**海南医学院第二附属医院**

**围术期科研数据系统用户需求书**

说明：《用户需求书》中标注有“★”号的条款为关键需求，需要投标公司实质性响应；“▲”号的条款为重要技术要求，响应程度将影响技术得分。

# 项目背景

我国医疗卫生事业发展迅速，服务能力和服务水平显著提升，但仍然面临一些突出问题。例如：科研、预防保健等方面滞后；常见病、多发病及疑难复杂疾病诊疗指南的开发与推广不足，缺乏基于中国人群特征的大规模流行病学数据及高质量临床研究结果，为疾病诊疗带来安全隐患。当前许多疾病在很大程度上仍然是经验性学科，而现代医学的发展要求运用循证医学的思想来进行研究。开展循证医学实践的一个不可缺少的重要组成部分是收集研究证据，而这些研究证据主要来源于临床病例资料。

海南医学院第二附属医院积累了大量的临床病例数据，对于这些繁多零碎的数据资料，传统的手工方式进行检索与统计繁琐费时、代价昂贵、效率低下，而且容易造成部分数据丢失，甚至产生错误数据。临床科研实践中需要一种简便、快捷、高效的方法来对数据资料进行储存、整理，并从中分析、提取有用的信息。自从上世纪八九十年代以来，国内外已有部分医院、研究中心开始进行疾病数据库建设。但由于样本量较大，整理、检索不方便，这些数据资料一直没有被充分利用。当前大数据、人工智能等新兴信息技术的迅猛发展为该问题的解决提供了机会和可能。

目前，麻醉科在临床研究的开展与实施过程中存在一些困难，影响着科研质量的整体提升。主要体现在：麻醉科临床研究涉及到的数据繁多，涉及多个业务系统，提取针对性的数据难以实现统一化和标准化，导致围术期临床研究数据利用率不高。

# 需求分析

## 2.1业务需求分析

此次项目建设以麻醉科科研难点为驱动，以需求为导向，充分发挥海南医学院第二附属医院麻醉科的基础优势，搭建一个适合麻醉科临床研究需要的系统化、规模化、集成化的研究系统，结合临床科研成果集成创新，促进学科、专科建设发展。

通过建设围术期科研数据系统，充分利用海南医学院第二附属医院丰富的麻醉与手术数据资源，提升麻醉科在医疗质控、临床科研和教学等方面的综合能力，推动麻醉学科发展。通过接入特定疾病患者诊疗、治疗数据，并根据该疾病发生发展规律、诊疗规律，进行疾病数据模型构建以及针对性的数据清洗、标准化、归一化、后结构化处理，形成统一规范的科研数据仓库。在此基础上构建临床科室需的所科研高级应用功能，通过权限配置支持全科科研人员使用，用于循证医学上的大样本量临床科学研究，提高临床科研生产力。同时，在数据数据采集与处理的过程中，根据课题的要求整合科研课题研究所需的临床病例数据，提供临床科研课题所需要的多学科协作的医学智能服务，包括如：课题方案设计咨询、辅助数据采集、辅助数据清洗处理、辅助统计分析等。

## 2.2功能需求分析

基于麻醉科临床科研建设要求，在遵循相关标准与规范的基础上，对手术患者临床诊疗信息，进行采集，以患者为中心，对数据进行融合、治理、存储与共享,形成围术期科研数据系统。覆盖的数据范围包括患者人口学信息、诊疗过程信息、检验信息、检查信息、手术信息、护理信息、随访信息等术前、术中、术后资料。覆盖的时间轴线从患者入院到行麻醉手术治疗，再转回病房康复出院的整个过程。基于患者疾病的发生发展规律及诊疗规律，对数据进行深度挖掘、结构化处理。

梳理围术期术语规范，建设结构化模板及数据采集系统，支持开展相关研究，支持建立符合国际规范的临床研究质量管理体系。基于围术期科研数据应用系统，实现一站式科研。功能包括：数据概览，队列管理，快速纳排，特征值分析与导出、病例对照研究在线分析、队列研究在线分析等。可方便临床研究者高效地进行复杂队列发现，可根据研究课题的要求导出需要的数据格式，支持多维度数据搜索，满足各种复杂条件联合检索功能，从而实现研究队列人群的快速发现及特征值的导出。

以2个实际临床研究课题为导向，基于多学科背景团队，提供从包括前期的研究方向评估指导，到课题开展需要的各方资源（研究数据提取、研究数据统计分析支持、机器学习专业知识辅导、IT应用研发）的支持，形成示范效应，带动科室临床研究的开展。

## 2.3系统对接需求

本项目需与HIS、LIS、PACS、EMR、手麻系统等进行对接，按照科研需要定时获取患者的病历、医嘱、体检、检验检查等信息。针对HIS、LIS、PACS、EMR、手麻、病理、心电、护理系统数据，需要从海南医学院第二附属医院各业务系统接入历年及增量患者诊疗数据，历史数据采取一次性主动ETL抽取的方式接入，增量数据根据应用需求，按一定频率进行增量更新。基础数据对接需求如下表所示：

|  |  |  |
| --- | --- | --- |
| **序号** | **基础业务系统** | **主要数据接入范围** |
| 1 | HIS | 门急诊病人登记表、挂号信息表、就诊表、诊断表、处方表、用药表、住院病人登记表、诊断表、医嘱表、用药表、转科表 |
| 2 | EMR | 病历文书、病案首页（包括病案诊断、病案手术等） |
| 3 | LIS | 检验申请表、检验报告主表，检验报告明细表，微生物(鉴定和药敏) |
| 4 | PACS | 检查申请表、超声B超报告、彩超报告、放射报告、内镜报告等 |
| 5 | 病理系统 | 病理诊断、病理报告等 |
| 6 | 心电系统 | 心电报告 |
| 7 | 护理系统 | 生命体征记录 |
| 8 | 手麻重症系统 | 病人手术记录、手术登记、麻醉事件、术中体征、麻醉记录、麻醉总结 |

## 2.4硬件资源需求

为满足本项目建设需求，项目配置如下硬件资源，如下表所示。

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **序号** | **硬件类型** | **数量** | **配置信息** | **备注** |
| 1 | 数据库服务器 | 1 | CPU：核心数≥40/内存：256G /硬盘： SAS 1.2T\*4/RAID：H730/网卡：1G网口\*4/Windows 2012企业版 64位， SQL Server 2016 企业版64位 | 用于部署数据仓库及应用系统 |

# 建设原则

（1）整体性原则

从医院的现实需求出发，统一规划、统一设计。围术期科研数据系统架构采用整体化的设计理念，可以覆盖业务系统数据的完整接入，并可进行逐一数据的校验服务以及对历史数据的关联性存储、数据利用、应用开发。可按需扩展和支撑第三方应用系统的数据利用，覆盖每个系统。

（2）支持原有业务信息系统

充分发挥已有系统的功能，利用现有的数据库，通过围术期科研数据系统提供数据集成和业务价值的增值。不仅能够保存业务系统的原有历史数据，而且能满足接入和管理现有及未来的数据增长的需求。

（3）安全性

具备全面的安全性设计，防止对数据的非法访问、破坏和泄露，如权限分级管理、查询系统操作日志监控、修改留痕管理等。同时应能够提供对数据保护，确保数据的完整性、保密性、不可抵赖性、一致性、可追溯性。由于围术期科研数据系统将管理医院相关患者数据，系统架构须提供严格、严谨和完善的安全和隐私管理策略和服务。

（4）标准化

围术期科研数据系统建设基于医疗行业标准，并支持面向SOA服务架构建设。所采用的医疗行业标准具备国际领先性，并与医院实际数据情况相结合，建立医院标准化的数据管理规范，提供相应的标准化组件功能实现标准化管理服务，这将节约持续发展的成本，简化持续发展的复杂性，使得复杂的医疗数据管理和整合成本降低、改善数据整体利用效果。

（5）扩展性

系统架构要求不仅可以满足目前科研的数据利用需要，而且需要为将来新的数据利用需求提供开放性。医疗科研的需求是随着时间的不断变化而动态改变，因此系统整体架构需要提供针对不同业务领域的可扩展性功能，不能单单仅满足目前需求，更需要满足未来新业务信息系统的接入。

（6）高效性

具备高性能大密度的计算、存储、并发处理能力。系统性能响应时间合理，完全满足业务需要，数据查询和分析和业务处理要求速度快、反应及时。需要保证系统的稳定可用性。

（7）灵活性

系统应采用模块化的设计，应满足目前医院的业务实际需求，也能充分考虑将来业务种类增长后的需求。便于医疗服务业务的变化和扩展，提供丰富的开发环境，支持多种应用的开发，能够实现跨系统数据库间的数据管理，具体良好的第三方软件支持功能，提供用户应用层的组件支持，如配置数据接入工具满足业务系统的更换需求，让用户可以自行使用配置工具进行数据接入。

总之，围术期科研数据系统设计与建设应充分考虑整合和利用现有各种资源，既要考虑系统的先进性和有效性，又要考虑其实现的可行性和经济性。系统的设计与建设要符合业务需求，系统结构、数据接口应可灵活扩展，为将来系统拓展提供空间。

# 建设目标

围术期项目在海南医学院第二附属医院麻醉科开展大数据系统建设与医学智能服务建设，在临床数据集成、自然语言处理、深度学习等多方面协助临床研究人员，使得临床人员在数据获取、数据挖掘、临床转化等多方面得到全面的提升。从而有力提升麻醉科数据科学和临床研究的水平，通过人工智能解决医疗问题、临床研究和工作中遇到的“痛点”，改善疾病的检测、诊断、治疗和预后管理，最终提高患者满意度。

通过建设麻醉科数据库使之具有：支持疾病相关的数据模型创建能力；快速集成多样数据源的能力；支持数据增量更新的能力；强力的预测分析能力；语义的文本挖掘能力。打通各业务系统之间的壁垒、建立数据标准体系，“唤醒”医院“沉睡”多年的宝贵临床数据，充分发挥数据的整合利用价值。并以此为基础，在科研创新方面，有针对性地进行大样本做临床回顾性研究。

支持开展创新的前瞻性临床研究，有效利用医院系统积累的大量宝贵的病例数据和院外病人康复和预后结局的数据，大幅度促进临床科研和患者健康服务的发展，改变过去以“单打独斗”为主的科研方式，朝着以区域性医疗机构为主体、以网络为依托，以聚焦慢性并、多发病、重大疾病，搭建高水平的临床医学研究系统。

# 采购清单

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **序号** | **产品名称** | **数量** | **单位** | **备注** |
| 1 | 围术期科研数据系统 | 1 | 套 | 含2个课题辅助分析服务。含验收后1年免费维护期。详见“6.1围术期科研数据系统技术参数要求”。 |
| 2 | 数据库服务器 | 1 | 套 | 3年维保服务，详见“6.2服务器技术参数要求”。 |
| 3 | 与院内业务系统数据接口 | 1 | 套 | 3年维保服务，详见“6.3与院内业务系统数据接口要求”。 |

# 技术要求

## 6.1临床数据集成与预处理

临床数据集成与预处理，主要实现手术患者相关的历史诊疗数据集成与融合处理，实现病人增量数据自动集成与更新。临床数据集成主要步骤包括数据源分析、数据仓库模型构建、数据接入（历史数据、增量数据）、数据清洗与整合等步骤。

### 6.1.1数据源分析

根据麻醉科科室特点、疾病特点构造学科数据仓库，科研数据库支持汇聚临床过程中产生的所有诊疗数据，以及其他相关数据。数据源分析主要是对科研数据库拟纳入的相关诊疗数据的来源、数据量、数据时间跨度、数据范围、数据质量、存储方式、数据之间的关联关系等相关属性进行分析，以供设计适宜的数据接入方案，并接入数据至科研数据仓库。

数据源主要分为院内电子化数据、院内非电子化数据、其他相关数据三大类。

#### 6.1.1.1院内电子化数据源分析

院内业务系统数据支持从各业务系统接入，通过对业务系统数据源进行分析，准确地查找和定位到所需的源数据表和视图，并分析数据之间的关联关系。

#### 6.1.1.2院内非电子化数据源分析

院内非电子化数据源主要包括记录有各重点专科科研关注的纸质文档、图片、视频、音频等信息。该类信息后期由临床医生根据具体研究课题数据应用需求，组织人员对特定信息进行针对性的抄录，形成电子化信息并接入数据库。

#### 6.1.1.3其他相关数据源分析

主要包括数据库未部署在院内的系统，以及来自第三方的数据等，根据临床研究需要，按需接入。

### 6.1.2数据接入

#### 6.1.2.1数据接入方式

支持业务数据自动接入，支持按固定格式导入或数据拷贝的方式接入。将所需数据从各类来源接入到围术期科研数据仓库中，院内业务系统数据可分历史数据接入、增量数据接入两个步骤进行。历史数据接入可根据医院应用实际，通过备份库、备份文件等多种方式接入（期间一般无须业务系统厂商配合），增量数据接入支持视图，WebService等多种接入方式。

#### 6.1.2.2数据接入范围

接入的数据范围包括但不限于HIS、EMR、LIS、PACS、病案、护理、手麻计划、手术主记录、手术术后并发症及处置、麻醉事件记录、麻醉总结、术前评估信息、插管到手术开始术中体征、手术开始半小时术中体征、手术结束到拔管术中体征、术中血压、术后急性肾损伤、术后心肌损伤、术后1天VAS量表、术后7天VAS量表、术后7天EQ-5D量表、术后7天MMSE量表、全手术期术中体征、血气分析等数据。

#### 6.1.2.3数据接入过程

##### 6.1.2.3.1确定入库规则

确定入库规则是指确定科研数据仓库中拟纳入的患者人群范围，即设置一定的筛选与排除条件，定义接入数据仓库的患者人群范围。

##### 6.1.2.3.2历史数据接入

历史数据指系统实施之前所产生的业务数据，历史数据采用ETL工具，一次性进行批量接入。在数据接入的过程中，调用患者主索引系统（EMPI）相关服务，获得入库人群的所有诊疗记录。

##### 6.1.2.3.3增量数据接入

增量数据接入指系统部署实施之后，新增加的诊疗数据，增量数据接入采用增量机制，通过定制化的数据治理工具实现自动异步更新。如何捕获变化的数据是增量抽取的关键，根据源数据的设计结构以及各主题数据内容的大小、更新频率等特点，可选择适配的捕获变更数据的方法，包括但不限于触发器、时间戳、全表比对、日志对比（CDC）等。

##### 6.1.2.3.4离线设备数据接入

离线设备数据接入指单机系统数据的接入，例如：呼吸机数据的接入。离线设备数据的接入采用原则上采用周期性导入的方法实现。

##### 6.1.2.3.5缺失数据补录

缺失数据补录指业务系统中未记录的，需要补充采集的相关数据，缺失数据一般通过系统补录操作界面或数据后台手动实现数据接入。

##### 6.1.2.3.6其他相关的第三方数据接入

其他相关的第三方数据接入，根据实际情况进行数据手工录入或导入实现数据接入。

### 6.1.3数据清洗

从数据源中抽取的数据不一定完全满足目的库的要求，例如数据格式的不一致、数据输入错误、数据不完整等，因此有必要对抽取出的数据进行数据清洗。

#### 6.1.3.1重复数据清洗

为了减少数据中冗余信息，首先对一定范围内数据进行排序算法，根据预定义的重复标识规则进行重复检测，最后完成重复数据的清洗工作。

#### 6.1.3.2不完整数据清洗

首先对数据表结构字段按照重要性等级进行判断，按照缺失比例和字段重要性制定方案，对于不完整的记录并且不需要的字段进行删除处理，对于重要字段并缺失的记录，通过计算分析对字段进行填充。

#### 6.1.3.3数据格式清洗

数据格式的清洗主要针对由人工搜集或用户填写的信息，对不符合规定的格式及内容进行清洗。主要包括时间、日期、数值等显示的格式，内容中不合理的字符等内容。

#### 6.1.3.4错误数据清洗

基于课题研究需要，用统计分析的方法识别错误值或异常值，如偏差分析、识别不遵守分布或回归方程的值，也可以用简单规则库检查数据值，或使用不同属性间的约束、外部的数据来检测和清理数据。

#### 6.1.3.5关联性数据清洗

当数据记录有多个来源时，需要进行关联性验证，如果在数据分析过程中发现数据之间互相矛盾，需要将关联性的相关数据进行调整或去除，通过对数据的分析检测，从而使得数据保持一致。

### 6.1.4数据整合

根据疾病的演变和诊疗过程，结合专业特点来综合分类,参考HL7 V3、RIM模型、CCR等标准来进行灵活构建模型，同种业务类型的数据进行归类标准化汇集。

#### 6.1.4.1标准化归一化

统一编码、统一字典、数据格式（时间、日期、数值、全半角）确定。

#### 6.1.4.2重复数据处理

处理数据重复,数据属性值不对,数据不一致等情况，规范标准化的制定；

#### 6.1.4.3数据冗余处理

按照数据流向和逻辑关系进行整合，确保数据的一致性和完整性。科室、病区、诊疗项目等关键业务编码的统一。

#### 6.1.4.4异构异质数据源的整合

充分考虑到数据源格式的多样性，比如数据库文件格式，文本文件格式，XML、JSON格式文件，结构化数据和非结构化数据的整合，标准化。

## 6.2数据后结构化

将系统文档中难以查询的、半结构化、非结构化的数据（如病史、手术记录、病理报告、影像报告等），抽取和分解为易查询的、结构化的、具有医学意义的数据。使得临床文本数据处理可以在一个非常精确的水平，进而使得医院和医生从这些数据中获得真正的临床价值和科研价值。

### 6.2.1入院记录后结构化

将非结构化的入院记录转换为结构化指标数据。

### 6.2.2出院记录后结构化

将非结构化的出院记录转换为结构化指标数据。

### 6.2.3既往史后结构化

将非结构化的既往史转换为结构化指标数据。

### 6.2.4个人史后结构化

将非结构化的个人史转换为结构化指标数据。

### 6.2.5家族史后结构化

将非结构化的家族史转换为结构化指标数据。

### 6.2.6体格检查后结构化

将非结构化的体格检查转换为结构化指标数据。

### 6.2.7术中事件后结构化

提取非结构化中的关键指标数据，按规则生成术中事件标签。

### 6.2.8术后结局后结构化

提取非结构化中的关键指标数据，按规则生成术后结局标签。

## 6.3患者主索引

储存各个业务系统的内部唯一患者标识，并且生成一个供专病数据库使用的唯一患者标识，将多个医疗信息系统查看同一个病人时能有效地关联在一起。

## 6.4数据质量核查

### 6.4.1信息化角度核查

通过数据关联性、唯一性等信息化角度进行数据质量分析、数据核查、出具数据质量报告。

### 6.4.2医学角度核查

基于医学知识角度进行数据质量分析，出具数据质量报告。如数据的值域范围核查、数据的业务逻辑核查。

## 6.5数据管理

数据管理维护主要包括：数据权限管理、数据更新管理、数据维护和备份等。

### 6.5.1数据权限管理

数据权限管理需考虑系统功能的使用权限、数据库的使用权限、数据文件的使用权限等方面。

数据库系统的权限管理采用了基于角色的访问控制方法。它的特点是先确定角色对服务所拥有的权限，然后将用户注册到角色中，或者说授予用户适当的角色，从而获得调用服务的权力。

### 6.5.2数据更新管理

数据的更新是指数据库存储信息的变动，更新操作包括：插入数据、修改数据和删除数据等。数据更新管理主要解决以下问题：

（1）基础信息及其相关核心数据同步更新

在系统运行过程中，及时更新变化的数据，使数据能够及时准确的表示信息状态。

（2）业务数据和分析型数据的定期更新

分析型数据与其来源的业务数据是紧密关联的，是业务数据加工提取而形成的。基础信息数据变更后，为保证分析数据的完整性、准确性，业务数据、分析数据必须定期更新。

（3）基础数据和专题数据集市的同步更新

专题数据集市从内容上是基础数据的一个子集，分析应用要求无论是存储于基础数据、还是存储于专题数据的数据，统一指标内容的数据必须保持全局的一致性。

（4）错误数据追加更新

在数据采集交换过程中经常出现由于业务上的不可抗拒的原因导致将错误的数据入库，发现错误需要重新修改。

### 6.5.3数据维护和备份

数据维护和备份策略主要包括下列工作：

（1）备份数据信息收集；

（2）定制数据备份策略及定期备份数据；

（3）收集正确的编目信息。

## 6.6数据模型构建

根据专科专病的诊疗规范以及相关疾病的发展规律，参考疾病相关信息学语义模型，构建面向疾病的数据模型，使数据的建模从传统的面向事件活动的方法演变到真正的面向疾病的建模，为后期的数据挖掘提供了数据保障。整体上是以电子病历为核心，医疗过成中产生的所有文档记录包括申请、报告、执行记录、知情同意、医疗和护理文书等看作是电子病历的一个组成部分。一个病人的电子病历由其历次的就诊记录构成，每一次就诊记录由本次发生的所有医疗和护理文档组成。

数据仓库模型可分为基础数据模型及专科数据模型两部分。基础数据模型涉及各专科临床科研大多数时候都需要采集的数据范围，主要包括就诊、医嘱及临床路径、用药、医疗观察、手术及治疗、预约、代码系统等几个方面。专科专病数据模型需要根据专科专病领域诊疗指南、相关文献以及我院临床实际情况进行设计，需要临床专科领域专家参与该项工作。基础数据模型构建如下：

### 6.6.1就诊

就诊信息描述了病人就诊的基本信息，包括门急诊和住院。就诊信息主要由病人基本信息，就诊信息两部分组成。就诊是整个信息模型的基础，病人在院发生的所有医疗活动都将以此为基础记录，通过就诊记录可用直接查找到病人的本次就诊的所有医疗记录内容。

病人信息描述了病人的基本信息，包括人口统计学信息，个体标识，危险因素，联系人，邮政及电信地址，在院注册内容等信息。其中个人信息是病人作为一个社会实体的基本信息，病人在院注册信息描述的是病人在本院登记的与就诊相关的信息，一个病人可以产生多次就诊记录。

就诊信息描述了病人一次就诊情况，包括就诊场所，接诊医疗人员，转科信息，床位信息，医保及资金账户等内容。与病人信息比较，就诊信息是每次就诊病人在院登记信息。

### 6.6.2医嘱及临床路径

医嘱及临床路径模型描述了医师在医疗活动中下达的医学指令，具体包括两部分内容即医嘱和临床路径。

医嘱是医生在医疗活动中直接下达的医学指令,包括的内容有完整的医嘱信息，相关的各种申请，医嘱执行记录，医嘱产生的费用等。

临床路径是医院的一种遵循询证医学的标准化治疗模式与治疗程序，临床路径的最终执行主要落实到医嘱和其他一些护理记录上。模型中描述的信息主要包括临床路径的组成，应用记录，对应的医嘱等。

### 6.6.3用药

用药模型主要描述病人就诊中的用药信息，包括用药申请和用药记录两部分内容。用药模型以药品信息为基础，申请和记录都不再存储额外的药品信息内容。

用药申请描述医嘱和处方等相关的用药信息，其描述的主要是医生下达的医疗制定涉及药品的部分，下达医疗指令相关的各种临床诊疗结果作为参考，同时对于不规则频率的用药也有具体的描述。

用药记录描述药品的使用记录，对应为医嘱和处方的执行记录信息。

### 6.6.4医疗观察

医疗观察模型描述临床上可测量的医疗活动，描述其申请和报告结果。具体包括检查检验的申请和报告，还包括诊断、过敏反应等信息。

观察申请主要指检查申请和检验申请，其组成部分包括申请需要描述的临床表现，严重程度，参照标准等内容。观察申请通过自身的引用实现观察项目的具体描述，同时对于住院病人观察申请会产生相应的医嘱。检验申请还包括样本内容。

观察报告检查检验的报告。观察报告包含临床表现，严重程度，参考标准等报告需要的具体内容，检验报告还包括样本和使用试剂的内容。观察报告和申请一样通过自身的引用实现报告项目的具体描述，观察报告与观察申请存在对照关系。

诊断和过敏反应作为一个单独的对象描述了临床诊断和过敏反应的内容，其包括一些备注信息的描述。

### 6.6.5手术及治疗

手术治疗描述的是临床上的医疗操作，模型中将医疗操作分为手术和治疗两部分，它们包含的信息项基本一致，所以采用相同的模型来描述。

手术治疗模型同观察申请一样也包括申请和治疗结果记录两部分。手术治疗申请指医生开出的治疗类申请单，具体内容包括相关的诊断，观察报告，申请人信息，临床诊断，相关材料申请等，申请对象主要记录操作的部位、方法、时间等。手术治疗申请的引用实现申请项目的详细描述。住院病人的手术治疗申请会产生对应的医嘱。

手术治疗事件描述医疗操作的结果，具体包括执行人，相关的系列结果，附加的医疗观察描述等内容。手术治疗事件通过自身的引用实现各结果间的前后顺序和相互关系，它与手术治疗申请间存在对照关系。

### 6.6.6预约

预约模型描述的是在日常的医疗活动中对医院各种资源，包括人、场地、物资、设备的等的预约申请及安排。该模型可以涵盖整个医疗过程的各种医疗活动比如检查项目、检验项目、手术治疗、就诊等等。

预约模型的核心在于将医疗资源的使用划分成为不同的时间片，然后有制定资源的维护人员根据医疗人员的申请合理安排，最佳地保持整个医院资源的最大使用率。

预约模型不仅支持对单个资源的预约，而且支持对多种资源的综合预约，预约的时间也支持不同的时间点，从而将一个病人的一次复杂的多资源预约系统地管理起来，实现整个预约数据的统一描述。

### 6.6.7代码系统

代码系统描述模型所用到的代码表结构。代码表由代码分类和基本代码表组成，代码分类标识代码表内容的分类如疾病、性别、国家等，基本代码表包含了代码表的基本属性如编码，名称，描述，拼音码，自定义码等。

不同的代码表往往具有自己特定的属性，所以对于有特定属性的代码表可以从基础代码表派生出来增加自己的属性。采用统一的代码系统有力于代码的统一管理，同时也为模型中的引用提供了便利。

## 6.7数据洞察

基于主流的可视化技术，展示数据系统数据整体情况，让用户对数据整体情况有初步的认知。

### 6.7.1数据的年份

展示纳入数据库的患者数据年份跨度及各年份数据量、患者数分布情况。

### 6.7.2病例数总量

展示纳入数据库的病例数总量及各年份分布情况。

### 6.7.3患者地域分布

通过地图直观地展示患者来源的地域分布情况。

### 6.7.4麻醉等级

通过统计图表展示患者麻醉等级分布情况。

### 6.7.5手术类型

通过统计图表展示患者手术类型分布情况。

### 6.7.6手术科室

通过统计图表展示各手术科室患者数量。

### ▲6.7.7词云、诊断分层

通过词云、分层图等展示用药、手术名称、诊断名称等分层信息。

## 6.8数据补录与校对

### 6.8.1数据标准化管理

根据业务信息，制定标准化术语库，不同数据源、数据形式，进行归一化及标准化处理。

### 6.8.2病例信息概览

支持按病程查看病例信息，支持基于患者就诊时间轴查看患者全息视图。

### ▲6.8.3病例信息填写质控

支持病例信息填写时质控稽查，如自动提醒填写范围等。

## 6.9特征提取

### 6.9.1数据概览

实现诊疗数据的查询和任意字段的筛选，包括病人基本信息、就诊记录、诊断、手术、检验、检查、用药、病理结果、病历文书、体征记录、手麻重症数据记录等视图，每个视图均需实现列设置、列统计功能以及数据导出功能。

### 6.9.2队列定义

提供灵活方便的组合条件查询筛选界面，以及多维度的诊疗数据检索，可以快速地找到科研病例集，再通过入选和排斥等集合运算实现研究人群的精确筛选。

### 6.9.3队列患者

集中展示已纳入队列的患者基本信息，包括如末次就诊日期、首次就诊日期、末次诊断描述、末次入院日期等。

### ▲6.9.4队列概览

通过可视化图表，展示队列人群性别分布、年龄分布、诊断排行、地域分布、手术词云、用药词云、诊断词云等。

### ▲6.9.5队列对比

通过可视化图表，直观地进行不同队列各维度指标内容的对比，如性别分布对比、诊断排行对比等。

### 6.9.6队列数据

展示入组队列数据详情，实现诊疗数据的查询和任意字段的筛选，支持列设置、列统计功能以及数据导出功能。

### 6.9.7指标提取

支持分析指标自定义、支持分析指标按照不同方法聚合后的提取；支持数据计算以及连续变量的离散化处理。

##### ▲6.9.8.1连续变量-时间段时长合计：

提取指定时间段内所需指标在特定阈值范围的总时长，如手术期间平均动脉压低于65mmHg的总时长。

##### ▲6.9.8.2连续变量-时间段时长拼接

提取指定时间段内所需指标在特定阈值范围内的时长拼接，如手术期间平均动脉压低于65mmHg的时长拼接。

##### ▲6.9.8.3连续变量-时间段计次

提取指定时间段内所需指标在特定阈值范围内的出现次数，如手术期间平均动脉压低于65mmHg的记录次数。

##### ▲6.9.8.4连续变量-时间段最大

提取指定时间段内所需指标在特定阈值范围内的出现的最大持续时间，如手术期间平均动脉压低于65mmHg的最大持续时间。

### ▲6.9.9自定义结局

支持患者心损伤、肾损伤、术后感染、术后脑卒中等常见围术期结局自定义、支持结局指标提取。

### 6.9.10队列模板

支持预定义特征提取范围，实现快速队列发现，支持模板管理。

## 6.10数据分析

### 6.10.1分析向导功能

支持向导式方式实现病例对照研究及队列研究。

### ▲6.10.2病例对照研究向导

病例对照研究数据分析步骤应包括：项目信息、导入数据集、多重插补、研究人群描述、回归建模五个步骤。

### ▲6.10.3队列研究向导

队列研究数据分析步骤应包括：项目信息、导入数据集、多重插补、队列匹配、研究人群描述、回归建模六个步骤。

### ▲6.10.4数据治理与插补

支持对特征提取后的二维表数据进行数据类型列验证、在线编辑修改、数据R值类型定义、实现数据的在线清洗和规整功能，如支持分类型数据的数据转换(将“吸烟”、“戒烟”赋值为“吸烟”）、支持数值型数据的数据赋值（将年龄“<=18”的标识为“0”，将年龄“>18”的标识为“1”)。且内嵌缺失值多重插补方法，方便后续分析研究。

### 6.10.5研究人群描述

支持对连续型变量依据正态分布的与否分别用均数（标准差）和中位数（四分位数）表示；对分类型变量，统计其在各组的频数与百分数。

### ▲6.10.6队列人群匹配

支持在队列研究中用倾向性评分匹配（PSM）的方法匹配病例组与对照组的基线特征（协变量），以消除这些协变量的混杂影响，在一定程度上达到随机分组的效果，提高研究可信度。

### ▲6.10.7多因素回归建模

支持针对不同的结局类型（生存期结局或二分类结局）分别用COX回归和logistic回归建立回归方程。队列研究中，可调整不同的协变量，观察危险因素（X）对结局变量（Y）的作用。它量化了在控制其它混杂因素后，危险因素（X）是否对结局变量（Y）有独立作用（p值），及其独立作用的大小（OR/HR值）。病例对照研究中可首先利用逐步回归进行危险因素初筛，后在此基础上用逻辑回归建立风险预测模型，并查看对应模型的效果，包括接受者操作特征（ROC）曲线、曲线下面积、及根据最大约登指数（Youden index）原则选择的最佳风险阈值。

### ▲6.10.8生存曲线

支持用Kaplan-Meier 法估计生存函数，在研究中比较病例组及对照组在特定时间间隔内事件（例如手术后1年内生存）的概率差异，并用Log-rankj检验比较两条生存曲线的差异显著性。

#### 6.10.9辅助课题研究服务

##### 6.10.9.1整体要求

（1）协助提供2个课题医学统计分析服务，协助为论文的“材料和方法”(Materials and Methods)部分撰写数据来源(Data Source)、数据收集方法(Data Collection)、统计分析方法(Statistical Analysis)；协助为论文的“结果”(Results) 部分制作统计结果图表。

（2）需要具有医学、统计学、IT工程师等多学科团队，要求通过配置以数据驱动的端对端的临床科研服务，提供以科研成果为导向的整体解决方案，向科研医生提供完整体系和最佳实践，进行科学的设计、测量与评价。

（3）为临床医生提供智能化全方位全链条本地化科研服务，助力医院学科建设，科研队伍建设，快速实现科研成果的产出与转化应用。

##### 6.10.9.2服务流程要求

（1）步骤一：项目立项

作为临床专家，首先根据本学科国内外最新进展、个人兴趣和工作经历、学科共同关心的科学问题，提出课题研究目的、构想和假说，完成临床研究项目合作意向表。

（2）步骤二：制定研究方案

医学研究者和医学团队专家讨论课题细节，结合相关文献，制定研究方案(Study Protocol)。

（3）步骤三：数据管理和可行性试探

由数据管理师和统计师（需IT工程师配合）探寻项目的主要数据（例如暴露/非暴露变量、结局变量、协变量等等）是否可以获得、数据是否结构完整、是否已经清理、是否还需要返回到原始数据以添增新变量等等一系列可行性问题。此步骤的结果将决定该研究是否可以先前推进到第四步，或因无可行性而停止，或改换成其他课题（回到第一步或第二步）。

（4）步骤四：统计分析操作

按照统计分析计划SAP，在第三步的基础上深入分析，完成统计操作，输出初步统计结果。

（5）步骤五：统计分析结果核实

统计师与医生深入讨论结果和解释结果的可能临床意义。对基于结果提出的统计要求执行进一步的数据探索和挖掘，直接双方最终确认可发表结果。

（6）步骤六：统计相关资料输出与论文撰写

统计师将继续参与部分论文的英文写作工作：

第一、为论文的“材料和方法” (Materials and Methods)部分撰写数据来源(Data Source)、数据收集方法(Data Collection)、统计分析方法(Statistical Analysis)；

第二、为论文的“结果”(Results) 制作统计结果相关图表。

（7）步骤七：审稿后修改的统计服务

统计师将继续跟踪服务项目，针对投稿后的审稿人提出的意见和要求，执行并完成进一步的数据探索和挖掘，帮助更新统计方法学描述部分和统计结果图表部分。

##### 6.10.9.3服务内容要求

1. 研究启动准备
2. 医学方案撰写（协助）
3. 医学顾问等协助医生完善方案
4. 医学方案撰写（主导）
5. 提供完善的课题方案，并和医生讨论通过
6. 样本量计算
7. 统计课题入选的人群数量
8. 专病数据库建立
9. 入库规则制定（入排标准）
10. 确定数据库入库人群入选和排除标准
11. 临床数据集成
12. 集成多源异构的业务患者临床诊疗数据
13. 临床数据补录系统，非电子化数据的手工补录，电子表格数据的数据库导入
14. 专科系统数据接入（如：内镜、高压氧舱等）
15. 离线数据导入
16. 单机离线系统数据接入
17. 数据标准化
18. 基于业务标准，对非标准化数据进行标准化处理。
19. 数据验证
20. 对数据逻辑关系，数据质量进行验证。
21. 术语标准化
22. 患者主索引，生成病人唯一主索引，关联病人历次诊疗记录。
23. 疾病数据建模，根据特定疾病的发生发展规律、诊疗路径，对疾病数据的组织、存储进行数据模型设计。
24. 数据导出服务。
25. 文本挖掘
26. 将文档中难以查询的、非结构化的数据（如影像所得），抽取和分解为易查询的、结构化的、具有医学意义的数据，使得临床文本数据处理可以在一个非常精确的水平，进而使得医院和医生从这些数据中获得真正的临床价值和科研价值。
27. 病历文本结构化：根据科研需求，针对历史诊疗数据中的病历文书进行后结构化。
28. 检查报告结构化：根据科研需求，针对历史诊疗数据中的检查报告进行后结构化。
29. 超声报告结构化：根据科研需求，针对历史诊疗数据中的超声报告进行后结构化。
30. 病理报告结构化：根据科研需求，针对历史诊疗数据中的病理报告进行后结构化。
31. 增量数据结构化：根据科研需求，针对增量诊疗数据中进行自动化后结构化。
32. 结构化数据对比及修改
33. 结构化数据质量核查与后结构化完善。
34. 应用部署
35. 课题管理：对科研课题进行创建、修改、权限设置等管理。
36. 患者360视图：在同一界面，集中展示患者全息诊疗数据，并可对检验指标进行可视化对比展示。
37. 数据库可视化：通过可视化界面，基于特定疾病所关注的指标内容，对数据库数据情况进行概要展示。
38. 疾病数据集一览：集中展示疾病数据集。
39. 疾病数据集检索：提供疾病数据集检索功能。
40. 队列发现：根据特定疾病的多维度诊疗数据自定义查询条件，再通过入选和排斥等集合运算实现研究人群的精确筛选，并且可以把多维度查询条件保存，以便满足条件的新病人自动入组，进行前瞻性队列研究。
41. 队列可视化：通过入组的研究队列人群，建立队列人群的可视化，用户可直接点击查看队列人群数据基本情况，包含队列人群的基本情况，如年龄、性别、诊断、手术等分布情况。
42. 队列对比：通过选择不同的队列，进行性别、年龄、诊断分层等内容对比，支持图形化的形式展示，支持图形的合并与拆解。
43. 队列探索：支持针对特定疾病队列，进行探索性分析。
44. Timeline：完整掌握患者历次就诊信息.实现全流程的进行患者病情追溯。以历次诊疗事件时间轴为主线，集中展示历次诊疗过程中的医疗事件。可进一步深入展示历次具体医疗事件内容。
45. 智能搜索：采用与医学临床紧密结合的中文分词技术，使得对非结构化文本的分词，既快速又符合医学实际；该全文搜索的结果，也可以作为多维度病例挖掘的手段之一。
46. 特征提取：支持通过可视化界面，提取课题关注的指标数据，支持数据的算法运算、聚合运算。
47. 探索分析：支持对课题数据进行探索性分析，不断优化完善课题方案。
48. 权限管理：实现对应用系统权限进行管理。
49. 数据管理
50. 数据导出：自定义表的数据导出到Excel。
51. 数据规整：导出的数据做课题挖掘、赋值等。
52. 异常值、空值处理：针对数据质量问题，进行数据预处理。
53. 可行性分析：基于课题方案，对课题数据进行探索性分析，确认课题的可行性。
54. 增加新变量：可能增添其他变量。
55. 统计分析
56. 统计分析（常规）：分析课题数据，出具常规统计结果。
57. 模型优化（预测）：探索更优的统计结果。
58. 制作统计结果图表：为统计分析出图表。
59. 撰写统计报告：撰写统计报告。
60. 分析结果讨论：按照统计报告，向医生解释统计结果，并讨论下一步工作。
61. 文章撰写与发表咨询
62. 文章撰写咨询：提供文章撰写咨询，定期与医生讨论。
63. 文章润色协助：协助临床医生对论文统计部分进行润色。
64. 论文会诊：组织相关领域专家，讨论文章可能存在的问题，提出修改指导意见。
65. 附加：培训
66. R语言培训：针对0基础用户，培训R语言在科研数据分析与统计方面的应用。

## 6.11服务器技术参数要求

CPU：核心数≥40/内存：256G /硬盘： SAS 1.2T\*4/RAID：H730/网卡：1G网口\*4/Windows 2012企业版 64位， SQL Server 2016 企业版64位

## 6.12与院内业务系统数据接口要求

### 6.12.1接口要求概述

围术期科研数据系统数据需求涉及医院所有手术患者的术前、术中、术后资料。一般包括：

1. 术前关注患者的入院诊断、各大系统合并症和功能状态、实验相关的检验结果、医学影像和超声检查结果、围术期的手术治疗方法等。该部分数据主要通过结构化或者非结构化进行数据提取和分析。
2. 术中资料主要通过手术麻醉信息系统提取。
3. 术后数据资料覆盖病房医生开启医嘱、病程记录和护理相关记录、重症监护室的医护医嘱和病程记录资料及评分表格等。

整个专病库覆盖患者人口学信息、诊疗过程信息、检验信息、检查信息、手术信息、护理信息和病例信息等围术期信息。围术期科研数据系统数据覆盖的时间轴线从患者入院到行麻醉手术治疗，再转回病房康复出院的整个过程。

所有数据为围术期科研数据系统单向接收，不向业务系统回传数据。要求支持的数据接入方式包括：开放备份库只读权限；开放源库相关数据表只读权限（夜间数据抽取）；开放源库相关视图只读权限；通过WebService服务接口读取数据。

系统实施过程中产生的第三方接口费用由中标方承担。

## 6.13数据安全保障

在处理和使用医院临床数据时，需要严格遵守医院的各项规章制度，还需要从数据脱敏与加密、网络安全、数据传输及存储安全、访问控制、监管审计五方面保证数据的安全。

（1）数据脱敏与加密

数据脱敏和加密支持对患者个人信息中的敏感信息（如：患者姓名、身份证号、电话、地址等）通过脱敏或加密规则进行数据的变形，实现敏感隐私数据的可靠保护，同时依然保持其它数据的格式和属性，保证其可识别性和可用性。

（2）网络安全

围术期科研数据系统和科研数据分析应用系统物理位置在医院机房内，部署在医院的私有云环境，需要建立安全通道运维专病库系统的服务和应用，为医院提供24\*7的安全运维服务，保证专病系统及其应用的正常运行。

（3）数据传输及存储安全

支持大数据系统使用加密存储和传输，即处理后的重要数据先加密再存储，防止重要数据被非法窃取或窥探；在网络传输中采用SSL/TLS安全协议为数据传输提供安全保障，如采用HTTPS协议保障，在使用浏览器搜索查看病历全文时，保障用户身份认证安全和数据传输安全。

（4）数据访问控制

支持对数据的访问权限全面的控制，不同用户、不同业务系统运维人员或系统管理员的权限不同，登录系统看到的数据亦不同。

# 商务要求

1. **售后服务要求：**

投标人提供的软件建设部分不少于两年的保修服务、硬件建设部分不少于五年的保修服务。保修期和系统技术服务自双方代表在项目验收单（终验）上签字之日起计算。保修期内所发生的一切费用包括系统技术支持，系统维护或升级，人员交通，差旅服务等费用全部包含在中标价内。免费服务期满后，每年的维护费不得高于项目软件价格的8%。

1. 系统故障的响应时间：投标人必须提供7\*24小时技术支持热线电话（固话，手机）。当发生故障时，自报障时起算，1小时内响应，若远程维护无法解决，投标人应安排人员在48小时内到达现场处理故障；
2. 维保期内，投标人向采购人提供免费软件升级和维护服务；
3. 维保期内，若因为政策变动而导致系统功能需要重建，则双方共同协商制订费用收取标准；若因为政策变动而导致系统功能需要增加或者修改，则包含在保修业务范围内。
4. **培训要求**
5. 所提供的培训课程表随投标文件一起提交；
6. 培训内容与课程要求：对系统的使用，操作，维护进行培训，并提供安装使用维护说明书，以确保采购人能够对系统有足够的了解和熟悉，能够独立进行系统的日常维护和管理。培训所需一切资料由中标商提供；
7. 培训费用：培训过程中所发生的一切费用（含培训教材费）均包含在中标价内。
8. **工期要求及付款方式**
9. 工期：合同签订后90个工作日内完成开发，实施和验收（该时间不包括等待第三方接口、硬件设施部署等非项目实施范围时间）。
10. 付款方式：

（1）合同生效后15天内，凭中标人提供的正式含税发票，采购人办理相关手续后，支付合同总额的30%；

（2）项目通过整体验收合格后15天内，凭中标人提供的正式含税发票，采购人办理相关手续后，支付合同总额的65%；

（3）项目质量保证期届满后，凭中标人提供的正式含税发票，采购人办理相关手续后，支付合同总额的5%。

1. **验收要求**

1) 验收按国家，行业有关的规定，规范进行；

2) 系统测试要求：中标人要对系统实际运行情况进行测试；

3) 软件系统的验收属于项目的合同验收，需符合海南省信息化项目相关验收管理办法的要求，同时需符合下列要求：

a. 满足合同和招标文件中列举的全部要求；

b. 实现合同和招标文件中列举的全部功能要求；

c. 达到合同和招标文件中列举的全部指标；

d. 文档齐全，符合合同和招标文件及相关标准要求，包括但不限于下列文档：需求说明书，概要设计说明书，详细设计说明书，数据库设计说明书，测试计划，测试报告，用户手册，项目计划书，用户培训计划，会议记录等。

4) 验收项目包括按照合同和招标文件中所标明的软件系统，及相关的技术维护文档，培训教材，使用说明书等。