**儿童感染性疾病症候群数据集信息化处理**

**服务概况及要求**

**一、项目概况**

“儿童感染性疾病症候群数据集信息化处理服务 ”， 利用大数据平台中疾病数据信息域编撰并发表儿童感染性疾病症候群标准数据集。设置科研域模块，包括：儿童感染性疾病的基础信息及就诊信息、临床表征、病史信息、体格检查、辅助检查、诊断信息及治疗信息、疾病评估及结局。服务内容包括制定数据集、数据集修稿处理、项目管理等工作。

**二、危急重症医疗平台功能描述**

**1、服务内容：**

服务内容包括数据集设计、数据集修稿处理、项目管理等工作，具体如下：

|  |  |  |
| --- | --- | --- |
| **序号** | **医学服务类别** | |
| 1 | 文献调研 | Uptodate、专病教科书、专病临床路径及专病诊疗指南、综述、领域文献调研 |
| 调研资料汇总及分析 |
| 2 | 数据集表单设计 | 数据集内容规划 |
| 数据集大纲及主题模块设计 |
| 数据集变量设计,包括数据集基础信息及就诊信息、临床表征、病史信息、体格检查、辅助检查、诊断信息及治疗信息、疾病评估及结局等主题模块的变量设计 |
| 发表小组的内部数据集表单质量评审及修改 |
| 3 | 数据字典设计 | 数据字典数据模块框架绘制 |
| 数据字典基础维度序号、域名、中文名称、英文名称、值域/格式、单位等维度的内容填充 |
| 数据字典基础维度变量类型、数据加工类型（数据提取方式）、数据来源、版本号等维度的内容填充 |
| 发表小组的内部数据字典质量评审及修改 |
| 4 | 发表版本内容交付 | 封面页、内页、主编简介、序言、目录、设计理念/数据集说明、纳排、附录：症状对照表、参考文献等出版物常规内容草拟 |
| 5 | 出版社审核及修订 | 配合出版社审核与修订、PDF排版及美化 |

**2、服务详细描述**

**2.1.文献调研**

**2.1.1.Uptodate调研**

于Uptodate中检索疾病相关知识内容，并汇总。

**2.1.2.专病教科书调研**

于对应教科书中查阅疾病相关知识，并汇总。

**2.1.3.专病临床路径调研**

于国家已发布的临床路径中搜索是否存在专病相关临床路径。

**2.1.4.专病诊疗指南、综述、领域文献调研**

于国内外文献数据库中检索项目相关文献，获取疾病诊疗指南、综述、领域文献。

**2.1.5.调研资料汇总及分析**

汇总所有文献，分析疾病领域相关研究内容、热点研究方向，初步圈定疾病领域研究框架，解构拆分研究的变量池。

**2.2.数据集表单设计**

数据集是指围绕着某一项特定任务或活动进行数据规划和设计，对其内容进行的系统归纳和描述的数据集合。数据集通常由数据元所形成的若干数据记录所构成，可通过不同的命名，将相同属性的主题归并在一起形成相同的类，将不同属性的主题区分开形成不同的类，最终通过规范的名称和标识符等对数据集进行标记识别。在内容层面可以认为数据集表单是大型队列数据库的eCRF表单集合。

**2.2.1.数据集内容规划**

参考文献调研整理的疾病领域研究框架、院方需求以及既往相关疾病的项目经验，规划数据集构建策略，最终与客户沟通确认。

**2.2.2.数据集大纲及主题模块设计**

基于客户确认的数据集构建策略，设计数据集表单主题模块及内容大纲，并与客户沟通确认。

**2.2.3.数据集变量设计**

遵循科学、创新、先进、可行、意义、准确、完整、规范等设计原则，基于调研阶段完成的研究变量池，完成所有数据集表单主题模块的变量筛选、特等细分研究变量的增补及最终的变量标化，包括：患者的基础信息及就诊信息、临床表征、病史信息、体格检查、辅助检查、诊断信息及治疗信息、疾病评估及结局等科研域模块。

**患者的基础信息及就诊信息**：这部分旨在制定针对患者基础情况的信息采集规范，是在临床诊疗活动和临床研究中，对患者常规采集的个人相关内容。此部分涵盖人口学信息、生物样本管理信息、就诊信息。其中人口学信息指人口学特征，即患者身份识别、民族、性别、看护人信息等特征信息；生物样本库相关的患者身份识别、等信息；就诊信息涵盖了急诊就诊信息（包括急重症就诊信息）、门诊就诊信息、住院就诊信息等3方面的内容。

**临床表征**：本章内容主要为患者临床表征信息（症状体征，疾病特异性表现）。其中涵盖发热呼吸道症候群特征、发热肠道症候群特征、发热伴出疹症候群特征、脑炎脑膜炎症候群特征、发热伴出血症候群特征等。

**病史信息**：本章内容主要为患者病史信息。其中包括但不仅限于：既往疾病史、特殊用药史、手术史、过敏史、输血史、月经史、家族史、预防接种史、个人接触史和出生史等。

**体格检查**：体格检查是临床实践过程中最基础、最直接且不可或缺的评估手段，它是医生运用自身感官（视、触、叩、听、嗅）或借助简单工具（如听诊器、叩诊锤、体温计、血压计、压舌板、手电筒等），对患者身体状况进行系统、有序的观察和评估的过程，通常是医生接触患者后进行的首要评估步骤，为后续的辅助检查选择和诊断思路提供最直接的线索和方向。主要包含：一般状态、全身评估、头颈部、胸部、腹部、脊柱四肢神经系统等。

**辅助检查**：辅助检查是临床实践过程中的核心组成要素，贯穿临床路径始终。对于儿童感染性疾病症候群，其包括但不限于以下内容：检验（一般检验、微生物培养、结核菌素试验、结核干扰素测定）、影像学检查（超声检查、放射、核磁）、电生理检查（心电图、脑电图）。

**诊断信息及治疗信息**：诊断信息主要关注儿童感染性疾病的一般诊断和症候群标识。其中症候群标识包括:发热呼吸道症候群、发热肠道症候群、发热伴出疹症候群、脑炎脑膜炎症候群、发热伴出血症候群。治疗信息内容包括儿童感染性疾病症候群相关的治疗方法，包括但不仅限于：药物治疗、特殊治疗方案（肾脏替代治疗、呼吸支持、体外膜肺氧合ECMO）等。

**疾病评估及结局**：为了更好地支持循证医学研究，这部分主要是针对疾病严重程度、进展、治疗和结局的评估，包括量表评估、疗效评估、并发症/重要医学事件。

**2.2.4.发表小组的内部数据集表单质量评审及修改**

内部审阅小组成员根据自身专业提出修改意见，制作小组完成修改且审阅小组无异议后，给到专家组(KOL)进行首轮内容审核，最终完成数据集版本初稿定稿。

**2.3.数据字典设计**

数据字典是对数据的数据项、数据结构、数据流、数据存储、逻辑处理进行定义和描述，对数据流程中的各个元素详细说明，是描述数据的信息集合。通过数据表单转化而来的数据字典内容可以进行发表、交换、管理和查询应用。这些数据可以由不同的物理存储格式来实现，按照数据元的定义与数据类型，在计算机系统中以数值、日期、字符、图像等不同的类型表达。

**2.3.1.数据字典数据模块框架绘制**

与客户确认数据字典需要包含的维度模块，并参考定稿版本的数据集表单完成数据字典各主题的框架绘制。

**2.3.2.数据字典基础维度的内容填充**

数据字典基础维度的内容填充，包括序号、域名、中文名称、英文名称、值域/格式、单位等维度。

**序号**：字段的编码字符串。由数据域-域模块-序号编码组成，举例：某数据元为数据集第二个数据域“临床表征”下第二个域模块“发热呼吸道症候群特征”中的排序第三位的数据元，则序号编码为“2.2.003”。。

**域名**：用于将同一使用类型的数据元进行归类。

**中文名称、英文名称**：数据元的标识方式，本着标准、明确、无歧义的原则进行命名，选择临床指南、文献或术语系统中出现的名称作为标准名，对于在临床中存在多个命名的数据元，会选取一个标准命名。

**值域/格式**：参考指南和文献兼顾本领域专家实用性，囊括最大可能范围的允许值的集合/格式，数值通常还需要有单位描述才能准备表达值的含义。

**单位**：数值型数据元所对应的数据单位。

**2.3.3.数据字典可选维度的内容填充**

数据字典可选维度的内容填充，包括变量类型、数据加工类型（数据提取方式）、数据来源、版本号。

**变量类型**：变量类型指该数据元的值对应的数据类型。

**数据加工类型（数据提取方式）**：根据数据来源及数据上层加工处理，将数据分为三类：分别为直接映射的数据，需结构化和归一的数据，手动填写数据。直接映射的数据：此部分数据为临床诊疗过程中产生，同时能够以ETL或NLP技术较为准确的进行自动提取的数据。需结构化和归一的数据：将大段非结构化文本，通过自然语言处理技术处理为标准的数据，并可进行统计分析的数据。手动填写数据：部分非院内数据，或临床诊疗过程中较少出现数据，需用户自行录入的数据。

**数据来源**：每个数据元后会标注数据源，包括但不限于国内外医学信息学信息标准、权威指南、相关文献和术语规范。

**版本号**：版本号由启用状态-收录年份-版本次数组成，其中A为启用/B为失效/C为预备版本，举例：①某数据元在2023年收录于本数据集，并正式启用，作为正式发布的第1版，那么该数据元目前版本号为：“A-2023-1”；②某数据元在2023年收录于本数据集，并在正式发布的第4次更新中标记失效，那么该数据元目前版本号为：“B-2023-4”。

**2.3.4.发表小组的内部数据字典质量评审及修改**

内部审阅小组成员交叉评审，制作小组完成修改且审阅小组无异议后，完成数据字典版本定稿。

**2.4.发表版本内容交付**

**2.4.1.出版物常规内容草拟**

出版物常规内容草拟，包括封面页、内页、主编简介、序言、目录、设计理念/数据集说明、纳排、附录：症状对照表、参考文献等。

**封面页：**本数据集作为出版物的基础信息部分，主要包含中英文书名、主编、主审、组织编写机构及出版社信息。

**内页**：本数据集作为出版物的进一步出版信息，主要包含书名、主编、主审、组织编写机构及编委信息，次页记录数据集设计的专家委员会具体信息。

**主编简介**：协助编委整理的本数据集主编简介及相关照片。

**序言**：由主编撰写，内容简明扼要、重点突出，主要包含编写的背景、指导思想、意图；图书主要内容、特点；图书编写的原则、方法、过程；图书读者对象、适用范围；致谢、自谦；作者署名及撰写时间

**目录**:本数据集参考同一使用类型的数据元进行归类，并设计整理的内容清单，主要体现数据集的结构层次及主要内容。

**设计理念/数据集说明**:

为了达到兼具实用性、通用性与创新性的制作的目的，数据集设计分别从医学维度与数据维度进行了考虑。旨在能贴合临床使用习惯，数据以统一化，标准化方式导出，更好的用于科研，教学及临床的使用。

从医学维度，本数据集参考国内外权威书籍、传染病领域诊疗指南及高影响因子或高引用率文献，同时结合近5年来国内外高质量研究所涉及指标并统合专家意见，依照疾病研究热点与科研主流方向综合收录符合临床实用需求的指标。

从数据维度设计，指标名称、编码、格式和结构等的标准化，不仅能促进不同机构、不同研究及不同地点之间的数据交换和整合，也会为数据使用者提供极大的便捷性。本数据集的构建遵循国家卫生行业统一标准，采用数据完整性标准、临床数据交换标准协会、国际疾病分类等行业标准规范，使其满足临床科研需求，有助于提高临床科研质量及效率，以实现标准化、结构化、规范化为最终目标，旨在构建专业领域中一个开放、共享、可扩展的标准数据集。

**纳排**:本数据集针对尽可能多的研究项目以及临床应用，同时考虑尽量减少无关数据给数据应用时带来的数据清洗工作量，将符合ICD10编码的相关诊断且诊断明确作为标准数据集纳排标准。

**附录：症状对照表**:儿童感染性疾病症候群标准数据集提供落地过程中临床表征变量提取时可能需要的症状对照表。

**参考文献**：儿童感染性疾病症候群标准数据集提供数据集中使用的中英文缩略词表、症状对照表及参考文献列表。

**2.4.2.交付件专家组(KOL)审核及修改**

发表版本内容给到专家组(KOL)进行多轮沟通内审及修改，最终输出定稿交付件.

**2.5.出版社审核及修订**

**2.5.1.出版物常规内容草拟**

配合医院负责人针对出版社提出的要求，进行内容或格式等修订、PDF版本排版及美化工作，直至数据集正式发表。

1. 质保及报价要求

1. 服务期限：合同签署后6个月内；

2.报价格式自拟，并加盖公章；

3.报价须为人民币报价，包含：服务费、车旅费、运输费（含装卸费）、保险费、安装调试费、医院信息系统对接接口改造费、税费、培训费、专用维修设备的使用费等所有费用。