

“2025 年创新专病库”

参赛类型： ☐ A 类（医学中心主导）

☒ B 类（科室主导）

申报部门： 儿科

项目名称： 基于多模态数据中心的过敏性疾病专病库建设及科
研和临床转化平台研发

负 责 人： _____ 联系方式： _____

牵头部门： 儿科

联合部门 耳鼻喉科、信息科、皮肤科

电子邮箱： _____

申报日期： _____

研究期限： _____

2025 年 4 月

1. 立项依据

1.1 过敏性疾病的现状、挑战及科学意义：

过敏性疾病概述与流行病学现状：近年来，我国过敏性疾病的患病率呈快速上升趋势，逐渐接近甚至超过西方发达国家。

挑战及科学意义：环境、遗传、肠道微生物之间的交互影响尚有很大的研究空间。

过敏性疾病往往不是单独存在，而是多种过敏性疾病同时出现在同一患儿身上，Ziyab[15]一项涵盖 1154 名，年龄范围在 18-26 岁的大学生过敏性疾病横断面研究显示，AA 和 AR 并存是最为常见的过敏性多发疾病，患病率为 5.1%，其次是 AR 和 AD，患病率为 1.5%，AA、AR 和 AD 三病共存的整体在整个研究样本中占比 0.6%；两种或多种（ ≥ 2 ）过敏性疾病的共患率为 7.9%。生命早期罹患 AD 可能增加过敏性疾病共病的风险。

Dharmage 等[16]研究发现，在 2 岁以内患有 AD 的婴儿中，在 6~7 岁期间与 AA 和 AR 的共病率显著增加，尤其是早发、持续性、IgE 阳性的 AD 患儿共患两病的风险更高。一项针对 2311 名加拿大儿童的出生队列研究表明[17]，1 岁时患儿 AD 的儿童成长至 3 岁同时罹患 AA 和 AR 的风险较 1 岁时无 AD 儿童分别增加 11 倍和 7 倍，这种不同过敏性疾病在时间长轴上表现出的一定前后顺序的现象也被称为特应性进程或过敏进程

（atopic march）。这种多病并存共发的现象往往导致患儿多脏器和系统受累，多病相互影响、相互加重、相互发展，严重影响患者的生活质量，增加医疗费用，延长了治疗周期，整体影响疾病控制，对儿童的成长发育产生长远影响，甚至到成人过敏性疾病仍伴随，如支气管哮喘可发展为哮喘慢阻肺综合征，进一步影响成人的心肺功能，造成活动耐力下降，很多工作不能胜任，部分患者还会出现严重的心理健康问题。六岁以内哮喘儿童若能得到早期实变，规范抗哮喘治疗，约有 70-80% 年长后症状不再反复，30-60% 的患儿可完全控制或自愈[18]，而成人阶段若仍有哮喘发生，则需终身用药抗哮喘治疗。故在生命早期儿童免疫功能尚未完全建立阶段，是有机会逆转过敏性疾病的自然进程的。

变应性疾病（哮喘、AR 和早期的特应性皮炎）与特定过敏原的 IgE 以及非过敏机制相关，这些机制可能共存，患者也会同时或连续患有多种变应性疾病。单一致敏和多重致敏个体在临床和免疫学方面存在显著差异。20 多年前提出的“同一气道，同一疾病”概念是对上、下气道变应性疾病的简略描述，越来越多的证据表明，**单纯 AR 和哮喘合并 AR 是具有显著差异的两类疾病。**表现在①. 在临床观察中，30% 的 AR 患者合并哮喘，而大多数哮喘患者合并 AR。②. 在出生和儿童队列中，对不同过敏原的单致敏和多重致敏代表不同疾病的表现：与单致敏相比，多重致敏与更强的全身 IgE 反应、疾病表型（单纯 AR 与哮喘+AR）、症状和发展轨迹相关。多病共存在一定程度上独立于 IgE 致敏，提示其存在不同的因果（基因和机制）途径。IgE 多重致敏与多病共存（包括发病年龄、变应性多病共存数量、疾病严重程度、嗜酸粒细胞水平和总 IgE 水平）之间存在关联。③. 基因组研究差异：两种方法（转录组学和 RNA 测序）在两个不同的队列[过敏发展机制（MeDALL）和波多黎各儿童哮喘的表观遗传变（EVA-PR）]中得出了相同的结果：哮喘合并 AR 与 T2 信号通路的 7 个基因相关：IL-5（嗜酸粒细胞）和 IL-33（多重致敏和嗜酸粒细胞增多）。研究者鉴定出 27 个与单纯性鼻炎相关的基因，包括 Toll 样受体（TLR）和 IL-17。这些研究表明，单纯性鼻炎是由局部 IL-17 驱动的疾病，而 2 型炎症相关的 AR 合并哮喘是系统性 IL-33 驱动的疾病。虽然变应性多病共存具有共享的表观遗传模式，这些模式仅在哮喘合并 AR 的儿童患者（而非单纯哮喘）中发现。④. 治疗方案差异：多重变应性疾病患者由于控制效果差，经常使用鼻用糖皮质激素和口服抗组胺药联合治疗。在 MASK-air 中，联合用药模式比单一治疗更难以

控制 AR。在综合症状-药物评分中, AR 与哮喘合并 AR 两类疾病的差异显著。

这些研究加深了人们对两种变应性疾病的认识:单纯 AR(局部炎症,与 IL-17 和 TLR 相关)和哮喘合并 AR(系统炎症,与 IL-33 相关)是两个相互独立的疾病[19]。不仅如此,AR 伴哮喘有别于单纯 AR 这一观点,被变应性疾病发病机制研究队列工作小组系统阐述。在患者临床治疗和管理上,单纯 AR 可通过药物控制,预后良好;而 AR 伴哮喘预后复杂,关注上下呼吸道的炎症控制需要采取更加全面和个性化的方案,才能有效提高患者的生活质量并减少病情的恶化风险。了解过敏性疾病共病与多病的概念,有助于初步理解其流行特征及病理生理形成过程,对促进疾病早期发现、预防及治疗有重要意义。但是基于近几十年的研究报道来看,当前有关过敏性疾病共病与多病的相关发病机制尚未达成共识,仍需探索更详细、更合理的机制来解释其特征,同时寻找可靠的生物标志物来跟踪和预测长期预防和治疗措施的有效性。

过敏性疾病是需要长期规范化的健康教育、监测、治疗、随访的,而传统的疾病诊疗模式,患者对疾病认识不够充分,多次就医也会带来很多不便利,这样患者依从性差,就很容易造成患者失访,影响治疗效果,甚至因长期用药导致一些不良反应而不自知,或者擅自停药等问题存在。我国儿科几乎没有专门的过敏免疫或变态反应专科医师队伍。相关知识的普及率就更低了。因此,加强健康教育,普及过敏科普知识,提高人们对过敏性疾病的防控意识,调动个体、群体及全社会的积极性,探索持续有效又经济的预防控制策略及长期随访管理。对于我国这个发展中且过敏性疾病发病率不断上升的国家,具有更为重要的意义。

综上,过敏性疾病作为常见病,多发病,病人基数大,虽然有部分患者经过系统的抗炎、抗过敏等治疗,症状可以得到暂时缓解,但难以根治,很大一部分人会终身患各种过敏性疾病,严重影响人类的生活质量,严重的哮喘发作甚至导致猝死。目前科学研究方面还有许多需要深入探讨的内容,如环境与遗传学基因交互,出生后过敏性疾病自然病程的队列研究,肠道菌群对过敏性疾病的影响等等,而这些研究的基础是需要大量的临床数据,故建立过敏性疾病专病库具有很重要的临床和科研意义,在这方面的研究也能带来多方面实用的转化效益。

2. 临床问题及意义

主要临床问题: 1. 临床需求:过敏性疾病病因复杂(环境、遗传、免疫等多因素交互),个体化诊疗需求

迫切,但缺乏系统化的临床数据支持。由于我国大多数医院没有独立的过敏反应科,缺少过敏反应专科医生,过敏性疾病往往分散在耳鼻喉科、呼吸科、皮肤科、消化科、儿内科等就诊,各科室对过敏性疾病的诊断及治疗标准不统一,误诊、漏诊、处理不当或治疗过度等情况时有发生。由于过敏性疾病存在多病并存共发的现象,将散发在多个科室的过敏性疾病患者归口管理形成多学科诊疗显得至关重要。

2. 科研瓶颈: 现有研究数据分散,缺乏标准化、多维度、长期随访的专病数据库,制约病因探索、新药研发和精准治疗。随着医院信息化建设的不断深入,大量重要的临床数据都积累在医院信息化系统中,如何利用这些数据服务于循证医学和临床科研是各级医疗机构面临的挑战和重要的研究方向。医疗信息化的传统建设模式是以医院信息系统(hospital information system, HIS)为核心,以科室为单位构建信息系统,其缺乏统一的数据采集标准,也没有相应的数据存储规范或指南,导致大部分数据以非结构化方式呈现,数据质量低下,临床资料搜集和整理工作仍主要依靠人工,效率不高,数据难以再利用,不能在

临床和科研工作中发挥应有的作用。

意义：多学科诊疗如何应用于疑难疾病的诊断与治疗已然成为近年来医学界讨论的热点之一，将多学科诊疗应用于过敏性疾病的诊疗，对于提升医院整体临床诊疗水平、改善患儿生活质量具有重要意义。因此，在信息化、互联网、人工智能等新技术的帮助下，多学科诊疗这一新型临床诊疗模式便成为提升过敏性疾病诊疗水平的重要突破点之一。

3. 本项目的特色与创新之处

- ①. 多维度数据融合：首次整合临床、环境、肠道菌群、基因组等数据，构建过敏性疾病全生命周期数据链。
- ②. 标准化与可扩展性：采用国际通用标准，支持未来向其他疾病数据库扩展。
- ③. 推动数据向新药研发、健康管理产品转化。

三、建设方案

1. 研究内容

1.1 建设覆盖临床诊疗、生物样本库、等多维度数据的过敏性疾病专病库，实现数据整合、共享与深度挖掘。

1.1.1 临床数据库：前期先进行既往数据筛选，后期可进行临床验证数据库的可行性，并逐步纳入更多过敏性疾病数据。筛选我院院 2020 年 1 月—2024 年 10 月的病历数据，纳入标准为临床诊断为哮喘（喘息）、过敏性鼻炎、湿疹或特应性皮炎、食物过敏的门诊及住院患者。数据来源包括医院的电子病历（electronic medical record, EMR）、HIS、实验室信息管理系统（laboratory information management system, LIMS）、影像存档与通信系统（picture archiving and communication system, PACS）、护理系统等。将收集到的数据通过提取-转化-加载技术对医院各个信息管理系统中的异构数据进行预处理，完成数据汇集。然后再经过映射、清洗等过程，对数据进行再审核、再校验，以提高数据的准确性和一致性。通过自然语言处理技术，对非结构化的数据进行结构化处理。最后，对患者相关信息、医生相关信息等隐私信息进行脱敏处理，完成专病数据库的数据处理。

1.1.2 生物样本库建设：在临床数据库基础上，建设生物样本库，包含患者血液、体液、组织标本，进行过敏性疾病相关基因组（易感基因如 IL-4、IL-13）测定、蛋白质组（如细胞因子谱）、代谢组（如组胺代谢物）、肠道菌群等的的数据，依据科研设计方案设计生物样本数据库 CRF 表格，逐步纳入患者数据。采用模块化电子 CRF 表格设计，涵盖样本采集（时间、部位、预处理方法）、存储条件（温度、冻存时长）、质量控制（RNA 完整性指数、DNA 浓度）及关联临床表型数据（疾病分期、过敏原特异性 IgE 水平）。并建立动态更新机制，为精准分型及靶向治疗研究提供资源支撑。

1.1.3 流调数据库：因过敏性疾病与生活环境关系密切，在该数据库 CRF 表格设计家庭环境监测、城市农村环境监测、城市农村疾病分布、家族史调查，并纳入其它与过敏相关的混杂因素（如饮食习惯、心理因素、维生素 D 摄入情况、疫苗接种、抗生素使用频率等）。具体包括：（1）环境监测数据：家庭环境（居住类型、通风条件、室内污染物如 PM2.5/VOCs、宠物接触、床品材质等）、城乡差异（GIS 定位的绿化率、空气污染指数、区域花粉浓度及气候特征）；（2）疾病分布数据：基于城乡分层的过敏性疾病发病率、严重程度及地理热图；（3）遗传与行为数据：详细家族史（三代内过敏及自身免疫病史的家系图谱）、饮食模式、营养状况（维生素 D 水平、日照时长）、医疗暴露（疫苗接种史、抗生素使用频率及生命早期用药记录）。

1.1.4 随访数据库：因过敏性疾病患者通常需要长期用药，除需密切监测药物不良反应外，还需评估治疗效果及时调整方案，故搭建过敏性疾病随访数据库，包括药物方案、药品不良反应监测、哮喘患者呼吸康复数据、疗效

评估问卷等，详见 CRF 表格专页。

1.2 标准化处理（详见设计依据内容 2.1.1）：

统一术语：遵循国际标准（如 SNOMED CT、LOINC）对过敏性疾病分类（如 ARIA 指南）。

结构化录入：设计标准化电子病历模板（CDR），减少自由文本依赖。

1.3 技术平台构建

数据存储架构：采用混合云架构，保障数据安全与共享效率。

使用 FHIR（Fast Healthcare Interoperability Resources）标准实现跨系统数据互通。

数据处理技术：自然语言处理（NLP）：从非结构化文本（如医生笔记）中提取关键信息。

数据脱敏：基于差分隐私或匿名化技术保护患者隐私。

1.4 数据治理与质量控制

质控流程：设立数据审核机制（如双录入校验、逻辑规则校验）。动态监测数据完整性（如缺失值报警）。

伦理合规：遵循 GDPR、HIPAA 等法规，建立伦理审查委员会（IRB）监督。

1.5 基于多模态数据的过敏性疾病临床研究与应用（详见 4.3 内容）：将利用数据库进行如下研究：环境-基因互作驱动的过敏性鼻炎风险分层与精准预防研究、粪菌移植（FMT）调控肠道菌群-免疫轴治疗儿童食物过敏的机制探索、全生命周期过敏自然病程的多模态队列研究、基于肺音 AI 分析的哮喘急性发作预警与分型管理、利用机器学习预测过敏原特异性免疫治疗疗效评估。

1.6 长期维护与迭代：

动态更新：设计自动化数据采集接口（如连接 IoT 设备实时获取环境数据）。

用户反馈优化：定期收集临床医生和研究者需求，迭代数据库功能（如新增数据字段或分析模块）。

2. 总体设计（以下框架可根据需要进行调整）

2.1 设计依据（包含标准、指南、共识、文章、临床经验等）

过敏性疾病专病库的设计需要基于多维度证据，整合国际指南、专家共识、研究文献及临床实践经验，以确保其科学性和实用性。以下是具体设计依据的详细分析：

2.1.1 国际/国内指南依据

2.1.1.1 过敏性疾病分类与数据标准化

ICD-11 编码：采用国际疾病分类标准，确保疾病名称、过敏原（如尘螨、花粉、食物）的术语统一化。

2012 年, 来自法国蒙彼利埃大学医院的研究者发起了建立过敏性疾病精准分类、统一术语和定义的修订行动, 包括世界过敏组织、美国过敏哮喘和免疫疾病学会、欧洲过敏与临床免疫学学会、亚太过敏哮喘和临床免疫学协会在内的多个学术组织均参与其中, 为过敏性疾病在新版国际疾病分类系统中实现全面精准表达提供科学分类建议、相关医学证据和临床支持[1]。经过努力, “变应性或过敏性疾病” 的详细分类作为独立一节, 成功被纳入了国际疾病分类第 11 次修订本死亡与疾病统计线性组合(International Classification of Diseases, Eleventh Revision for Mortality and Morbidity Statistics, ICD - 11 MMS) 中的第 4 章节免疫系统疾病中[2,3] (图 1), 该小节具体包括 9 个类目, 各个类目包含的疾病实体和术语数量详见表 1, 以 “累及呼吸道的变应性或过敏性疾患” 小节为例, 分类的层级关系详见图 2, 具体包含的疾病实体和相关医学术语数量详见表 2。

在过敏原分类方面, 基于逻辑观察标识符、名称及代码 (Logical Observation Identifiers Names and Codes, LONIC) 中关于过敏原检查项目的成熟分类, 从临床相关性、患病率、理解难易程度、ICD - 11 框架适用性角度遴选具体过敏原进入 ICD - 11 的扩展码章节[4]。在 2023 年 1 月发布的 ICD - 11 用于疾病与死因统计线性组合中, 过敏原分类在病因和物质的两个小节中均可以被索引到, 具体按照相对分子质量大小分为低分子量半抗原(含 100 个实体)、低分子量半光抗原(含 18 个实体)和高分子量过敏原(含 29 个实体), 如谷物和面粉作为过敏原时的分类编码为 XM3XH8[5]。

ICD - 11 是一套基于本体模型、具有多层级结构、能够支持数字化应用的分类系统[6]。其编码结构较 ICD - 10 进行了重大调整, 通过引入后组配方式建立一种由主干码、扩展码、特殊连接符号组成的结构化编码模型, 实现了多维度信息全球标准化采集路径, 更有利于计算机对医疗大数据的分析和解读[7]。ICD - 11 修订的另一个亮点是引入了单独的扩展码章节, 从而实现在不过度增加主干码数量的基础上, 大幅提升过敏性疾病相关信息采集的精度和丰度, 有利于获得过敏性疾病疾病谱数据和相关死因的全貌信息, 继而实现过敏性疾病数据在精细化层面的全球可比, 满足临床科研、流行病学监测、医疗质量与安全、医保支付、精细化管理等不同应用目的下对数据的需要[8]。

对 2023 年 1 月发布的 ICD - 11 MMS 版本中 “变应性与过敏性疾病” 一节中的分类编码情况进行分析, 具有后组配的主干码共计 235 个, 占 58. 3%, 且主干码后组配维度数量最多可达 4 个。后组配维度中涉及采集的信息点具体包括解剖部位、侧别、严重程度、严重过敏反应分级、病程、过敏原、职业因素、导致过敏的药物、昆虫种类、外因、发生机制等(图 3), 如图 4 中的案例 1, 对于变应性接触性皮炎, 通过 ICD - 11 可以记录皮炎发生的具体解剖部位、相关的过敏原及是否与职业因素有关等多种信息。此外, ICD - 11 中专门设置了 “伤害 - 原因 - 机制” 三组分编码模型, 以标准化的方式记录更多维度信息, 实现对医疗质量和患者安全的测量, 尤其可为药物引起的过敏性不良反应的识别和监测提供帮助, 为医政管理、医疗质量促进、医学研究和评价提供有力抓手[9]。例如对某种药物引起的药物性严重过敏反应的情况进行分类时, 可以采用编码簇的方式捕捉药物性严重过敏反应的有害效应、具体外因, 引起的药物种类、引起有害效应的机制, 详见图 4 中的案例 2。

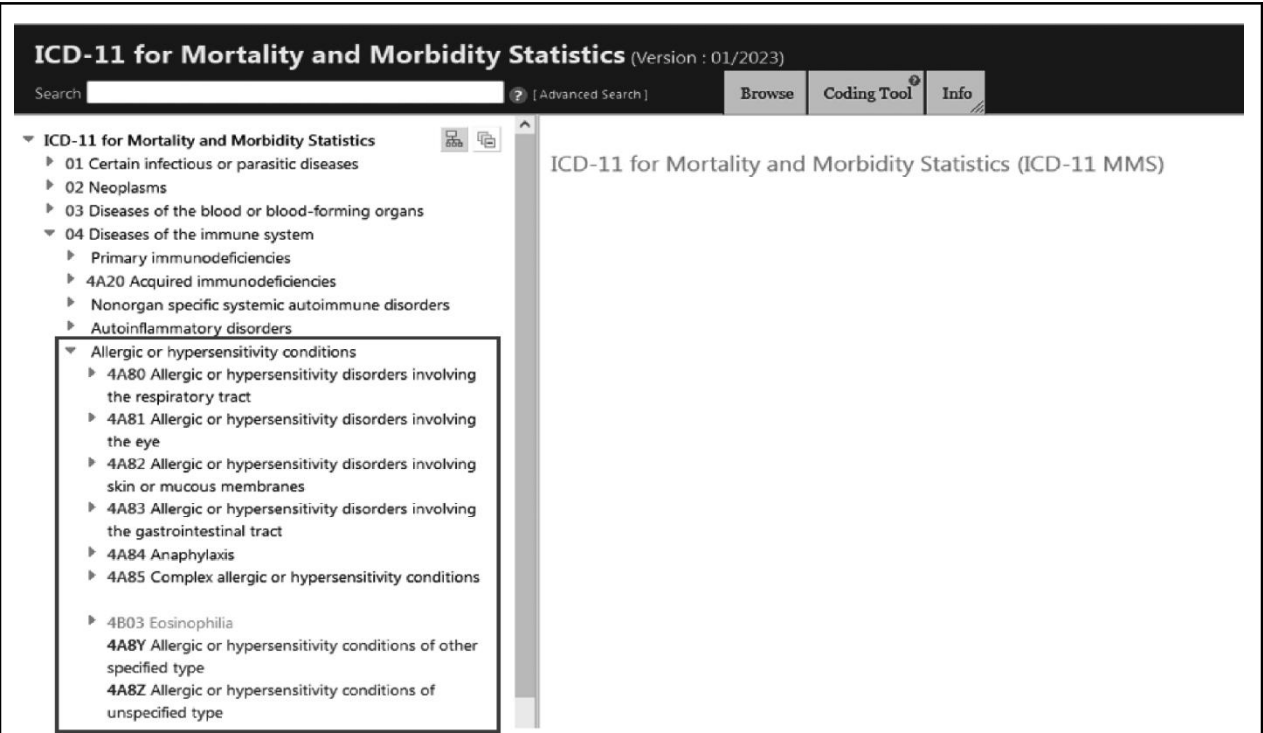


图 1 ICD - 11 用于疾病与死因统计线性组合中“变应性或过敏性疾病”的分类图示[22]

表 1 ICD - 11 第 4 章“变应性或过敏性疾病”各类目包含实体和术语的数量分布

类目名称	类目编码	实体数	术语数
累及呼吸道的变应性或过敏性疾病	4A80	175	543
累及眼的变应性或过敏性疾病	4A81	7	35
累及皮肤或黏膜的变应性或过敏性疾病	4A82	215	394
累及消化道的变应性或过敏性疾病	4A83	35	65
严重过敏反应	4A84	20	64
复合变应性或过敏症	4A85	97	189
嗜酸性粒细胞增多症	4B03	14	16
其他特指类型的变应性或过敏症	4A8Y	1	1
未特指的变应性或过敏症	4A8Z	3	12

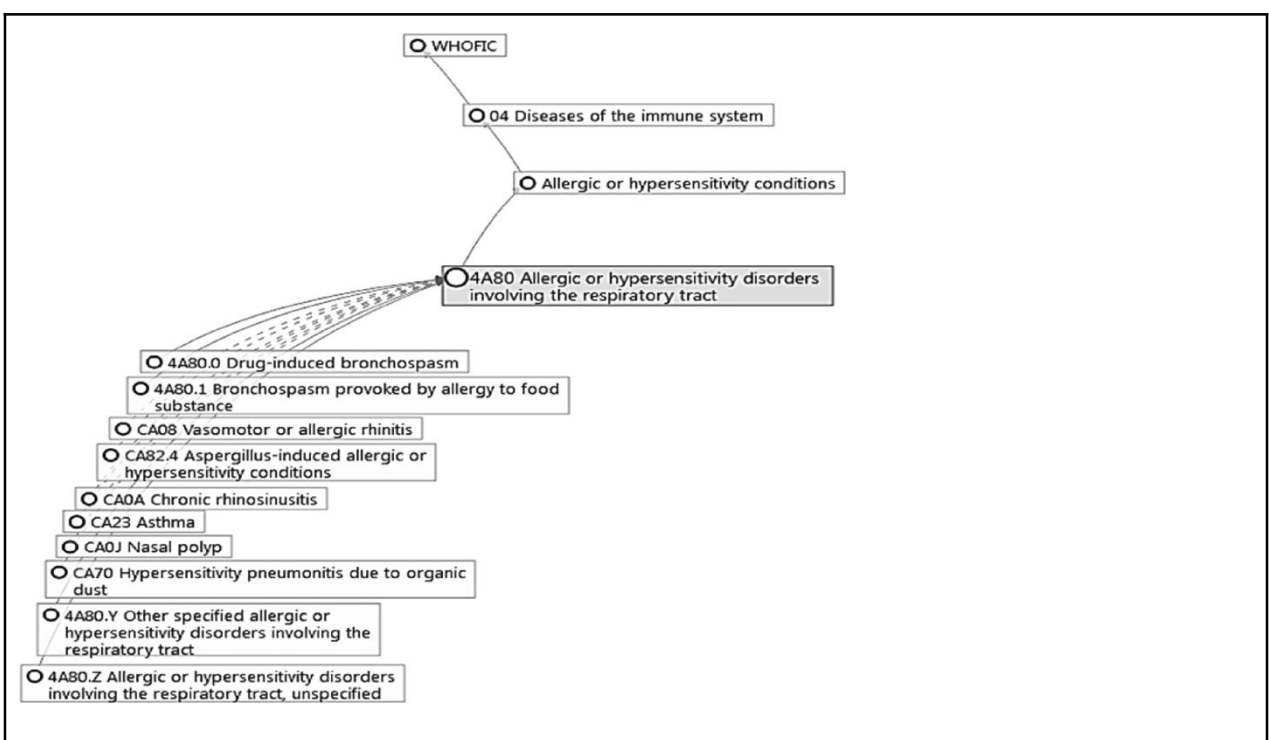


图 2 ICD - 11 中“累及呼吸道的变应性或过敏性疾病”小节的分类层级展示

表 2 ICD - 11 中的“累及呼吸道的变应性或过敏性疾患”各分类下实体和相关术语数量分布

下级分类名称	ICD-11 编码	实体数	术语数
药物诱导的支气管痉挛	4A80. 0	1	3
食物过敏诱发的支气管痉挛	4A80. 1	1	6
血管收缩性或变应性鼻炎	CA08	43	91
曲霉菌所致的变应性或过敏性状态	CA82. 4	2	4
慢性鼻窦炎	CA0A	47	151
哮喘	CA23	32	139
鼻息肉	CA0J	18	28
有机粉尘引起的过敏性肺炎	CA70	26	138

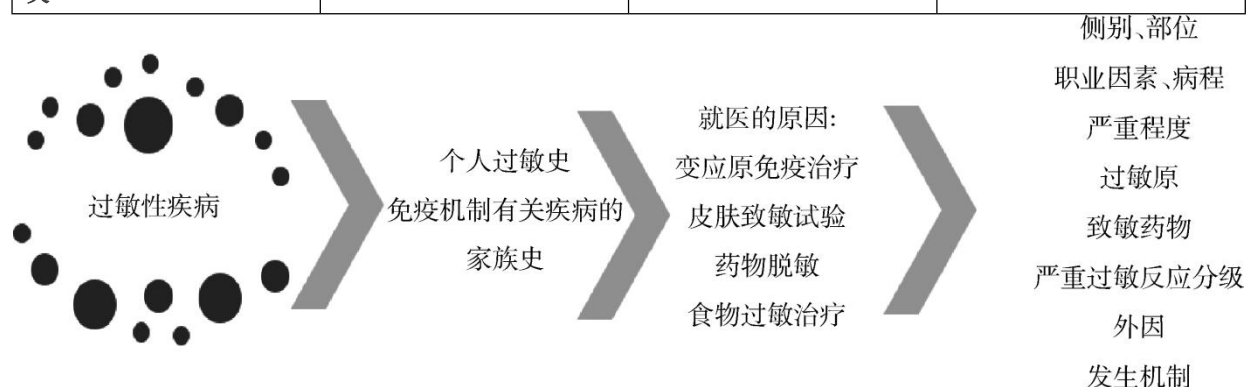


图 3 ICD - 11 可用于过敏性疾病相关信息分类的范畴

案例1：因脸部使用了含有“对羟基苯甲酸乙酯”的化妆品，引起了口周区红疹，最终被诊断为变应性接触性皮炎

目标主干码：EK00.1 化妆品或香料引起的变应性接触性皮炎

└后组配解剖部位：XA1A48 口周区

└后组配关联情况-职业相关性：XB80非职业相关性

└后组配关联情况-过敏原：XM1CG8对羟基苯甲酸乙酯

ICD-11编码簇：EK00.1&XA1A48&XB80&XM1CG8

案例2：既往有青霉素过敏史患者，因医生疏忽，仍然被注射青霉素药物，引起Ⅲ级严重过敏反应

主干码1：4A84.1药物性严重过敏反应

└后组配关联情况：PL00在治疗过程中药物、药剂或生物制剂有关的损伤或损害

└XM7Q57盘尼西林（任何）

└后组配关联情况：PL13.7 患者禁忌使用的药物或物质作为引起损伤或损害的机制

└后组配-严重分级：XS2Y 严重过敏反应Ⅲ级

ICD-11编码簇：4A84.1&XS2Y/PL00&XM7Q57/PL13.7

主干码2：QC44.2药物、药剂或生物制品变态反应个人史

后组配-药物：XM7Q57 盘尼西林（任何）

ICD-11编码簇：QC44.2&XM7Q57

图 4 ICD - 11 在过敏性疾病分类中的应用实例

2.1.1.2 过敏性疾病的共病与多病关联指南，指导数据库设计时建立多系统关联字段，便于单一过敏性疾病与多重过敏性疾病共患的发病机制、流行病学、遗传学、治疗效果等方面进行临床研究。

过敏性疾病共病与多病定义为：同一患者罹患两种及以上具有共同致病机制的过敏相关疾病，多指过敏性鼻炎、过敏性哮喘、特应性皮炎，“共病”强调同时伴随存在的情况，“多病”侧重多病发生于同一患者，包括同时及相继发生的两种情况。多数研究倾向于用过敏性疾病的多病性（allergic multimorbidities）来精炼概况过敏性疾病的这一特质[10, 11]。当重点研究一种过敏性疾病与另一种或多种过敏性疾病同时存在的机制时，“共病”一词就更为贴切[12]。

变应性鼻炎及其对哮喘的影响（Allergic Rhinitis and its Impact on Asthma, ARIA）指南是由世界卫生组织（World Health Organization, WHO）支持并参与的国际指南，是汇集了耳鼻喉科、呼吸科、免疫学、变态反应学等多个学科的专家共同制定的标准化的诊断和治疗方案。2024年7月，变态反应学领域顶级期刊The Journal of Allergy and Clinical Immunology: In Practice 出版了“变应性鼻炎及其对哮喘的影响 2024 最新观点”[Concepts for the development of person-centred, digitally-enabled, artificial intelligence-assisted ARIA care pathways（ARIA 2024）[13]，由全球 56 个国家和地区的 220 名专家共同参与，强调以患者为中心、数字化赋能、人工智能（artificial intelligence, AI）辅助的诊疗理念，使其更具适用性和可持续性。

ARIA 旨在制定适用于 AR 和哮喘多病共存的疾病管理路径，并采用“推荐、评估、发展和评价分级体系”（Grading of Recommendations, Assessment, Development, and Evaluation, GRADE）整合 RCT 和真实世界证据（real-world evidence, RWE），制定以患者为中心的指南。

ARIA-EAACI 工作组联合推荐使用数字生物标志物管理呼吸道变应性疾病。

表 3 呼吸变应性疾病数字生物标志物

疾病	短期监测（时间）	长期监测（时间）
变应性鼻炎	CSMS（1d）	ARCT（1 周），RCAT（1 周）
支气管哮喘	E-DASTHMA（1d）	ACQ（1 周），ACT（4 周）
变应性鼻炎合并哮喘	—	CARAT（4 周）

注：CSMS 为联合症状-药物评分；e-DASTHMA 为日常电子哮喘症状-药物评分；ARCT 为变应性鼻炎控制测试；RCAT 为鼻炎控制评估测试；ACQ 为哮喘控制问卷；ACT 为哮喘控制测试；CARAT 为变应性鼻炎和哮喘控制测试；—为无数据

故在患者随访数据库的设计中，将结合 ARIA 推荐的数字生物标志物进行管理随访患者。

2.1.1.2 基于专家共识及指南推荐各类过敏性疾病诊断标准：

支气管哮喘诊断标准[14]：（符合第 1~4 条或第 4~5 条）：1. 反复喘息、咳嗽、气促、胸闷，多与接触变应原，呼吸道感染，冷空气，物理或化学性刺激，运动以及过度通气（如大笑和哭闹）等有关，常在夜间和（或）凌晨发作或加剧。2. 发作时双肺可闻及散在或弥漫性，以呼气相为主的哮鸣音，呼气相延长。3. 上述症状和体征经抗哮喘治疗有效或自行缓解。4. 除外其他疾病所引起的喘息、咳嗽、气促和 胸闷。5. 临床表现不典型者（如无明显喘息或哮鸣音），应至少具备以下 1 项：（1）证实存在可逆性气流受限：①支气管舒张试验阳性指吸入速效 β_2 受体激动剂后 15 min，FEV1 增加 $\geq 12\%$ ；②抗炎治疗后肺 通气功能改善指给予 ICS 和（或）抗白三烯药物治疗 4~8 周，FEV1 增加 $\geq 12\%$ 。（2）支气管激发试验阳性。（3）PEF 日间变异率（连续监测 2 周） $\geq 13\%$ 。

不典型哮喘的诊断：1. 咳嗽变异性哮喘的诊断：以咳嗽为唯一或主要表现。诊断标准（以下前 4 项为诊断基本条件，如果条件允许，强调尽可能进行肺通气功能检查，以获取可变呼气气流受限的客观诊断依据，避免诊断不足和诊断过度）：（1）咳嗽持续 4 周以上，常在运动、夜间和（或）清晨发作或加重，以干咳为主，不伴有喘息；（2）临床上无感染征象，或经较长时间抗菌药物治疗无效；（3）抗哮喘药物诊断性治疗有效；（4）排除其他原因引起的慢性咳嗽；（5）肺功能检查证实存在 AHR；（6）个人或一、二级亲属过敏性疾病史，或变应原检测阳性，增加诊断可能性。

胸闷变异性哮喘（chest tightness variant asthma, CTVA）的诊断：以胸闷为唯一或主要症状。儿童 CTVA 的诊断主要根据临床表现和肺功能检查存在可变呼气气流受限，并排除可能引起胸闷的其他疾病，诊断标准：（1）胸闷或长叹气为唯一或主要的临床表现，持续时间 >4 周；（2）胸闷发作时肺部听诊无哮鸣音；（3）胸部 X 线片无器质性改变；（4）肺功能检查证实存在 AHR；（5）排除心血管系统、消化系统及精神心理等因素引起的胸闷；（6）抗哮喘治疗有效。没有条件进行肺通气功能检查或不配合检查者，可使用抗哮喘药物进行诊断性治疗，如治疗后 症状缓解，停药后症状反复，支持 CTVA 诊断。

过敏性鼻炎的诊断标准[15, 16]：应根据患者家族史和典型过敏史、临床表现以及与其一致的实验室检测结果制定（1）症状：喷嚏、清水样涕、鼻痒和鼻塞出现 2 个或以上每天症状持续或累计在 1 h 以上，可伴有呼吸道症状（咳嗽、喘息等）和眼部症状（包括眼痒、流泪、眼红和灼热感等）等其他伴随疾病症状（2）体征：常见鼻黏膜苍白、水肿，鼻腔水样分泌物（3）实验室检测：过敏原检测至少 1 种过敏原 SPT 和（或）血清特异性 IgE 阳性；鼻分泌物检测高倍显微镜下嗜酸粒细胞比例 >0.05 为阳性。

由于婴幼儿皮肤点刺或者血清特异性 IgE 检测阴性率较高，同时婴幼儿 NAR 的发病率较低，因此婴幼儿 AR 的诊断，皮肤点刺或者血清特异性 IgE 检测可不作为必要条件，仅根据过敏史、家族史，典型的症状及体征即可诊断。

食物过敏在世界范围内受到广泛关注。自 2010 年来，世界上很多学术团体发布了关于食物过敏，特别是牛奶蛋白过敏（cow's milk protein allergy, CMPA）的指南、共识

或建议。包括 2010 年世界过敏组织 (World Allergy Organization, WAO) 发布的《牛奶蛋白过敏诊断和理论依据》[17], 2012 年 8 月欧洲儿科肠胃病学、肝病和营养学会 (European Society for Paediatric Gastroenterology, Hepatology, and Nutrition, ESPGHAN) 发布的《婴幼儿和儿童牛奶蛋白过敏的诊断方法和治疗原则》[18], 中华医学学会 2013 年发布的《中国婴幼儿牛奶蛋白过敏诊治循证建议》[19] 和 2017 年发布的《食物过敏相关消化道疾病诊断与管理专家共识》[20], 2017 年英国牵头发布的《国际 MAP 指南 (international MAP Milk Allergy in Primary Care guideline) 》, 以下简称《iMAP 指南》[21]。这些指南、共识或建议为食物过敏, 特别是 CMPA 的诊治提供了证据性指导, 而且为临床医师提供了切实可行的临床诊治方案。

欧洲过敏和临床免疫学会 (European Academy of Allergy and Clinical Immunology, EAACI) 以《食物过敏和严重过敏反应指南》和《IgE 介导食物过敏的免疫治疗指南》为基础, 发表了最新版的 IgE 介导食物过敏管理指南[22]。该指南对食物过敏、脱敏 (desensitisation)、缓解或持续无反应 (remission or sustained unresponsiveness) 以及耐受 (tolerance) 进行了定义。**食物过敏**是由免疫系统介导对食物的不良反应。根据发病机制不同, **食物过敏可分为 IgE 介导、非 IgE 介导及 IgE/非 IgE 混合机制介导**。IgE 介导的食物过敏是指接触食物过敏原 2h 内出现皮肤、胃肠道和 (或) 呼吸道症状, 当再次暴露相同食物过敏原后症状再发, 且存在 IgE 致敏和 (或) 效应细胞 (如肥大细胞、嗜碱性粒细胞) 激活的证据。IgE 介导食物过敏除需要与非 IgE 介导和混合机制介导的食物过敏鉴别外, 尚需要与非免疫机制介导 (如代谢性、药理性和毒性) 导致的食物不良反应及其他疾病 (消化系统疾病、心因性疾病等) 相鉴别。**脱敏**代表治疗期间的效果, 即患者在接受治疗后获得的能够摄入一定量食物而不出现过敏症状的能力。**缓解或持续无反应**代表停止治疗后的长期效果, 即停止治疗后能够摄入一定量的食物而不出现过敏症状的能力。**耐受**是指不受次数和数量限制, 无症状地摄入食物的能力。并**建议 IgE 介导食物过敏的过敏原检测方法, 列举了食物过敏问诊的关键信息, 指导数据采集中的免疫学指标字段设计 (如 sIgE、皮肤点刺试验结果) 。**

特应性皮炎 (atopic dermatitis, AD) 的诊断标准: AD 是一种异质性疾病, 表现多种多样, 诊断需要一定标准。目前国外常用的诊断标准包括 Hanifin-Rajka 标准和 Williams 标准 (主要标准: 皮肤瘙痒; 次要标准: ①屈侧受累史, 包括肘窝、腘窝、踝前、颈部 (10 岁以下儿童包括颊部皮疹); ②哮喘或过敏性鼻炎史 (或在 4 岁以下儿童的一级亲属中有特应性疾病史); ③近年来全身皮肤干燥史; ④有屈侧湿疹 (4 岁以下儿童面部颊部/前额和四肢伸侧湿疹); ⑤2 岁前发病 (适用于 >4 岁患者)。确定诊断: 主要标准+ 3 条或 3 条以上次要标准) [23]。

张建中[24]等提出的中国 AD 诊断标准: ①病程超过 6 个月的对称性湿疹; ②特应性个人史和/或家族史 (包括湿疹、过敏性鼻炎、哮喘、过敏性结膜炎等); ③血清总 IgE 升高和/或外周血嗜酸性粒细胞升高和/或过敏原特异性 IgE 阳性 (过敏原特异性 IgE 检测 2 级或 2 级以上阳性)。符合第 1 条, 另外加第 2 条或第 3 条中的任何 1 条即可诊断 AD。此标准在诊断青少年和成人 AD 方面敏感性高于 Hanifin-Rajka 标准和 Williams 标准。

姚志荣[25]等提出的中国儿童 AD 临床诊断标准: ①瘙痒; ②典型的形态和部位 (屈侧皮炎) 或不典型的形态和部位同时伴发干皮症; ③慢性或慢性复发性病程。同时具备以上 3 条即可诊断 AD。典型的形态和部位 (屈侧皮炎) 包括儿童面部和肢端受累; 非典型的形态和部位包括: ①典型的湿疹样皮疹, 发生在非屈侧部位 (头皮皮炎、眼睑湿疹、乳头湿疹、外阴湿疹、钱币状湿疹、指尖湿疹、非特异性手部或足部皮炎/特应性冬季足、甲或甲周湿疹和身体其他部位的湿疹样皮疹) ②非典型湿疹样皮疹。单纯糠疹、唇炎、耳下和耳后/鼻下裂隙、痒疹、汗疱疹、丘疹性苔藓样变异。此标准的敏感性也高于 Hanifin-Rajka 标准和 Williams 标准。

2.2 数据库内容设计（包含 CRF 表等）

2.2.1 临床数据库

2.2.1.1 数据集的建立

数据集是专病数据库信息模型的顶层设计，代表了数据库中数据的精度和广度，决定了专病数据库建成后的可用性、可扩展性和未来潜在应用价值。本专病数据库拟对多种来源数据进行整合，制定统一的数据标准（参照国际过敏性疾病数据标准如 WAO 指南），涵盖患者基本信息、病史、过敏原检测、诊疗过程信息、实验室检查、器械检查、治疗方案、随访记录等 8 个主题数据模块，共包括 18 个条目及各条目记录的 66 个字段内容，并纳入环境因素（如 PM2.5、花粉浓度）、生活方式（饮食、居住环境）等多模态数据。（表 1）。

表 1：临床数据库的数据条目及条目记录的内容

序号	条目	条目记录内容
1	患者基本信息	就诊号、就诊时间、姓名、性别、出生日期、年龄、家庭住址、身高、体重、出生体重、生活方式
2	医保情况	医保类型、医保号
3	联系人	姓名、联系方式
4	医疗服务机构标识信息	机构名称、机构地址
5	医疗服务者标识信息	姓名、职称
6	既往史	食物及吸入过敏原阳性史，过敏性疾病史
7	家族史	一代或二代亲属过敏史 一代或二代亲属过敏性疾病史
8	症状及体征	症状描述、严重程度、症状频率、症状开始时间、诱因、体征描述、环境因素
9	实验室检查	检查项目名称、检查日期、检查值
10	肺功能检查	检查名称、检查日期、检查值
11	呼出气一氧化氮检测	检查名称、检查日期、检测值
12	影像学检查	影像学检查名称、检查日期、检查报告
13	疾病诊断	疾病名称、严重程度、诊断时间、
14	用药	药物名称、药物规格、一次药物的使用剂量、药物使用天数、给药频率、给药途径、给药部位、治疗开始时间、治疗终止时间、用药指导、用药注意事项、
15	既往用药史	药物名称、药物使用天数
16	其他辅助治疗	治疗方法、开始时间、终止时间、注意事项 1
17	随访信息	过敏症状改善情况、身体发育情况、药物不良反应、不良反应处理方法
18	费用	费用类别、费用

2.2.1.2 信息化平台的搭建

开发专病库管理系统，支持数据录入、清洗、检索、分析与可视化。应用区块链技术保障数据安全，符合《个人信息保护法》和医疗伦理要求。

本专病数据库数据来源 HIS、EMR、LIMS、PACS 等中除结构化数据外，必然会包含大量的文本信息。

为此，我们将利用自然语言处理技术，结合医学人工标注、医学图谱等，建立适用于本专病数据库的自然语言处理模型，对检查检验报告和病历文本进行结构化数据转换。表 2 列出了部分字段如症状及体征、实验室检查、用药情况、其他辅助治疗等对应的数据项。本专病数据库会严格按照该数据项的内容对文本信息进行标注和提取，从而实现数据的规范化、结构化。

表 2 专病数据库部分字段对应的数据项

条目	字段	数据项
症状及体征	哮喘	喘息；咳嗽；呼吸困难；活动受限；夜间觉醒；胸闷；夜间或晨间症状加重；症状由病毒感染、锻炼、过敏原暴露、气候改变、大笑、刺激物触发；听诊闻哮鸣音
	过敏性鼻炎	清水样鼻涕；鼻痒；鼻塞；阵发性喷嚏；嗅觉减退；清咽喉；吸鼻子；刺激性干咳；咽异物感；眼痒；眨眼；鼻腔黏膜肿胀；鼻腔黏膜色苍白或充血；有鼻道分泌物；鼻甲肥大；黑眼圈；“过敏性鼻炎敬礼”；“过敏性鼻炎皱褶”；睡眠障碍；日常活动受限；在校表现变差或工作能力下降
	湿疹/特应性皮炎	慢性复发性瘙痒性皮肤病；有屈侧或面颊部皮炎史；皮肤干燥；干皮症；耳根裂纹；鱼鳞病；掌纹症；毛周角化症；有皮肤感染（特别是金黄色葡萄球菌和单纯疱疹病毒感染）倾向；旦尼－莫根眶下褶痕；眶周黑晕；苍白脸；白色糠疹；颈前皱褶
实验室检查	过敏检查	皮肤点刺试验；血清特异性免疫球蛋白 E 水平；血清总免疫球蛋白 E 水平；外周血嗜酸性粒细胞计数
	哮喘专项检查	胸部 X 线检查；肺功能测试；呼出气一氧化氮测试；支气管舒张试验；呼气峰值流量变异率；锻炼激发试验；支气管激发试验
	过敏性鼻炎专项检查	鼻黏膜激发试验；调节性 T 细胞计数、免疫球蛋白 G4 水平、辅助性 T 细胞 1 计数、辅助性 T 细胞 2 计数；嗜酸性粒细胞阳离子蛋白水平
用药情况	哮喘治疗	吸入性糖皮质激素；长效 β_2 受体激动剂；白三烯受体拮抗剂；茶碱；短效 β_2 受体激动剂
	过敏性鼻炎治疗	抗组胺药；糖皮质激素；白三烯受体拮抗剂；肥大细胞膜稳定剂；抗胆碱能药
	湿疹治疗	糖皮质激素；钙调神经磷酸酶抑制剂
	食物过敏治疗	糖皮质激素、抗组胺药
其他辅助治疗		过敏原特异性免疫治疗；抗免疫球蛋白 E 抗体治疗

2.2.1.3 临床数据库 ER 图：

2.2.2 流调数据库

2.2.2.1 数据集的建立

本数据集基于流调 CRF 表格设计，用于收集和分析儿童过敏性疾病的相关数据，涵盖患儿基本信息（如人口学特征、过敏史）、医疗调查服务者信息（数据采集者资质）、家庭/学习/城乡环境暴露（过敏原、污染物、生活习惯）、家族遗传史、潜在混杂因素（饮食、心理、早期暴露等）以及环境监测方法（检测设备与技术）。数据来源于医院门诊/住院患儿、社区筛查、环境监测站、家庭问卷调查及实验室检测报告，通过电子 CRF 表格（如 REDCap、EpiData）和环境传感器数据导出进行采集，并经过双人录入校验、逻辑核查（如年龄与诊断合理性）等质量控制措施，确保数据准确性。数据集采用结构化电子表格形式存储，支持后续统计分析、机器学习建模及环境-疾病关联研究。

搭建 CRF 表格设计：

序号	条目	条目记录内容
1	患儿基本信息	<p>调查时间、姓名、性别、年龄、居住地类型：城市<input type="checkbox"/>农村<input type="checkbox"/>，居住时长（年），家庭住址、身高、体重、既往过敏史：有<input type="checkbox"/>无<input type="checkbox"/>，具体疾病名称：过敏性鼻炎<input type="checkbox"/>，支气管哮喘<input type="checkbox"/>，特应性皮炎<input type="checkbox"/>，过敏性结膜炎<input type="checkbox"/>，荨麻疹<input type="checkbox"/></p> <p>本次调查疾病诊断名称：过敏性鼻炎<input type="checkbox"/>，支气管哮喘<input type="checkbox"/>，特应性皮炎<input type="checkbox"/>，过敏性结膜炎<input type="checkbox"/>，荨麻疹<input type="checkbox"/></p>
2	医疗调查服务者标识信息	姓名、职称（住院医师 <input type="checkbox"/> 主治医师 <input type="checkbox"/> 副主任医师 <input type="checkbox"/> 主任医师 <input type="checkbox"/> ）
3	家庭环境暴露信息（过去一年）	<p>居住信息：</p> <p>房屋类型：平房<input type="checkbox"/>，楼房<input type="checkbox"/>，别墅<input type="checkbox"/>，其他<input type="checkbox"/></p> <p>建筑材料：木质<input type="checkbox"/>，混凝土<input type="checkbox"/>，其他<input type="checkbox"/></p> <p>通风频率：每日<input type="checkbox"/>每周<input type="checkbox"/>偶尔<input type="checkbox"/></p> <p>空调：使用季节：几乎不用<input type="checkbox"/>夏季每日<input type="checkbox"/>全年高频<input type="checkbox"/></p> <p>频率：每日<input type="checkbox"/>每周 3-5 天<input type="checkbox"/> 偶尔<input type="checkbox"/></p> <p>温度设定：<24℃<input type="checkbox"/>24-26℃ <input type="checkbox"/> >26℃<input type="checkbox"/></p> <p>模式：制冷<input type="checkbox"/>制热<input type="checkbox"/>通风<input type="checkbox"/>除湿<input type="checkbox"/></p> <p>空调滤网清洁频率：每月 <input type="checkbox"/>每季度 <input type="checkbox"/>每年<input type="checkbox"/>从未清洁<input type="checkbox"/></p> <p>暖气使用频率：类型：集中供暖（水温/室温） <input type="checkbox"/>电暖器 <input type="checkbox"/>燃气壁挂炉<input type="checkbox"/></p> <p>室温范围：<input type="checkbox"/> <18℃ <input type="checkbox"/> 18-22℃ <input type="checkbox"/> >22℃</p> <p>地毯：有<input type="checkbox"/>无<input type="checkbox"/>，毛绒玩具：有<input type="checkbox"/>无<input type="checkbox"/></p> <p>过敏原检测：</p> <p>PM2.5：家庭室内年均浓度（$\mu\text{g}/\text{m}^3$）</p> <p>花粉：室内花粉种类（如蒿草、豚草等）及浓度（颗粒/m^3）</p> <p>甲醛：室内甲醛浓度（mg/m^3）及检测时间</p> <p>尘螨：床垫/沙发尘螨浓度（单位：个/g 灰尘）</p> <p>其他（霉菌、宠物皮屑、蟑螂过敏原等</p> <p>生活习惯：</p> <p>运动频率与强度：每周几次，每次多少时间，强度：轻<input type="checkbox"/>中<input type="checkbox"/>重<input type="checkbox"/></p> <p>游泳有<input type="checkbox"/>无<input type="checkbox"/></p> <p>睡眠质量与时长：每日几小时睡眠，失眠：有<input type="checkbox"/>无<input type="checkbox"/>，昼夜节律紊乱：有<input type="checkbox"/>无<input type="checkbox"/></p> <p>吸烟/饮酒/电子烟使用情况：主动<input type="checkbox"/>被动暴露<input type="checkbox"/></p> <p>清洁频率：每日<input type="checkbox"/>每周<input type="checkbox"/>每月<input type="checkbox"/></p> <p>宠物饲养：猫<input type="checkbox"/>狗<input type="checkbox"/>鸟<input type="checkbox"/>其它<input type="checkbox"/>，数量：<input type="checkbox"/>只</p> <p>吸烟者同居情况：有<input type="checkbox"/>无<input type="checkbox"/></p> <p>烹饪燃料：天然气<input type="checkbox"/>煤气<input type="checkbox"/>生物质燃料<input type="checkbox"/></p>

		空气净化器使用：是 <input type="checkbox"/> 否 <input type="checkbox"/> 吸尘器使用：是 <input type="checkbox"/> 否 <input type="checkbox"/> 过敏原回避措施执行情况：优 <input type="checkbox"/> 良 <input type="checkbox"/> 差 <input type="checkbox"/>
4	学习/工作环境信息（过去一年）	学校/单位位置：城市中心 <input type="checkbox"/> ，郊区 <input type="checkbox"/> ，农村 <input type="checkbox"/> 教室/办公室通风情况：良好 <input type="checkbox"/> 一般 <input type="checkbox"/> 差 <input type="checkbox"/> PM2.5 暴露水平：年均值， $\mu\text{g}/\text{m}^3$ 花粉暴露：季节性峰值浓度（颗粒/ m^3 ），种类：桑科 <input type="checkbox"/> 豚草 <input type="checkbox"/> 樟树 <input type="checkbox"/> 松科 <input type="checkbox"/> 水稻花粉 <input type="checkbox"/> 柏树 <input type="checkbox"/> 甲醛：新装修教室/办公室检测值： 尘螨暴露：地毯有 <input type="checkbox"/> 无 <input type="checkbox"/> 、布艺家具使用情况：有 <input type="checkbox"/> 无 <input type="checkbox"/>
5	城乡环境监测数据（过去一年）	城市环境监测： PM2.5：年均浓度、季节性变化 花粉：主要种类、月度分布、峰值浓度 其他污染物（NO ₂ 、SO ₂ 、臭氧等） 监测点位置：交通区 <input type="checkbox"/> 、工业区 <input type="checkbox"/> 、居民区 <input type="checkbox"/> 农村环境监测： PM2.5：秸秆焚烧期浓度、年均值 花粉：农作物相关花粉（如油菜花、杨树、柳树等） 生物质燃料使用导致的室内污染：有 <input type="checkbox"/> 无 <input type="checkbox"/> 饮用水源类型：自来水 <input type="checkbox"/> 纯净水 <input type="checkbox"/> 井水 <input type="checkbox"/> 其他 <input type="checkbox"/>
6	家族史	一级亲属过敏性疾病史（父母、兄弟姐妹）：疾病名称：哮喘 <input type="checkbox"/> 湿疹 <input type="checkbox"/> 鼻炎 <input type="checkbox"/> 食物过敏 <input type="checkbox"/> ，关系：父 <input type="checkbox"/> 母 <input type="checkbox"/> 哥哥 <input type="checkbox"/> 弟弟 <input type="checkbox"/> 姐姐 <input type="checkbox"/> 妹妹 <input type="checkbox"/> 二级亲属过敏史（祖父母、叔伯等）：疾病名称：哮喘 <input type="checkbox"/> 湿疹 <input type="checkbox"/> 鼻炎 <input type="checkbox"/> 食物过敏 <input type="checkbox"/> ，关系：爷爷 <input type="checkbox"/> 奶奶 <input type="checkbox"/> 外婆 <input type="checkbox"/> 外公 <input type="checkbox"/> 叔叔 <input type="checkbox"/> 伯伯 <input type="checkbox"/> 舅舅 <input type="checkbox"/> 姑姑 <input type="checkbox"/> 姨姨 <input type="checkbox"/> 家族环境居住相似性：共同暴露因素：吸烟有 <input type="checkbox"/> 无 <input type="checkbox"/>
7	其他潜在混杂因素	饮食习惯：高加工食品 <input type="checkbox"/> 传统饮食 <input type="checkbox"/> 维生素 D：每日 400u <input type="checkbox"/> ，口服周期：多少年 Omega-3 脂肪酸摄入量： 心理因素： 长期压力水平（量表评估，如 PSS-10） 焦虑或抑郁状态（PHQ-9/GAD-7 量表） 疫苗接种史：按计划接种：是 <input type="checkbox"/> 否 <input type="checkbox"/> 抗生素使用频率：每月 <input type="checkbox"/> 每季度 <input type="checkbox"/> 每年 <input type="checkbox"/> 早期生活暴露情况：母乳喂养：是 <input type="checkbox"/> 否 <input type="checkbox"/> ，具体月数 分娩情况：自然分娩 <input type="checkbox"/> 剖宫产 <input type="checkbox"/> 宠物接触年龄：婴儿期 <input type="checkbox"/> 幼儿期 <input type="checkbox"/> 学龄前期 <input type="checkbox"/> 学龄期 <input type="checkbox"/> 农场环境暴露：有 <input type="checkbox"/> 无 <input type="checkbox"/> ，具体月数 疾病负担： 因过敏导致的缺勤或缺课天数/年

		医疗费用支出：年人均
8	环境监测方法	PM2.5 检测仪型号， 花粉采样方法：重力沉降法 <input type="checkbox"/> 体积采样法 <input type="checkbox"/> 分子检测 <input type="checkbox"/> 尘螨检测方法：过敏原蛋白检测方法 <input type="checkbox"/> 显微镜直接计数法 <input type="checkbox"/> <input type="checkbox"/> PCR 分子检测法 <input type="checkbox"/> 尘螨代谢产物检测法 <input type="checkbox"/> 空气采样法 <input type="checkbox"/> 环境传感器技术 <input type="checkbox"/>

2.2.2.2 信息化平台的搭建：

基于流调 CRF 表格设计的数据采集需求，搭建儿童过敏性疾病环境暴露信息化管理平台，实现多源数据（临床信息、环境监测、家族史等）的电子化采集、存储、分析与共享。平台采用 B/S 架构，支持医院、疾控中心、环境监测站等多机构协作，满足流调数据实时录入、质量控制及科研分析需求。

2.2.2.3 流调数据库 ER 图：

2.2.3 随访数据库

2.2.3.1 数据集的建立

基于 CRF 表格设计，建立结构化随访数据库数据集。该随访数据库系统性地整合了儿童过敏性疾病诊疗全周期的结构化数据，涵盖基础信息（人口学资料/生长发育指标）、临床症状学特征（鼻部/眼部/呼吸道/皮肤症状的时序性记录）、体征检查（呼吸系统/皮肤科专科查体）、实验室检测（血常规/过敏原/IgE 等免疫指标）、功能评估（肺功能/FeNO/鼻阻力）、影像学资料（X 线/CT/DICOM 文件）、多维度诊断（哮喘/鼻炎/皮炎等 ICD 标准化诊断及分期分级）、药物治疗（吸入/口服/生物制剂等 18 类药物使用明细）、辅助治疗、生活质量量表（ACT/C-ACT 等 7 种标准化问卷）以及药物不良反应监测，形成覆盖门诊随访、检查检验、治疗干预和疗效评估的一体化科研数据库体系。

随访数据库搭建 CRF 表格设计：

序号	条目	条目记录内容
1	患儿基本信息	门诊号、就诊时间、姓名、性别、年龄、家庭住址、身高、体重、初诊 <input type="checkbox"/> 复诊 <input type="checkbox"/>
2	医疗服务者标识信息	姓名、职称（住院医师 <input type="checkbox"/> 主治医师 <input type="checkbox"/> 副主任医师 <input type="checkbox"/> 主任医师 <input type="checkbox"/> ）
3	症状	症状：鼻痒 <input type="checkbox"/> 、喷嚏 <input type="checkbox"/> 、流清涕 <input type="checkbox"/> 、鼻塞 <input type="checkbox"/> 、眼痒 <input type="checkbox"/> 、流泪 <input type="checkbox"/> 、眼红 <input type="checkbox"/> 、眼灼热感 <input type="checkbox"/> 、咽痒 <input type="checkbox"/> 、发热 <input type="checkbox"/> 、张口呼吸 <input type="checkbox"/> 、打鼾 <input type="checkbox"/> 、咳嗽 <input type="checkbox"/> 、喘息 <input type="checkbox"/> 、气促 <input type="checkbox"/> 、胸闷 <input type="checkbox"/> 、长叹气 <input type="checkbox"/> 、活动受限 <input type="checkbox"/> 、夜间憋醒 <input type="checkbox"/> 、夜间及晨起症状加重 <input type="checkbox"/> ，皮疹 <input type="checkbox"/> ，瘙痒 <input type="checkbox"/> ，疼痛 <input type="checkbox"/> ，睡眠障碍 <input type="checkbox"/> 严重程度：轻度 <input type="checkbox"/> 、中度 <input type="checkbox"/> 、重度 <input type="checkbox"/> 、极重度 <input type="checkbox"/> ； 症状频率：一日 <input type="checkbox"/> 次、持续 <input type="checkbox"/> /反复 <input type="checkbox"/> ， <input type="checkbox"/> 天/ <input type="checkbox"/> 月/年、一年发作 <input type="checkbox"/> 次； 症状开始时间： <input type="checkbox"/> 年 <input type="checkbox"/> 月 <input type="checkbox"/> 日； 诱因（接触吸烟 <input type="checkbox"/> 、尘螨 <input type="checkbox"/> 、花粉 <input type="checkbox"/> 、病毒感染 <input type="checkbox"/> 、霉菌 <input type="checkbox"/> 、冷空气 <input type="checkbox"/> 、气候改变 <input type="checkbox"/> 、运动 <input type="checkbox"/> 、哭闹 <input type="checkbox"/> 、食物 <input type="checkbox"/> 、药物 <input type="checkbox"/> 、其它 <input type="checkbox"/> ）
4	体征	体温： <input type="checkbox"/> ℃，脉搏： <input type="checkbox"/> 次/分，血氧饱和度： <input type="checkbox"/> %（有无吸氧），神志（清醒 <input type="checkbox"/> 、谵妄 <input type="checkbox"/> 、昏睡 <input type="checkbox"/> 、昏迷 <input type="checkbox"/> ），精神（良好 <input type="checkbox"/> 、倦怠 <input type="checkbox"/> ），呼吸（匀称 <input type="checkbox"/> 、急促 <input type="checkbox"/> 、不规则 <input type="checkbox"/> ），外观（消瘦 <input type="checkbox"/> 、正常 <input type="checkbox"/> 、肥胖 <input type="checkbox"/> ），鼻腔粘膜：苍白 <input type="checkbox"/> 、

		<p>水肿□、肿胀□、清涕□、评分□，三凹征（有/无），左肺呼吸音（清/粗），左肺闻及干啰音（有/无，满布/散在），左肺闻及湿啰音（有/无，满布/散在），右肺呼吸音（清/粗），右肺闻及干啰音（有/无，满布/散在），右肺闻及湿啰音（有/无，满布/散在），心率□次/分，心律（整齐/不齐），心脏闻及杂音（有/无，收缩期/舒张期/无），腹（平坦、隆起、凹陷），腹压痛（有/无），反跳痛（有/无），</p> <p>皮疹分布：头颈：红斑□水肿/丘疹□渗出/结痂□表皮剥脱□苔藓样变□干燥□及其他；</p> <p>上肢：红斑□水肿/丘疹□渗出/结痂□表皮剥脱□苔藓样变□干燥□及其他；</p> <p>躯干：红斑□水肿/丘疹□渗出/结痂□表皮剥脱□苔藓样变□干燥□及其他；</p> <p>下肢：红斑□水肿/丘疹□渗出/结痂□表皮剥脱□苔藓样变□干燥□及其他；</p> <p>生殖器：红斑□水肿/丘疹□渗出/结痂□表皮剥脱□苔藓样变□干燥□及其他；</p> <p>电子呼吸音（音频植入）</p>	
5	实验室检查	<p>检查项目名称、检查日期、检查值</p> <p>血常规（检查日期）：白细胞计数、中性粒细胞计数、淋巴细胞计数、嗜酸性粒细胞计数、中性粒细胞百分比、淋巴细胞百分比、嗜酸性粒细胞百分比、血红蛋白、红细胞计数、血小板、C反应蛋白</p> <p>过敏原检查（检查日期）：总 IgE：□，尘螨组合□，霉菌组合□，宠物毛屑组合□，常见食物组合□，坚果组合□，过敏原单项检查（检查日期）：屋尘螨□粉尘螨□花粉□杂草□烟曲霉□链格孢□蟑螂□猫皮屑□狗毛屑□鸡蛋白□牛奶□虾□螃蟹□大豆□芝麻□小麦□坚果□其它□</p> <p>皮肤点刺试验（检查日期）：屋尘螨□，粉尘螨□</p> <p>（检查日期）痰嗜酸性粒细胞计数□</p> <p>（检查日期）总 IgE：数值（嵌入报告）</p>	
6	肺功能检查	<p>检查名称、检查日期、检查值</p> <p>肺功能测试（检查日期）：IC□，MEF50□，FEV1□，FVC□，MMEF75/25□，FEV1/FVC□，MEF25□，MEF75□，MVV□，FEV1/VCmax□，VCmax□，结论：</p> <p>支气管舒张试验（检查日期）：阳性□阴性□</p> <p>支气管激发试验（检查日期）：阳性□阴性□</p> <p>潮气肺功能检查（检查日期）：WT/kg□，RR□，Ti□，Te□，Ti/Te□，TPTEF□，VPTEF□，TPTEF/TE□，VPEF/VE□，PTEF□，TEF50/TIF50□，TEF50□，TEF25□，TEF25-75□</p>	
7	呼出气一氧化氮检测	<p>检查名称、检查日期、检测值</p> <p>呼出气一氧化氮（检查日期）：FeN050□，FeN0200□，</p>	

		CaNO□	
8	电子鼻咽喉镜	检查名称、检查日期、检查值 结论	
9	听力检查	检查名称、检查日期、检查值 声导抗□、纯音听阈测定□	
10	鼻阻力	检查名称、检查日期、检查值 结论：	
11	影像学检查	影像学检查名称、检查日期、检查报告 胸部 X 线（检查日期）：影像导入、结论 胸部 CT（检查日期）：影像导入、结论	
12	疾病诊断	疾病名称、严重程度、诊断时间 支气管哮喘：急性发作期□，临床缓解期□，慢性持续期□，轻度□，中度□，重度□，极重度□ 过敏性鼻炎：□轻度□中-重度□季节性□常年性□间歇性□持续性□ 特应性皮炎□：轻度□中重度□ 鼻窦炎：急性□慢性□ 腺样体肥大□ 扁桃体肥大□ 鼾症□ 阻塞性呼吸睡眠暂停低通气综合症□ 过敏性结膜炎□ 其它	
13	用药	药物名称、药物规格、一次药物的使用剂量、药物使用天数、给药频率、给药途径、给药部位、治疗开始时间、治疗终止时间、用药指导、用药注意事项 糖皮质激素 ：布地奈德（阿斯利康□正大天晴□）□使用方法，丙酸氟替卡松气雾剂□使用方法，布地奈德/福莫特罗 80/4.5□使用方法，布地奈德/福莫特罗 160/4.5□使用方法，沙美特罗氟替卡松吸入剂 50/100□使用方法，沙美特罗氟替卡松气雾剂 25/125□使用方法，沙美特罗氟替卡松吸入剂 50/250□使用方法，泼尼松□使用方法，甲泼尼松龙□使用方法，甲强龙□使用方法，地塞米松□使用方法， 鼻用糖皮质激素 ：布地奈德鼻喷雾剂□使用方法，糠酸莫米松鼻喷雾剂□，丙酸氟替卡松鼻喷雾剂□使用方法 外用激素软膏：糠酸莫米松乳膏□，丙酸氟替卡松软膏□，卤米松三氯生乳膏□，曲安奈德益康唑乳膏□，钙调磷酸酶抑制剂：他克莫司□，吡美莫司□ 磷酸二酯酶抑制剂：克立硼罗□ 莫匹罗星软膏：□ 度普利尤单抗（开始时间、剂量）：□ 阿布希替尼片（开始时间、剂量）：□ 乌帕替尼缓释片（开始时间、剂量）：□	

		支气管扩张剂： 特布他林吸入剂□使用方法，沙丁胺醇吸入剂□使用方法，复方异丙托溴铵吸入剂□使用方法，丙卡特罗□使用方法 抗组胺药： 西替利嗪□，氯雷他定□，酮替芬□，赛庚啶□，氯苯那敏□ 白三烯受体拮抗剂： 孟鲁司特□ 祛痰药： 氨溴索□，乙酰半胱氨酸□，桉柠蒎□，浓盐水雾化□ 抗生素： 头孢地尼□，头孢克洛□，阿奇霉素□，阿莫西林克拉维酸钾□ 过敏原特异性免疫治疗（开始时间）： 畅迪□，安脱达□，屋尘螨片剂□ 抗免疫球蛋白 E 治疗（开始时间、剂量）： 奥马珠单抗□ 靶向治疗（开始时间、剂量）：	
14	既往用药史	药物名称、药物使用天数 头孢类抗生素□，大环内酯类抗生素□，糖皮质激素□，抗组胺药□，白三烯受体拮抗剂□，祛痰药□，环孢素□，阿布希替尼□，乌帕替尼□	
15	其他辅助治疗	治疗方法、开始时间、终止时间、注意事项 支气管镜检查□（检查日期）结论，图片导入 吸氧（有/无）：鼻导管□，面罩□，高流量给氧□，机械通气□	
16	调查问卷	ACT 评分问卷（≥12 岁），C-ACT 评分（4-11 岁），TRACK 评分（≤5 岁），症状药物评分，儿童皮肤病生活治疗指数，儿童特应性皮炎 Scoard 评分，婴儿皮肤病生活质量指数，鼻炎 VAS 评分，评分时间□年□月□日，	
17	随访信息	过敏症状改善情况、身体发育情况、药物不良反应、不良反应处理方法	
18	费用	费用类别、费用	
19	药物不良反应	药品名称，不良反应发生时间，药物用法用量，服用天数，皮疹□瘙痒□腹痛□恶心□呕吐□胸闷□心悸□呼吸困难□其它□ 严重程度：轻度□中度□重度□	

2.2.3.2 信息化平台的搭建：

基于流调 CRF 表格设计，搭建儿童过敏性疾病随访数据信息化管理平台，实现多源数据（临床信息、实验室检查、肺功能、影像学、用药记录、生活质量评估等）的电子化采集、结构化存储、智能分析与多中心共享。平台采用 B/S（Browser/Server）架构，支持医院、科研机构、随访中心等多机构协作，满足临床随访数据实时录入、自动质控、动态分析及科研数据导出需求。

2.2.3.3 随访数据库 ER 图：

2.3 数据采集与质控（包含多模态数据等）

2.3.1 数据采集实现策略

2.3.1.1 多模态数据源定义与整合：

本专病库采集的数据类型包括：**临床数据**：电子病历（EMR）、实验室报告（IgE 检测、皮肤点刺结果）、影像学（如胸片、肺部 CT、鼻窦 CT）、患者报告结局（PROs）。**组学数据**：基因组（GWAS 或外显子测序）、蛋白质组（血清细胞因子谱）、代谢组（尿液/血液代谢物）、肠道菌群鉴定。**环境数据**：我省气象台整合的过敏原浓度（如花粉、PM2.5、尘螨等环境数据）、气象数据（温湿度）、空气污染指数（PM2.5、NO2）。**随访数据**：包括电话或微信随访信息（如药物使用、PEF 监测、药物不良反应、吸入剂使用依从性和方法、疗效评估问卷等），后期可在数据库基础上研究设计可穿戴设备（哮喘监测 APP）、IoT 传感器（家庭尘螨监测仪）等。

数据采集的工具与相关技术如下：不同的数据来源不同，如**临床数据**：可通过医院信息系统（HIS/LIS）API 对接，或标准化电子数据采集表（EDC）如 REDCap、OpenClinica。**组学数据**：利用实验室信息管理系统（LIMS）导出结构化数据（如 FASTQ、mzML 格式）。**环境数据**：依托公共数据库（如 NOAA、WHO 空气质量平台）自动爬虫，或第三方 API 实时接入。**患者自报告**：定制化移动端 APP（集成 VAS 评分、症状日历）或微信小程序可实时将患者呼吸音数据、用药、家庭环境监测等数据自动上传。

2.3.1.2 标准化数据录入流程

结构化模板设计：遵循 CDISC（临床数据交换标准）或 OMOP（观察性医疗结果伙伴）模型，诊断名称依据 ICD-11 编码标准，过敏原名称采用 WHO/IUIS 标准命名，确保字段统一，例如过敏性鼻炎症状字段按《ARIA 指南》分类（鼻塞、流涕、喷嚏、鼻痒）。

多学科协同采集：建立数据字典（Data Dictionary）和操作手册（SOP），明确各学科职责（详见 2.4）。

2.3.1.3 自动化与半自动化采集

自然语言处理（NLP）：从非结构化文本（如医生病程记录、出院小结、入院记录等）提取关键信息（如“尘螨过敏史” → 结构化字段 allergen: Dermatophagoides）。**物联网（IoT）集成**：通过蓝牙/Wi-Fi 连接患者家中的空气质量传感器，自动上传 PM2.5 和温湿度数据。**区块链存证**：关键数据（如知情同意书、原始检测报告）上链，确保数据不可篡改。

2.3.2 数据质量控制（QC）全流程

2.3.2.1 质控前：在数据标准文档中定义数据范围（如 tIgE 值单位必须为 kU/L，单项检测 IgE 单位 KUA/L）、允许值域（如年龄 ≥ 0 且 ≤ 120 ）、逻辑关系（如“尘螨阳性”患者必须填写 sIgE 值）等，系统可自动实变错误，及时提醒医生纠正。搭载**自动化校验工具**，使用开源工具（如 Great Expectations）或商业软件（如 Informatica）配置校验规则，实时拦截异常数据。

2.3.2.2 质控中：分阶段质量控制：

录入阶段质控：在医务人员输入数据时触发逻辑检查（如日期格式错误提示）。如果输入关键字段（如过敏原类型）可由两名独立录入员比对，差异 $>5\%$ 则触发人工复核。

传输阶段质控：通过哈希值（MD5/SHA256）验证文件传输检测数据完整性，并记录数据转换日志（如缺失值填充方法），确保可追溯。

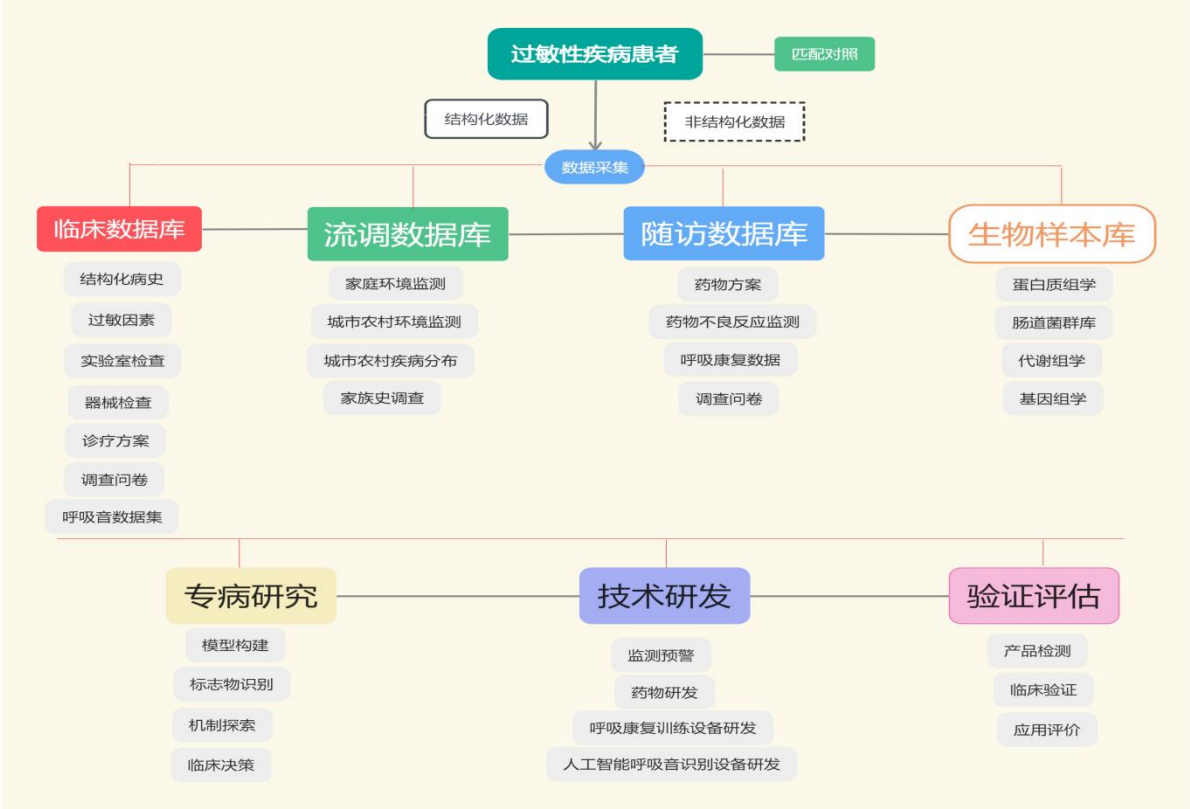
存储阶段质控：采用 RAID 10 或分布式存储（如 Hadoop HDFS）防止数据丢失，使用 Git 或 DVC 管理数据版本，记录每次更新内容。

2.3.2.3 质控后：设置动态数据看板，此构建可视化仪表盘（如 Tableau、Power BI），实时监控数

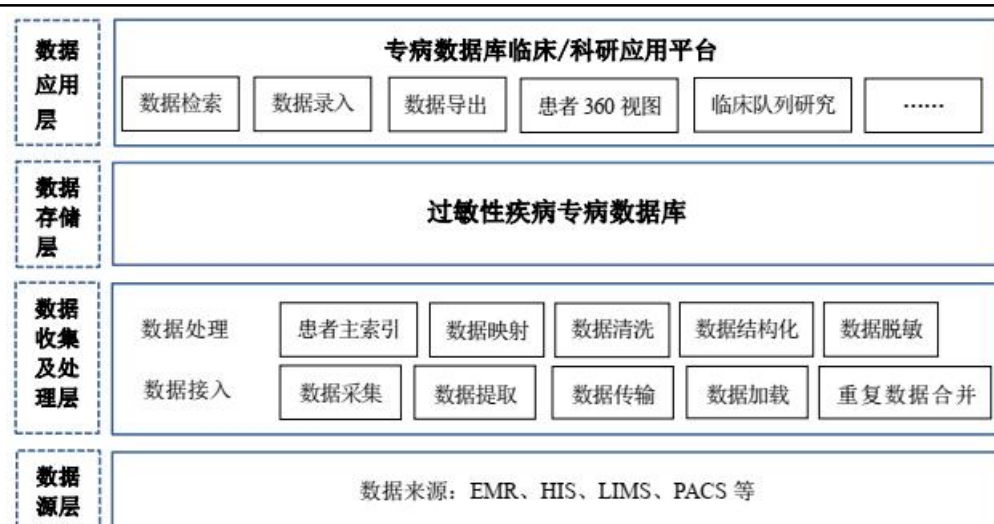
据质量指标，为保证数据**完整性**，降低缺失率（如症状评分缺失<5%）。后期若有多中心合作，则跨中心数据分布差异（如不同医院 IgE 检测方法校准）应保持**一致性**。保证数据**实时性**，数据更新延迟（如环境数据滞后<24 小时）。定期进行**人工抽样复审**，随机抽取 5%~10%病例，由临床专家复核数据准确性（如将原始病历与录入数据比对）。

3. 技术路线：

3.1 专病库建设总体框架图



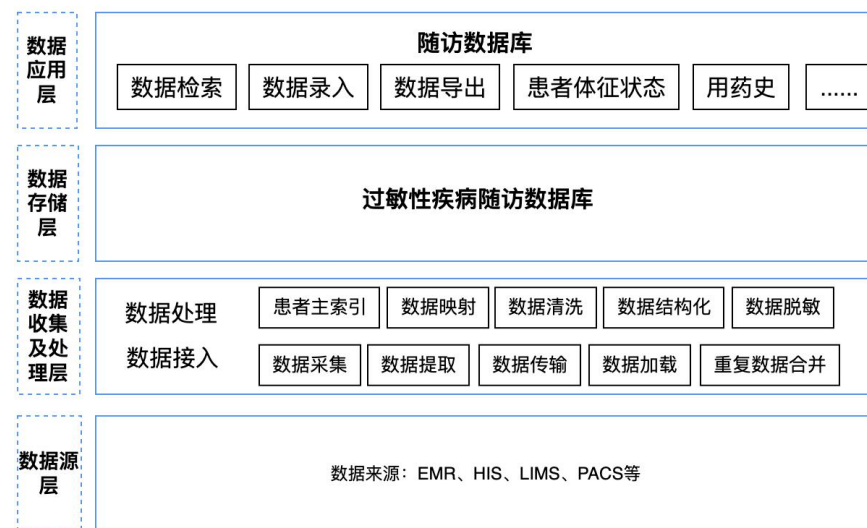
3.2 临床数据库建设设计框架图：



3.3 流调数据库建设设计框架图:



3.4 随访数据库建设设计框架图:



4. 数据库建设方案

4.1 数据库预期规模（包含预期样本量、变量数量等）

过敏性疾病数据库的预期规模需根据研究目标、资源投入和实际应用场景综合设计，本专病库建设预期规模定如下内容：

4.1.1 样本量规划

（1）核心样本量估算

首先评估基础研究需求样本量，目前我院单中心数据库，在短期（1-3 年），采集 3,000-5,000 例患者（涵盖过敏性鼻炎、哮喘、特应性皮炎、食物过敏等主要疾病）。长期规划（5-10 年），采集 10,000-20,000 例过敏患者病例信息，覆盖地域扩展。后期若条件成熟，可考虑成立多中心联盟数据库（待定），初始阶段采集约 10,000-30,000 例（联合 3-5 家大型医疗机构）。技术成熟后，可扩展至 50,000-100,000 例，并可考虑覆盖全国或国际协作网络。

依据统计效力考量样本量，若用于机器学习模型训练，每个亚组（如特定过敏原或疾病亚型）需 ≥ 500 例以满足模型泛化需求。若探索基因-环境交互作用，需 $\geq 10,000$ 例以检测中等效应量（OR=1.2-1.5）。

（2）不同过敏性疾病的样本量类型分布：依据疾病类型分类，过敏性鼻炎占（50%-60%）、哮喘占（20%-30%）、特应性皮炎占（10%-15%）、食物过敏占（5%-10%）。依据人群分层，年龄分层（儿童/成人/老年）、地域分布（城市/农村）、过敏原类型（尘螨/花粉/宠物皮屑/霉菌等）。

4.1.2 变量数量与维度

（1）核心变量涵盖了临床数据库、随访数据库中的类别，其中基线变量包含 50-100 个字段（如人口学信息、病史、过敏原检测结果、合并症、环境监测数据）。动态监测变量：20-30 个/次随访（如症状评分、药物使用、不良反应）。

中心实验室负责生物样本库建设，收集组学数据，包括基因组：1,000,000+ SNPs（全基因组关联分析）或 200-500 个候选基因（靶向测序）。蛋白质组：100-200 种血清蛋白标志物（如细胞因子、免疫球蛋白）。代谢组：500-1,000 种代谢物（LC-MS/GC-MS 检测）。肠道菌群鉴定（微生物群及其代谢产物主要是短链脂肪酸）

流调数据中的环境监测数据，依据福建省气象台监测数据，及我院花粉监测，居家尘螨数据，每日/每周动态变量约 10-20 个，包括花粉浓度、PM2.5、温湿度。

（2）数据库总体变量规模，其中每位患者约 200-500 个核心变量（临床+环境+随访）构成结构化数据。每位患者附加 1,000,000+ 变量）生物组学数据。若包含 10,000 例患者，结构化数据规模约 200 万-500 万数据点，组学数据达百亿级。

4.1.3 数据采集的时间跨度，单次采集也就是基线数据占比 60%-70%。随访周期：1-5 年，每 3-6 个月更新一次，包含症状、环境暴露、治疗反应、疗效评估等。动态变量占比 30%-40%。

4.1.4 预计样本量动态扩展数量，估计我院单中心每年新增 2,000-5,000 例，若有实现多中心，则每年样本量可增加 10,000-20,000 例。随研究进展扩展新模式（如肠道菌群宏基因组、单细胞转录组），可每新增模式 100-1,000 变量/样本。

4.2 建设过程

4.2.1 首先是用户需求分析，进行专病库建设规划阶段

(1) 前期先进行临床调研，明确数据库用途（如疾病机制研究、疗效预测、公共卫生决策支持等），确定核心研究问题（如环境-基因交互作用、长期治疗随访等）。

(2) 评估数据来源的可行性，评估医院与合作大学、实验室、省气象站的可及性，其次评估福大合作团队在数据工程、机器学习、隐私保护方面的技术储备。评估预算与项目执行周期，综合考量硬件（服务器、云存储）、人力（数据工程师、临床专家）、时间成本等。

(3) 初步设置项目执行周期，可分阶段目标，1 年完成基础临床数据库建设，3 年内扩展至多模态数据。制定标准化文档，如数据字典、伦理审查流程、SOP 操作手册等。

4.2.2 第二阶段初步完成数据库设计与建模

(1) 针对不同的数据来源选择不同的数据处理方式，依据关系型数据库，可参考使用 PostgreSQL 存储临床表单数据，形成结构化数据。利用 NoSQL（如 MongoDB）或数据湖（如 AWS S3）存储组学原始数据、影像文件，进行非结构化处理。采用 OMOP-CDM 或 FHIR 标准，确保跨机构兼容性，形成标准化框架。

(2) 采用分层方式进行数据架构：分为**原始层**存储未经处理的原始数据（如 DICOM 影像、测序 FASTQ 文件）。**整合层**存储经清洗、转换后的结构化数据（如症状评分表、环境指标时序数据）。**应用层**存储面向研究者的可视化查询接口或 API 服务。在技术上，可使用 Apache Kafka 处理实时数据流（如可穿戴设备数据），采用分布式存储（如 Hadoop HDFS）应对高维组学数据。

4.2.3 第三阶段进行数据采集与整合

(1) **对多模态数据进行录入**，通过医院 HIS/LIS 系统 API 对接，或人工录入标准化 EDC 系统（REDCap），使用 NLP 工具（如 CLAMP）解析非结构化病历文本，进行**临床数据**采集。与实验室 LIMS 系统集成，自动导入测序或质谱数据，形成**组学数据**。可爬虫抓取公共数据库（如中国环境监测总站、福建省气象站），或 API 接入气象局实时数据，采集**环境数据**。利用医院到访、微信、电话等方式采集随访数据集。

(2) 采集后的数据需进行**标准化处理**，过敏原名称统一为 WHO/IUIS 标准，疾病编码采用 ICD-11，将不同单位的检测结果（如 IgE 值）转换为统一计量单位（kU/L），将环境数据与患者症状时间戳匹配，填补缺失时段（线性插值或均值填充）。

4.2.4 第四阶段进行技术平台的搭建