsk AK, et al. Epigenetic activation of antiviral sensors and effectors of interferon response pathways during SARS-CoV-2 infection [J]. Biomed Pharmacother, 2022, 153: 113396.
- [18] Yu S, Lin Y, Li Y, et al. Systemic immune profiling of Omicron-infected subjects inoculated with different doses of inactivated virus vaccine [J]. Cell, 2023, 186(21): 4615-4631.
- [19] Su CY, Zhou S, Gonzalez KE, et al. Circulating proteins to predict COVID-19 severity [J]. Sci Rep, 2023, 13(1): 6236.
- [20] Quah J, Jiang B, Tan PC, et al. Impact of microbial Aetiology on mortality in severe community-acquired pneumonia [J]. BMC Infect Dis, 2018, 18(1): 451.
- [21] Zhao J, He X, Min J, et al. A multicenter prospective study of comprehensive metagenomic and transcriptomic signatures for predicting outcomes of patients with severe community-acquired pneumonia [J]. EBio Medicine, 2023, 96: 104790.
- [22] Fritzsche LG, Nam K, Du J, et al. Uncovering associations between pre-existing conditions and COVID-19 Severity: A polygenic risk score approach across three large biobanks [J]. PLoS Genet, 2023, 19(12): e1010907.
- [23] Willette AA, Willette SA, Wang Q, et al. Using machine learning to predict COVID-19 infection and severity risk among 4510 aged adults: a UK Biobank cohort study [J]. Sci Rep, 2022, 12(1): 7736.
- [24] Han Y, Zuo Y, Luo Z, et al. Positive sputum culture of *Candida* spp. as a risk factor for 30-day mortality in patients with hospital-acquired pneumonia: a propensity-score matched retrospective clinical study [J]. Chin Med J Pulm Crit Care Med, 2023, 1(2): 113-118.

(收稿日期:2025-09-10;修回日期:2025-09-15)

(本文编辑:彭羽)