

# 交大一附院-驼人共建医疗器械研究院

## 研究申请书

项目名称： 基于 SurClaw 的微创外科医保质控与费控  
图文报告智能生成系统

项目负责人： 彭子洋

所在科室： 临床医学（医工学方向）

联系电话： 17809212035

邮箱： 1061201104@qq.com

（本页单面打印）

一、科室意见（本页单面打印）

**申请者承诺**

我保证本申请书填报内容的真实性。如果获得资助，我与本项目组成员将严格遵守医院临床研究课题管理办法的有关规定，切实保证研究工作时间，按计划认真开展研究工作，按时报送有关材料。

申请者（签章）：

2026 年 04 月 22 日

**科室意见**

本科室已审查该项目的科学性、创新性与伦理可行性。该项目紧密结合国家 DRG/DIP 医保支付方式改革的宏观战略需求，提出的基于多模态大模型的微创外科图文报告智能生成系统具有重大的临床转化与卫生经济学管理价值。科室具备完成该前瞻性多中心临床试验所需的临床病例资源、软硬件支撑平台及研究时间。同意申报。

科主任（签章）：

请科主任签名并加盖科室公章

年 月 日



### 三、研究方案

#### 1. 立项依据

##### 1.1 国内外研究现状与临床行业痛点

微创外科以其创伤小、恢复快等显著优势，已成为当前肝胆、胃肠、胸外及泌尿等实体肿瘤治疗的主流手术范式。微创手术记录作为反映手术全过程、术中关键解剖发现、特殊技术处置及高值耗材实际使用的法定医疗文书，是医疗质量控制、医保费用审核、科研数据沉淀以及医疗纠纷举证的核心法律基石。然而，一个长期存在且极为尖锐的矛盾是：尽管现代外科手术器械、内镜显示技术（如 4K/3D/荧光）与微创理念飞速迭代，但手术记录的形式与载体却长期停滞于“模板化纯文字”的原始阶段。这种高度依赖主观回忆、以文字为主体的记录模式在描述复杂的血管解剖变异、肿瘤浸润边界及三维空间手术逻辑时，存在不可避免的严重维度损失。这种损失导致海量珍贵的“手术实境”客观影像数据与最终生成的文本记录之间发生严重割裂，形成了当前医院信息化建设中难以打破的“信息孤岛”。

在卫生经济学与国家医保政策的宏观维度上，这一痛点显得尤为致命。当前，国家医疗保障局正以前所未有的力度推进 DRG/DIP（按疾病诊断相关分组/按病种分值付费）支付方式改革，明确提出至 2025 年底，DRG/DIP 支付方式将覆盖所有符合条件的开展住院服务的医疗机构，基本实现病种和医保基金的全覆盖。这一深刻的政策转型要求医疗机构由过去粗放式、规模扩张式的运营机制，强制转向注重内部成本控制、强调医疗服务技术价值的内涵式高质量发展阶段。然而，在实际临床与审计运行中，由于缺乏具有数字时间戳的客观影像证据链作为支撑，医保监管端及医院内部物价管理部门难以对术中高值耗材，如超声刀头、各型切割吻合器及配套钉仓、血管闭合夹等的实际消耗情况进行精准核查。据 2024 年医疗保障事业发展统计快报显示，2024 年全国医保基金总支出高达 2.97 万亿元，其中手术费用及耗材支出占据极大比例。近年来，国家及省级医保局已通报多起大型三甲医疗机构因涉嫌“虚构医疗服务”、“串换项目”或“虚记骨科高值医用耗材”而被查处的恶性案件，单起案件骗取或违规使用医保基金的金额高达数千万元，如华中科大同济医学院附属同济医院案、芜湖市第二人民医院案。单纯的文本记录无法自证清白，导致医院在面临严苛的医保飞行检查或复杂的医疗纠纷，例如胆总管医源性损伤导致的手术责任界定时，往往因“举证不能”而陷入极度被动的法律境地，最终承担巨额赔偿与行政处罚。在此背景下，一种能够自动生成防篡改“可视化凭证”的系统，成为赋能 DRG/DIP 精准支付审核、保障国家医保基金安全的行业刚性需求。

##### 1.2 前期研究结果与核心技术演进路径

为了从根本上改变这一落后现状，西安交通大学第一附属医院吕毅教授团队于 2015 年及 2025 年两次牵头，联合中华医学会外科学分会制定并发布了《倡用图文外科手术记录专家共识》，在国际上首次确立了“图文手术记录”的定义与行业规范，强烈推荐以文字描述为主体结合术中高清照片，使医疗记录能够客观、真实地反映手术实境。然而，早期的图文记录工作站高度依赖手术医生在术中手动脚踏截屏，并在术后耗费大量精力进行

人工回忆、图像筛选与文本拼接。生成一份符合规范的高质量图文报告往往需要耗时 15 至 30 分钟，这在日益繁忙的外科临床 workflow 中极大地增加了医师的文书负荷，导致“图文并茂”的理念难以实现规模化、常态化的临床普及。

随着计算机视觉与生成式人工智能技术的爆发，医疗 AI 经历了从传统卷积神经网络向基于 Transformer 架构的跨越。尽管 Transformer 显著提升了图像的全局上下文建模能力，但其核心的自注意力机制在数学原理上决定了其计算复杂度会随序列长度呈二次方指数级增长 ( $O(N^2)$ )。在处理微创外科动辄长达数小时、帧率高达 60FPS 的 4K/8K 超高清视频流时，Transformer 架构会面临严重的显存爆炸与实时计算瓶颈，根本无法满足微创手术室对于“手眼协调”所需的毫秒级（小于 50ms）低延迟严苛要求。



图 1 术者与手术图文记录系统合影

本项目团队在此前沿交叉领域取得了颠覆性的原创突破，成功研发出“SurClaw 智能图文报告生成系统”。该系统依托团队历时数年构建的、目前国内规模最大的 1.5PB 多模态手术影像与电子病历数据集，涵盖 692 例多中心患者，精细语义标注超 24 万帧，创新性地引入了目前国际最前沿的、具有线性计算复杂度 ( $O(N)$ ) 的 Mamba 状态空间模型架构。以此为基础开发了 STSH-Net 时空协同感知网络，使得边缘端异构计算设备能够在保持极低功耗与显存占用的同时，对超长时程的手术视频进行具有因果逻辑的全局记忆与实时阶段推理，处理帧率高达 52 FPS。此外，针对腹腔镜手术中因高频电刀切割组织产生的焦烟遮挡这一世界级行业公认的“卡脖子”难题，团队首创了 Yun-Trans 自适应去雾网络。该网络利用动态专家混合机制感知雾气浓度特征并动态调度子网，将重度烟雾极端视觉条件下的解剖结构识别准确率由传统的 35.66% 实现了跳跃式提升至 74.15%，确保了机器视觉感知的连续性与鲁棒性。在最为关键的自然语言报告生成端，团队采用了微调后的 DeepSeek 及 Qwen-VL 等多模态大语言模型，并深度融合了 RAG 检索增强生成机制。通过将实时的视觉特征与权威的 SAGES 等医学指南知识库在统一的语义空间中进行跨模态对齐，彻底从底层逻辑上抑制了生成式 AI 在严谨医疗场景中致命的“幻觉”现象，实现了秒级、

全自动生成高度结构化且完全符合 HL7 国际医疗标准的图文病历。



图 2 新华社对智慧手术术中导航系统的相关报道

### 1.3 本研究的意义与拟解决的关键科学与临床问题

基于上述背景，本研究的实施具有深远的理论意义与重大的现实应用价值：

#### 研究意义：

- 临床增效与安全防范：** 验证系统通过 AI 自动化流转显著释放外科核心生产力的效能。预期将图文报告的生成时间由人工拼凑的 20 分钟以上极致压缩至 30 秒以内，将医生从繁重的文书劳动中解放出来回归临床。同时，验证系统通过术中实时提供“安全关键视野”辅助导航，能否在真实世界中将微创手术达到国际公认的“教科书式结局（TO）”的比例由基线的 76.7%提升至 90.0%，并显著降低医源性胆管损伤及大出血等围手术期并发症发生率。
- 医保审计与社会治理重塑：** 验证基于计算机视觉定量识别的手术步骤与高值耗材消耗日志，能否作为具备不可篡改数字时间戳的“可视化证据链”。这种数据要素不仅直接响应国家关于医疗纠纷司法质证的最高要求，更将为政府及医保局实施 DRG/DIP 精细化控费提供无死角的审计凭证。据初步测算，该系统在区域范围内展现出每年节约数亿元，经测算陕西省约 9.2 亿元医保基金的巨大控费潜力，为构建数字社会诚信医疗体系提供硬核管理支撑。

#### 拟解决的关键问题：

- 多中心多术式泛化能力验证：** 在多中心、不同品牌内窥镜设备，如 Storz、奥林巴斯、国产唯精机器人等的复杂物理环境下，基于 DeepSeek 及 Mamba 架构的 SurClaw 系统在自动摘取关键帧及生成标准化医疗术语时的鲁棒性与语义精准度。
- 临床获益的因果机制阐明：** 量化评估 SurClaw 系统的智能预警干预与微创外科手术质量 TO 达成率、CVS 达成率提升及并发症下降之间的直接因果关系。
- 卫生经济学与管理效能评估：** 验证由系统全自动生成的数字化证据链在真实医院病案管理及医保控费，如 DRG/DIP 入组准确率、高值耗材虚记漏记的防范率等工作流中的直接经济学效益。

## 2. 研究目的

本项目旨在开展一项多中心、前瞻性、随机、平行对照的临床试验，全面、客观地评估“基于 SurClaw 的微创外科医保质控与费控图文报告智能生成系统”在真实临床世界中的有效性、安全性、易用性及卫生经济学价值。各项具体研究目的细化如下：

### 1. 首要有效性目的：

验证使用 SurClaw 系统能够显著缩短微创外科术后医疗文书（图文病历）的撰写与整理总耗时，同时验证系统生成病历在结构化、专业度及解剖要素覆盖率上不劣于甚至优于由高年资专科医师人工撰写的病历（对标《倡用图文外科手术记录专家共识（2025 版）》标准）。

### 2. 临床安全性与质控评价目的：

评估应用该智能一体机系统进行术中实时影像监测（涵盖高危操作步骤预警、解剖变异识别及 CVS 视野推荐）后，能否实质性地提高患者的手术安全性。具体表现为微创手术的“教科书式结局（T0）”达成率的显著提升，以及术后 30 天内整体围手术期并发症（如出血、胆漏、吻合口瘘、感染等）发生率的统计学下降。

### 3. 卫生经济学及医保费控目的：

验证系统依托 AI 计算机视觉自动提取的关键节点快照与高值医用耗材（如各型腔镜下切割吻合器、超声刀头、施夹器等）识别日志，是否能有效提高病案首页在 DRG/DIP 结算审核中的一次性通过率。评估其在降低医保拒付率、防范耗材异常计费（虚记、多记、串换）方面为医疗机构带来的合规保障效益及医保基金节约效应。

### 4. 医学教育与带教赋能目的：

通过量化量表探索高质量的 AI 智能标注图文报告作为标准“教学图式”，对降低低年资住院医师外在认知负荷、缩短复杂微创术式学习曲线及提升临床考核成绩的实际作用。

## 3. 研究设计

### 1.1 研究类型

本研究类型为： 临床试验。

在具体的试验设计方法学上，本研究采用多中心、前瞻性、随机化、平行对照、单盲（结局评价者与数据分析师设盲）的设计方案，以提供循证医学证据。



图 3 手术图文记录系统在临床中的使用

### 1.2 研究总体设计和流程图

研究将无缝整合当前公立医院的常规临床路径与智能化的医保监管需求。整个研究流程在时间轴上分为入组随机、术中干预、术后即刻评估及长期随访四个核心节点。

流程阶段	关键操作与随访内容	暴露/预测因子与混杂因素控制
阶段一：患者筛查与入组 (Day -7 至 Day 0)	临床医师基于纳排标准在门诊或病房筛查患者。获取伦理知情同意书，采集患者基线人口学资料、既往病史、术前影像学与实验室检查数据、术前 ASA 分级。	记录可能的混杂因素：年龄、BMI、肿瘤分期、既往腹部手术史（可能导致严重粘连）、基础合并症等。
阶段二：随机分组 (Day 0, 术前)	采用基于网络的中央随机化系统，运用区组随机化方法，按临床中心及拟实施的手术亚专科进行分层，1:1 比例随机分配至 SurClaw 组或传统对照组。	确保各临床中心及不同类型复杂手术在两组间的基线特征均衡分布，避免选择偏倚。
阶段三：手术实施与干预 (T0)	SurClaw 试验组：手术室内部署嵌入式异构一体机，无损采集 4K 视频流。STSH-Net 与 Yun-Trans 算法实时运行，提供去雾与	记录手术时间、术中失血量、中转开腹率、高值耗材实际拆包使用数量。AI 系统在后台静

	阶段预警。对照组：按常规流程实施微创手术，不开启 AI 视觉预警。	默生成机器日志以备核查。
阶段四：术后文书生成与即刻评估 (T1, 术后 2-24 小时)	SurClaw 试验组：手术结束缝皮后，系统在 30 秒内全自动抓取关键帧并生成标准 HL7 图文报告，推入 HIS 系统。对照组：由手术医师凭回忆手动调阅 PACS 抓图，在 HIS 中敲击键盘撰写手术记录。	主要终点评估点：由专职质控员使用秒表精确记录“临床文书生成总耗时”。双盲评估专家对两组病历的解剖覆盖度、结构化规范性进行评分。
阶段五：出院时评估 (T2, 术后 3-14 天)	收集患者术后临床恢复数据，包括术后首次下床活动时间、肠道排气时间、引流管拔除时间、术后住院总天数。	初步判定是否达成教科书式结局 (T0) 的部分指标要求。
阶段六：术后长期随访 (T3, 术后 30 天)	患者门诊复查或电话随访。记录术后 30 天内的任何不良事件，包括但不限于出血、感染、胆漏/肠痿、非计划二次手术及 30 天内再入院情况。	安全性终点评估点：汇总计算最终的围手术期并发症发生率及完整的 T0 达成率。
阶段七：医保与财务审计 (T4, 术后 30-60 天)	提取 HIS 系统中该病案的最终费用清单及病案首页编码。对接医保局 DRG/DIP 模拟审核系统。对比计费耗材数量与实际手术 (录像/AI 日志) 使用数量的差异。	卫生经济学终点评估点：计算账物相符率、DRG/DIP 入组准确率及医保初步拒付比例。

#### 4. 研究人群

##### 4.1 样本来源

本研究依托国家级科研平台及专科联盟资源，样本将以西安交通大学医学院第一附属医院（牵头单位）为主，联合国内其他具备高水平微创外科能力的顶尖三甲医疗机构，基于团队前期在全国 5 家医院实现系统规模化部署的基础，共同招募多中心病

例资源。病例科室涵盖但不限于普通外科（胃肠、甲状腺）、肝胆外科、胸外科及泌尿外科。

#### 4.2 纳入标准

拟入组的受试者必须同时满足以下所有标准：

1. 年龄在 18 岁至 75 岁之间（含界值），性别不限；
2. 经完整的术前临床、影像学（CT/MRI/超声）及病理学或细胞学评估，具备明确的微创外科手术指征，拟行常规微创手术（如腹腔镜解剖性肝切除术、腹腔镜胃肠癌根治术、胸腔镜肺叶切除术等）；
3. 拟行择期手术，且术前麻醉风险评估 ASA（美国麻醉医师协会）分级判定在 I~III 级之间，具备耐受全身麻醉下气腹手术的生理条件；
4. 充分理解本研究的目的与流程，同意将围手术期病案数据（含经过严格自动化脱敏处理的手术视频影像、语音指令及文本信息）用于临床评价与 AI 模型训练优化，并由本人或法定代理人自愿签署由伦理委员会批准的临床试验知情同意书；
5. 患者依从性良好，同意并有能力遵守本研究方案中规定的术后随访与门诊复查计划。

#### 4.3 排除标准

受试者若符合以下任一标准将被排除于本研究之外：

1. 因急性创伤、消化道大出血、急性化脓性梗阻等急腹症需行紧急限期手术的急诊患者；
2. 术前影像学或多学科会诊评估认为局部病灶极其巨大、周围重要脏器或大血管受侵犯严重，极有可能在术中早期即中转开腹的手术病例；
3. 属于复杂二次手术或翻修手术（如胆道重建失败后再手术），预期手术区域存在极其严重的致密粘连，正常解剖结构完全丧失者；
4. 患有严重的心血管系统并发症（如半年内发生过心肌梗死、严重心力衰竭）、终末期肾病（需长期规律透析）、严重肝功能衰竭（Child-Pugh C 级）或其他预期生存期极短（小于 6 个月）的基础疾病，导致术后随访极有可能无法完成者；
5. 认知功能存在严重障碍、患有未控制的精神疾病或存在语言沟通障碍，无法有效理解并签署知情同意书者；
6. 在入组前 3 个月内参加过或同期正在参与其他任何可能对本研究主要终点（如手术并发症率、手术时间）产生严重混杂影响的药物或医疗器械临床试验者。

#### 4.4 研究对象终止或者退出研究的标准

为保护受试者权益并保证研究数据的严谨性，在以下情况下受试者将终止或退出研究：

1. **知情撤回：** 受试者在研究的任何阶段可无需理由、无条件撤回知情同意，主动要求退出研究，且不影响其后续应得的常规医疗照护。
2. **重大方案违背或临床策略变更：** 患者在微创术中发生不可预见的极严重情况（如因

发现肿瘤腹膜广泛播散转移而决定仅行探查活检，放弃原定根治切除术；或发生大出血紧急开腹），导致既定评估的复杂微创步骤未被执行。

3. **失访：** 患者术后未能按计划完成关键的随访节点访视，尤其是术后 30 天并发症评估，且经研究人员通过电话、短信等多种途径连续三次联系未果。
4. **研究者判断或安全性叫停：** 经主要研究者医学判断，受试者发生了极其严重的术后不良事件，继续参与研究随访可能损害其最佳健康利益；或在多中心试验实施过程中，若发现实验设备存在可能威胁患者安全的重大软硬件系统缺陷，伦理委员会或申办方有权紧急叫停研究。

## 5. 研究分组

本试验基于统计学严谨性，采用**前瞻性、随机、平行对照**的设计架构。

所有通过严格的纳入/排除标准筛查并成功签署知情同意书的受试者，将由独立于临床操作的专职数据协调员，通过经过验证的基于网络的中央随机化系统（IWRS）获取分组信息。为控制各分中心手术技术差异及不同亚专科手术难度带来的混杂偏倚，随机化过程采用区组随机化（Block Randomization）方法，并按“临床中心”及“手术大类（如肝胆、胃肠、胸外）”进行双重分层。患者按 1:1 的比例被随机分配至以下两组：

- **试验组（SurClaw 智能辅助组，n=140）：** 在此组中，手术室内将部署完全自主可控的 SurClaw 软硬件一体机（“轻量化盒子+软件赋能”模式）。该设备通过 6G-SDI 无损传输技术截取内窥镜设备的 4K 视频流。在手术全程中，基于 Mamba 架构的 STSH-Net 与基于 MoE 的 Yun-Trans 算法在边缘计算端（FPGA+GPU+CPU）实时静默运行，为医师提供去雾增强与高危节点预警。当手术结束、伤口缝合完毕后，系统将依据内部具有时间戳的视觉目标检测日志与术中语音指令，触发微调后的多模态大模型（DeepSeek 等）配合 RAG 机制，在 30 秒内全自动检索、匹配医学知识库，生成包含核心解剖结构、关键操作截屏、异常事件与高值耗材明细的结构化 HL7 标准图文手术报告。该报告经主刀医师一键审阅电子签名后，通过 API 接口无缝推送至医院 HIS/PACS 系统及医保结算前置接口。



图 4 术者在术中使用手术图文记录系统来完成术中影像实时分割步骤

- **对照组（传统人工记录组，n=140）：**患者在对照组接受与试验组遵循完全相同临床治疗规范标准的高质量微创外科手术。然而，该组的手术过程影像仅作为常规屏幕显示，不进行 AI 实时分析干预。术后医疗文书的生成完全依赖传统的“经验驱动” workflow：主刀医师或助手需要在术后凭借个人的主观回忆，在 HIS 系统的病历模板中手动输入大段文字描述；若需附加图像，则需通过 U 盘或网络手动从内窥镜主机中筛选、拷贝并粘贴 2-3 张照片至文书系统。手术期间的器械及高值耗材计费，则依然遵循传统的由巡回护士进行手工核对与条形码扫描录入的方式。

## 6. 研究评价指标

研究评价指标体系立足于将科研成果转化为医疗生产力，并紧密围绕国家对创新医疗器械在临床有效性及卫生经济学维度的严苛考核标准，设定如下：

### 6.1 主要评价指标

#### 1. 临床文书生成总耗时：

- **测量定义：** 定义为从手术缝皮结束、患者被推出手术室的那一刻起计算，直至该病例的最终版手术记录（包含图文部分）在医院 HIS 系统中完成撰写、校对、排版并施加主刀医生电子签名锁定的绝对时间差。
- **临床意义：** 此指标是量化评价系统释放医生核心医疗生产力、“极致提效”能力的最直观证据。预期将验证单份高质量图文报告生成耗时由传统人工模式的 15-30 分钟大幅缩短至 AI 辅助下的 30 秒以内的显著性。

#### 2. 医保 DRG/DIP 计费审核通过率及耗材账物相符率：

- **测量定义：**（1）账物相符率：提取 AI 系统利用计算机视觉定量识别出的术中高值耗材（如一次性超声刀头、不同颜色的吻合器钉仓等）实际击发与使用数量日志，与最终上传至医保系统的收费明细清单进行精准比对计算的吻合度。（2）审核通过率：该病案在出院结算时，一次性无争议通过国家或省级医保智能审核规则库（含过度诊疗、虚计多记核查）的比例。
- **临床意义：** 通过验证 AI 提供的防篡改可视化证据链，证明其为防范医院高值耗材虚记、串换等违规行为提供的硬核技术支撑，体现其在保障医保基金安全及精细化控费方面的核心价值。

### 6.2 次要评价指标

#### 1. “教科书式结局”达成率：

- **测量定义：** T0 是一种能够综合反映外科手术卓越质量的复合型“全或无”终点指标。本研究中，患者必须在围手术期严格同时满足以下所有条件方被判定为达成 T0：① 无术中不良事件发生；② 术后无严重并发症（Clavien-Dindo 分级  $\geq$  III 级，即需要手术、内镜或放射介入治疗）；③ 无计划外因并发症重返手术室；④ 术后住院时间在各亚专科规定的 DRG 标准预期范围内；⑤ 术后 30 天内无再入院；⑥ 若为恶性肿瘤手术，病理切缘必须为 R0 阴性且淋巴结清扫数目达标。

- **临床意义：** 用于论证 AI 实时监测预警系统对提升微创外科手术整体同质化、安全化水准的因果效应。研究预期证明试验组的 TO 达成率能从既往历史基线的约 76.7% 显著跃升至 90.0% 以上。

## 2. 安全关键视野达成率及并发症发生率：

- **测量定义：** 以胆囊切除等特定手术为例，通过两名盲法独立专家盲审手术录像，判定术中肝胆三角等关键解剖结构的游离显露是否达到了国际公认的 CVS 安全标准。同时精确统计随访期内整体手术事故及相关并发症率的下降幅度。

## 3. 病案质控与结构化内涵评分：

- **测量定义：** 邀请外部资深病案质控专家，基于 2025 版《共识》标准，对系统生成的病历从“关键解剖要素遗漏率”、“逻辑时序正确性”、“医学术语规范度”三个维度进行盲法评分。旨在证明 AI 大模型结合 RAG 机制后生成的文本在专业度上无“幻觉”且优于人工。

## 6.3 其他研究因素

- **医学教育带教效能：** 针对参与该手术过程或事后进行病案复盘的规培医师/住院医师，采用技术接受度模型量表及 NASA-TLX 认知负荷量表进行问卷调查。评估结构化、图文并茂的客观记录在加速年轻医生理解解剖变异、缩短其获得独立主刀资格培训周期方面的教育学价值。

## 7. 研究步骤

本研究规划执行周期为三年（2026 年 01 月至 2029 年 12 月），依据 GCP 规范及医疗器械临床试验管理规定，严密划分为以下四个递进阶段实施：

### 1. 第一阶段：项目启动、伦理审批与基线确立（第 1-6 个月）

- **伦理与合规：** 向牵头单位及各参研多中心的临床研究伦理委员会提交完整的研究方案、知情同意书及设备安全性报告，完成多中心伦理审查备案并获得批件。完成项目在中国临床试验注册中心的预先注册公开。
- **设备部署与软硬件联调：** 在各中心选定的数字化核心手术室内部署 SurClaw 异构边缘计算一体机。由技术工程师完成 6G-SDI 硬件采集线缆的无损布设，并与院方信息科协同，打通与 HIS、PACS 及医保前置审核平台标准 API 数据接口。
- **人员培训与同质化管理：** 组织开展多中心研究者启动会。针对所有参与研究的外科医师、麻醉师、巡回护士及专职临床研究协调员进行严格的方案培训、GCP 规范培训及系统操作规程演练，确保数据录入与干预实施的高度同质化。

### 2. 第二阶段：受试者招募、干预实施与随访执行（第 7-24 个月）

- **招募与入组：** 严格按照纳排标准在各中心连续招募符合条件的患者，充分告知风险收益后签署知情同意书，通过 IWRS 系统进行在线中央随机化分配。

- **手术实施与数据采集：** 受试者进入手术室后，试验组全程开启系统进行多模态视觉 AI 分析与去雾强化；术后主刀完成伤口缝合即刻，系统秒级生成图文报告及耗材证据链并上传。对照组则维持传统手工图文提取与键盘打字撰写模式。专职人员同步使用秒表等工具精准记录 TRG 等即时终点数据。



图 5 手术图文记录系统在记录、分析临床数据并上传等方面发挥重要作用

- **术后随访推进：** 术后出院时记录恢复节点数据。术后 30 天由独立随访护士通过门诊复查或电话完成患者并发症评估，并将源数据实时双录入至符合 21 CFR Part 11 标准的电子数据采集系统。
3. **第三阶段：数据清理、医保审计回溯与盲态审核（第 25-30 个月）**
- **医保与财务对账：** 在患者出院 30 至 60 天后的结算周期内，由专门的卫生经济学研究员在医院财务及医保科协助下，调取研究队列的 DRG/DIP 分组反馈结果、耗材收费账单明细，对比机器视觉生成的客观凭证，统计拒付率与账物相符率差值。
  - **数据监察与锁库：** 临床监察员完成所有分中心的 100%源数据核查，处理完所有数据疑问后，召开盲态审核会议，确定各分析数据集，由独立数据监察委员会签署文件后正式锁定数据库。
4. **第四阶段：统计分析、总结报告撰写与器械注册推进（第 31-36 个月）**
- **破盲与统计分析：** 移交至独立统计分析团队进行破盲，运用 SAS/R 语言运行既定统计分析计划，输出多维度图表。
  - **成果转化：** 撰写多中心临床研究总结报告。形成高质量的真实世界证据，为 SurClaw 系统从当前的创新型二类医疗器械向高阶的“三类手术术中智能导航设备”的阶梯式演进及国家药监局注册申报筑牢底座基础。
  - **学术发表：** 整理核心临床论据与算法创新点，撰写高水平学术论文，拟投稿于《Annals of Surgery》、《中国医疗设备》等国内外权威期刊，引领行业标准。

## 8. 统计分析

## 8.1 样本量估计

本临床试验的样本量估算是依据主要评估维度的核心临床终点——“教科书式结局（TO）达成率”进行严密计算的。该研究设计为检验试验组（SurClaw 辅助组）相较于对照组（传统组）的优效性。

依据项目前期开展的小样本探索性研究结果，以及领域内权威文献（例如团队发表的专病库预评估数据<sup>2</sup>），采用传统手术记录及纯经验质控工作流的对照组（ $P_c$ ），其微创手术 TO 达成率的历史基线水平约为 **76.7%**。在引入具备实时风险预警及高质量循证复盘能力的 SurClaw 系统后，预期试验组（ $P_t$ ）的 TO 达成率能够发生质的跃升，提升至 **90.0%**。

统计学参数设定如下：

- **检验水准（双侧  $\alpha$ ）**： 设定为 0.05（允许 5% 的假阳性错误）。
- **检验效能（ $1 - \beta$ ）**： 设定为 80%（确保有 80% 的概率发现真实的优效性差异）。
- **优效性界值（Margin,  $\delta$ ）**： 本研究旨在证明纯粹的临床获益绝对提升，故设定为 0。
- **分组比例**： 试验组与对照组为 1:1 等比例分配。

依据两独立大样本率比较的优效性样本量估算经典数学公式：

$$N_{per\_group} = \frac{(Z_{1-\alpha/2} + Z_{1-\beta})^2 \times [P_t(1 - P_t) + P_c(1 - P_c)]}{(P_t - P_c - \delta)^2}$$

将各项参数代入公式推演计算：

1.  $Z_{1-\alpha/2}$ （当  $\alpha = 0.05$  时的标准正态分布分位数） $\approx 1.96$
2.  $Z_{1-\beta}$ （当  $1 - \beta = 0.8$  时的标准正态分布分位数） $\approx 0.84$
3. 分子部分系数： $(1.96 + 0.84)^2 = 7.84$
4. 率的方差和：  

$$0.900 \times 0.100 + 0.767 \times 0.233 \approx 0.2687$$
5. 分母（效应量平方）： $(0.900 - 0.767)^2 \approx 0.017689$
6.  $N \approx 7.84 \times 0.2687 / 0.017689 \approx 119.09$

计算得出，为保证统计学效能，理论上每组至少需要成功随访 119 例受试者。考虑到外科临床试验由于极度复杂的手术中转、术后长达 30-60 天随访期内固有的失访脱落以及极其少见的方案违背现象，保守估计设置约 **15% 的脱落率**。

最终确定计划入组样本量为：每组 **140 例**，两组总计招募 **280 例** 患者入组。

（补充说明：该 280 例的大样本量设置，对于另一项呈现连续变量分布的终点——“临床文书生成总耗时”，在面对预期将从均值 15-20 分钟断崖式下降至 0.5 分钟前后的超大效应量差异时，其统计检验效能将无限趋近于 100%，完全满足评价要求。）

## 8.2 数据分析方法

本研究将由独立第三方统计团队执行，使用国际通用且经验证的统计分析软件（R 语

言 4.3 版本) 进行所有数据处理。所有显著性检验均采用双侧检验, 检验水准预先设定为  $P < 0.05$  以确定统计学显著性差异。

#### 1. 分析集定义与缺失值处理:

- **意向性分析集/ 全分析集:** 包含所有被随机化分配且至少进行过一次手术干预操作的受试者。对于 FAS 集中随访脱落产生的缺失值, 采用多重插补法或最末次观测值结转法进行处理。有效性主分析将基于 FAS 集。
- **符合方案集:** 仅包含严格遵循入排标准、全程依从临床干预方案、顺利完成手术且未发生重大方案违背、并成功收集到所有关键终点数据的受试者。PPS 将用于支持性及敏感性分析, 以验证结论的稳健性。
- **安全分析集:** 用于统计不良事件, 包含所有接受过手术且有至少一次安全性评估记录的患者。

#### 2. 定性数据分析:

- 对于主要终点之一的“DRG 审核通过率”、“耗材账物相符率”, 以及次要终点“TO 达成率”、“CVS 达成率”和“并发症发生率”等分类数据, 采用频数和百分率(%) 进行统计描述。
- 组间基线与终点比较, 采用 **Pearson  $\chi^2$  检验**。若样本出现理论频数较小(如  $< 5$ ) 的格子, 则采用 **Fisher 确切概率法**。
- 为校正多中心临床试验带来的潜在中心效应及不同术式难度, 如肝切除与阑尾切除差异造成的混杂, 将采用 **CMH 检验**进行分层校正分析。

#### 3. 定量数据分析:

- 对于“临床文书生成总耗时”、“病历结构化评分”等连续性数值变量, 首先进行数据的正态性检验及方差齐性检验。
- 对于服从正态分布且方差齐的数据, 采用均数±标准差描述, 两组间比较采用 **独立样本 t 检验**。
- 若数据呈偏态分布或方差不齐(预估文书生成时间极可能呈极度偏态), 则采用中位数及四分位间距描述, 组间差异比较采用非参数的 **Wilcoxon 秩和检验**。

#### 4. 生存期与时序随访数据分析:

- 对于随访期间“发生术后并发症的时间点”、“术后首次排气/下床时间”、“非计划再入院间隔”等含有删失数据的时序相关终点, 采用 **Kaplan-Meier 乘积极限法** 绘制生存分析曲线。
- 组间差异的统计学检验采用 **Log-Rank 检验**。

#### 5. 多因素回归模型校正分析:

- 由于真实世界外科患者存在复杂的异质性, 不可避免地存在残余混杂。将建立 **多元 Logistic 回归模型**, 针对 TO 达成、医保违规拒付等二分类终点 或 **多重线性回归模型**, 针对耗时等连续终点。
- 模型中将“是否使用 SurClaw 干预”作为核心自变量, 同时强制纳入年龄、性别、BMI、术前 ASA 评分、恶性肿瘤 TNM 分期、基础代谢合并症及手术大类作为协变量进行多因素校正。计算干预措施的独立优势比或回归系数 ( $\beta$ ) 及其 95%可信

区间（95% CI），深度剖析系统获益的独立因果机制。

## 9. 研究是否涉及特殊医疗或检测技术

[√] 有。本项临床试验深度融合了当前人工智能技术领域最高阶的“多模态时空大模型认知架构”与“软硬件异构协同边缘计算”两项特殊前沿计算机科学技术。

- 1. 架构维度：自主可控的“端-边-云”异构医疗边缘计算一体机平台。** 为彻底打破数字设备对网络带宽的依赖，并最高限度地满足《网络安全法》对患者隐私医疗数据“不出室、不上云”的严苛物理隔离要求，本系统创新性地采用了工业级的“FPGA + GPU + CPU”三位一体嵌入式异构计算底层平台。其中，FPGA 作为底层视觉中枢，直接并发生效处理经 6G-SDI 无损线缆透传而来的 4K/60FPS 超高清视频信号流，实现“零延迟”的裸流采集，确保主刀医生手眼协调的绝对临床操作安全。
- 2. AI 认知技术维度：突破传统范式的 Mamba 状态空间模型与 LLM 知识增强技术。**

- **STSH-Net 与 Mamba 架构：** 抛弃了传统医疗 CV 依赖的 Transformer 二次方计算复杂度架构，在全球范围内率先将具备线性计算复杂度 ( $O(N)$ ) 的 Mamba 结构深度引入医疗领域，构建 STSH-Net。配合 Yun-Trans 自适应去雾网络及 BiMRF-ADS 形变追踪框架，解决了超长几小时手术视频流中的因果逻辑关联衰减及重度烟雾下的软组织形变追踪难题。

**交互式多模态大语言模型与 RAG 机制：** 在报告自然语言生成端，系统不仅仅是简单的图像分类，而是跨界引入了 DeepSeek、Qwen-VL 等参数量庞大的多模态视觉大语言模型进行医学专属指令微调。最为特殊的是，系统集成 RAG 机制，在生成文书时强制模型外挂实时检索权威的外科手术标准指南及《专家共识》解剖图谱库。通过统一语义空间的 Q-Former 多模态对齐映射，从算法底座上彻底根治了通用大模型在医疗辅助决策中容易瞎编乱造的“幻觉”顽疾，确保输出的图文病历客观、严谨，满足 HL7/FHIR 等国际最高级别医疗证据文书互认标准。

## 四、创新性和可行性分析

### 1. 本项目特色及创新性

本项目的理论深度、算法先进性与政策响应度在同类研究中均处于国际领跑身位。其特色与创新性集中体现在以下三个核心维度：

#### 1.1 理论范式创新：打破“被动机械记录”，构建基于大模型语义驱动的手术时空因果建模新思想

长期以来，学术界与医疗产业界对微创手术录像的处理，仅仅停留在将其视为由静态像素帧组成的“物理影像流”，多关注于存储或简单的器械目标检测，缺乏对长序列手术操作因果关系的深入解读。本课题在全球首次提出将长达数小时、高度动态的微创手术全过程视为一种蕴含着特定医学语法结构与时空因果逻辑的“动态文本”。系统不再是被动



地“机械录像”，而是通过多模态时空大模型的感知与认知重构，将庞杂非结构化的影像数据，实时、精准地转化为机器可理解、可分析的“结构化解剖知识流”。这一颠覆性的认知跃迁，使得系统兼具了类人的逻辑推理能力，从根本上为外科医疗文书从“传统经验依赖模式”向“数据智能驱动精准科学范式”转型提供了极其坚实的理论思想支撑。

### 1.2 核心算法创新：攻克长视频处理“不可能三角”，引领多模态大模型的外科临床级落地

本项目在人工智能算法框架的设计上构筑了极深的技术护城河，重点攻克了微创外科面临的数项公认“卡脖子”技术壁垒：

- **计算架构突破：** 针对数小时 4K 级手术长视频实时分析中“高精度、长时序依赖与低算力消耗”的“不可能三角”瓶颈，创新采用具备完美线性计算复杂度 ( $O(N)$ ) 的 Mamba (状态空间模型) 作为视觉核心骨架，开发出 STSH-Net 网络，彻底取代了算力开销庞大的传统 Transformer，实现了在边缘计算资源受限下的 52 FPS 超高帧率实时推理。
- **抗干扰感知突破：** 针对腹腔镜手术中高频电刀切割组织产生大量焦烟导致视野致盲的重大临床痛点，首创 Yun-Trans 自适应去雾网络。该网络利用最前沿的 MoE (动态专家混合) 路由机制，动态识别烟雾物理属性并智能切换清雾算力，将重度干扰环境下的病灶及解剖结构识别精准度从极其低下的 35.66% 跳跃式拔高至 74.15%。
- **大语言模型融合落地：** 在报告生成链路中，系统通过 RSMR-Net 深度集成微调版 DeepSeek 多模态大模型，并创造性采用 RAG 知识增强检索机制，完美解决了 AI 生成医疗文书时的“幻觉”现象。单份排版精美、高度结构化的 HL7 标准图文病历生成耗时由人工时代的 15 分钟断崖式缩减至 30 秒以内，关键步骤识别精度稳超 90%，成为全球首个将多模态大语言模型成功且稳健应用于外科手术报告领域的临床级标杆案例。

图 6 中国教育报对微创智能手术辅助系统的相关报道

### 1.3 医保管理与数字社会治理创新：打造 DRG/DIP 支付改革与诚信医疗的“可视化上帝视

角”

本项目的社会治理与卫生经济学创新远超其纯粹的技术价值。在国家医保支付体系向 DRG/DIP 全面转型的深水区，面对医院普遍存在的术中高值耗材监管困难、虚假计费及串换骗保等侵蚀医保基金的乱象，系统的诞生重塑了医院精细化管理体系与政府监管机制。它通过对术中高价值业务动作及耗材消耗的机器视觉定量监控，为每一台手术自动凝炼出带有“不可篡改数字时间戳”的关键帧证据链。这不仅一举击碎了医疗纠纷中因手术室“黑盒”状态而导致责任难以界定的司法质证难题，更直接为医保局、卫生健康委等监管部门实施 DRG/DIP 客观审计提供了拥有“上帝视角”的防伪客观凭证。据权威数据要素价值测算，仅仅依托于该系统对高值耗材使用情况进行精准监控、杜绝虚记漏报一项，在省级推广应用后，每年即有望为单一省份（如陕西省）的医保基金池安全优化出高达约 9.2 亿元的支出空间。这一管理模式创新将为构建国家数字健康诚信医疗体系提供极其强大的技术抓手与可广泛复制推广的“交大方案”。

## 2. 可行性分析

本项目的顺利实施、验证及最终的高质量产业化转化，拥有极为坚实的数据底座、成熟的工程原型、顶配的交叉学科人才梯队以及深厚的政策红利支持。各项可行性保障具体分析如下：

### 2.1 独一无二的大规模高质量数据“护城河”与行业标准背书

人工智能大语言模型在任何垂直医学领域的成功调优与能力泛化，均高度、甚至是绝对依赖于多模态数据要素的规模广度与标注精度。本项目研发底座汇聚了由西安交大一附院等全国多中心协作、历经数年积累形成的目前国内已知规模最为庞大的专病影像数据集。该数据集体量高达惊人的 1.5PB，不仅完整涵盖了 692 例多中心肝胆、胃肠等真实世界患者的超高清视频流，更针对多达 24.3 万余帧的手术核心画面，由高级职称专科医师团队逐帧完成了精细化的多层级语义（器械-动作-目标）三元组标注。这种极其稀缺且具备极高置信度的专病影像数据库，构筑了本项目训练出的多模态大模型具备卓越鲁棒性、能彻底根除医疗“幻觉”的最强底层基石保障。此外，团队牵头撰写发布了 2015 版与更新的 2025 版中华医学会《倡用图文外科手术记录专家共识》，这种在全国范围内作为行业金标准推广的指导文件，为系统最终生成的报告提供了无懈可击的专业权威性与合规性背书。

**倡用图文外科手术记录专家共识（2025 版）**

中华医学会外科学分会外科手术学学组

2015 年中华医学会外科学分会外科手术学学组制定了第一版《倡用图文外科手术记录专家共识》，使手术记录能更客观、详细地反映手术实境。近年来，信息技术的迅猛发展为图文手术记录的进一步的推广提供了技术保障，经过多方、多轮次专业的讨论、调研，制定《倡用图文外科手术记录专家共识》，旨在促进图文手术记录的标准化和规范化，促进外科领域手术记录的标准化和规范化，并鼓励有条件的单位逐步推广和应用。

**关键词：**图文式手术记录 医院管理 电子病历 手术步骤 标准化

**Expert consensus on advocating the use of graphic surgical records**

**(2025 edition)**

**Group of Operation, Society of Surgery, Chinese Medical Association**

In 2015, Group of operation, society of surgery, Chinese Medical Association formulated the first edition of the "Expert Consensus on Advocating the Use of Graphic Surgical Records", which enables surgical records to reflect the surgical reality more objectively and in detail. In recent years, the rapid development of information technology has provided technical support for the further promotion of graphic surgical records. Therefore, our academic group plans to revise the consensus in order to promote the standardization and normalization of graphic surgical records, and encourage eligible its application.

**Key Words:** graphic surgical records, hospital management, electronic medical record, surgical procedure, standardization

图 7 《倡用图文外科手术记录专家共识（2025 版）》摘要

## 2.2 成熟的工业级软硬协同平台与前期规模化验证基础

本课题并非停留在纯理论阶段的概念验证（PoC），而是建立在已经成熟运转的工程化系统之上。项目组已经完成基于 FPGA+GPU+CPU 工业级三位一体异构计算架构的一体机样机研发，克服了高算力下的散热与电磁兼容难题。由于采用了“轻量化盒子+底层软件赋能”的技术路线，该国产化替代方案系统具备极高的泛化性与经济性，能够无缝兼容包括德国 Karl Storz、日本奥林巴斯以及国产唯精多臂手术机器人等各品牌型号现役内窥镜设备，而其单台部署综合成本被极具竞争力地压缩至仅为进口同类导航设备的十分之一。在真实的临床场景可行性方面，SurClaw 智能系统展现出了惊人的落地能力，目前已成功跨区域辐射至全国 16 家顶尖三甲医疗机构的 5 大核心外科科室。系统累计已为超过 4 万人次的微创手术患者提供了安全不间断的 AI 实时辅助，常态化规模生成了 3000 余份高质量图文病历报告。这不仅证实了系统的普适性与稳定性，更意味着该项目具备在短期内完成预期大样本量入组及收集真实世界证据的强大执行可行性。



图8 上合组织成员国外宾参访并了解手术图文记录系统

### 2.3 顶尖的“医工企交叉”复合型科研团队配置

本研究的人才梯队深度践行了国家提倡的产学研医交叉融合模式。项目骨干团队依托于西安交通大学“双一流”工程学科的强大算力支撑，以及全国首批国家医学中心建设单位西安交大一附院无与伦比的临床病案资源。项目负责人彭子洋博士具有极其扎实的临床医学背景，并深耕于医工学交叉的人工智能外科诊疗一体化管理领域前沿，此前已在该领域取得了丰硕的研究成果，并屡获“优秀研究生标兵”及“产教融合之星”等殊荣，具备卓越的项目统筹及执行能力。此外，课题获得了其导师、国内微创与磁外科领域领军人物吕毅教授的鼎力支持（曾主导相关领域国家自然科学基金重大仪器专项、国家科技进步二等奖、教育部技术发明一等奖），从顶层架构上确保了项目的临床前沿高度。加之西安交通大学未来技术学院王志博博士等人工智能骨干团队在Mamba核心算法框架与大模型指令微调维度的硬核算力与技术护航，这种“临床提需求、工科解难题”的无缝对接团队结构，使得攻克该课题的所有技术拦路虎变得切实可行。

# 驼人-交大一共建医疗器械研究院

“未来”：教学、科研、产业的集合

从成为西安交大未来技术学院本科生的那天起，这些学生就拥有了一路读到博士的资格，本硕博贯通培养机制让他们潜心于科研。

“我们要用十年磨一剑，培养可堪重用的未来科技领军人才。”王小华说。

为了提升人才培养质量，西安交大在课程体系、教师队伍建设、学位评价体系等方面不断做着改进，其中很多细节也足见用心，比如推行课一项一赛一体化、注重培养学生的家国情怀、提升学生领导能力等等。



博士研究生彭子洋与王志博参加第六届智慧医疗创新大赛全国总决赛答辩现场

“未来技术学院可以说是一个小型的教育、科技、人才一体化发展的示范点，它有三重功能：育人的学院，科学技术的研究院，未来产业、未来学科的孵化器。”未来技术学院副院长方敬介绍，学院目前已孵化出医工交叉学科专业，在行业有一定引领作用。

医工学方向的博士研究生彭子洋和王志博认为，他们是学科交叉培养模式最大的受益者。两人联手研发设计了首个国产微创腔镜外科手术导航体系，并因此获得6项国家级竞赛奖项，申报专利8项，以第一作者发表中国科学院SCI收录论文4篇，他们均被评为首届西安交通大学“产教融合之星”，全校仅有10人获此荣誉。

图 9 陕西省科学技术厅对交大医工学博士进行报道

## 2.4 深度契合国家政策导向与无阻碍的产业化成果转化通道

本临床项目在宏观发展环境上，踩准了国家深化医改与数字新基建的时代节拍。系统的推广应用高度吻合国家《“健康中国 2030”规划纲要》要求，并直击国家医疗保障局关于推进 DRG/DIP 支付方式精细化改革、强化医保基金智能监管的核心痛点与刚性政策红利。在产业化潜能上，本课题支撑的核心技术体系已累计获得 4 项国家发明专利（如 ZL 2024 1 1527344.8：一种微创腔镜图文报告生成系统及方法等）和 2 项核心多模态大模型及病历报告生成相关的软件著作权授权，构筑了完备的知识产权保护壁垒。同时，项目此前已大放异彩，先后斩获了 2024 年国家数据局“数据要素×”大赛医疗健康赛道一等奖、中国国际大学生创新大赛（国赛）金奖等一系列极具含金量的国家级科技荣誉，并获得了数千万元级别的直接研发资金与行业头部器械厂商（如联影医疗）的战略意向投资。依托这些极其丰厚的政策认可、资金背书以及前期的临床应用检测数据，在项目结题后，推动该二类创新医疗器械在国家药监局（NMPA）完成临床注册、进而向三类术中智能导航系统进阶的产业转化道路将畅通无阻，具备百分之百的可行性与极高的社会经济回报前景。

近日，陕西省教育厅、学位委员会公布了陕西省第八届研究生创新成果展高质量成果名单。西安交大获评27项高质量成果，其中A档7项。外科梦工场团队再传捷报，获评2项A档、1项C档。研究生创新成果展是我省研究生科研水平的集中展示平台，勉励高校研究生勇于探索创新，为加快科研成果转化贡献力量。参展研究生与评审专家、同行交流互动，促进了学术思想的碰撞和融合，推动学科领域的发展。成果展同时吸引着众多企业和科研机构的关注，为校企合作和科研成果转化提供了契机和平台。

成果档次	作品名称	主要完成人	指导教师	学院(系、中心)
A档	微创手术导航与记录平台	郭子博、王志刚、张雷雷、王强、周松、李俊强、张博、魏东、徐雯、魏建博、黄山、王强	杜毅、王小平、冯家豪、杨理	未来技术学院
A档	可变形自适应智能骨科设计仿真实用	张磊磊、冯强、郭子博、郭子博、王强	尹全刚、杜毅	医学部
C档	微创骨科-3D打印骨科精准定制治疗系统	张磊磊、郭子博、王志刚、周松、郭子博、王志刚、周松、李俊强、杨理	杜毅、王小平	未来技术学院

图 10 微创手术导航与记录平台作品荣获陕西省第八届研究生创新成果展 A 级

## 五、进度计划和预期目标

### 1. 进度计划

本项目研究周期为3年（2024年X月至2027年X月），分为四个阶段推进，各阶段任务及时间节点如下：

阶段	时间节点	主要任务	量化考核指标
第一阶段： 筹备与启动	第1-6个月	①完成牵头及分中心伦理审查，获得批件并完成 ChiCTR 注册。 ②完成各中心 SurClaw 一体机部署、6G-SDI 线缆铺设及与 HIS/PACS/医保前置接口的联合调试。 ③召开多中心研究者会议，完成所有研究人员 GCP、方案及设备操作的标准化培训与考核。	①获得全部中心伦理批件。 ②软硬件系统 100%通过院内信息科安全及接口联调测试。 ③研究人员培训考核通过率 100%。
第二阶段： 临床入组与干预	第7-24个月	①依据纳排标准连续入组 280 例受试者，通过 IWRS 系统随机分组。 ②按方案实施手术干预、术后文书生成及数据采集。 ③完成全部受试者术后 30 天随访，数据实时录入 EDC 系统。 ④在第 14 个月进行中期分析，由 IDMC 评估主要终点趋势与安全性，决定是否继续或调整试验。	①完成全部 280 例受试者入组。 ②30 天随访完成率 $\geq 90\%$ 。 ③源数据核查疑问率 $< 5\%$ 。
第三阶段： 数据核查与锁库	第25-30个月	①完成所有病例的医保费用对账与 DRG/DIP 审核结果收集。 ②CRA 完成 100%源数据核查，解决所有数据质疑。 ③召开盲态审核会议，确认方案违背、分析数据集，锁定数据库。	①数据核查完成率 100%。 ②数据库锁定。
第四阶段： 总结与转化	第31-36个月	①移交独立统计团队，完成统计分析报告。 ②撰写临床研究总结报告及分中心小结。 ③整理核心成果，投稿 SCI 及中文核心期刊论文 1-2 篇。 ④准备 NMPA 医疗器械注册申报材料，启动创新医疗器械特别审查程序申请。	①完成 1 份总统计分析报告。 ②撰写 1 份临床研究总报告。 ③发表高水平论文 1-2 篇。 ④形成完整的医疗器械注册申报资料包。

## 2. 预期目标

### 总体目标：

通过本项多中心 RCT，产出来自真实世界的高等级临床证据，验证 SurClaw 系统在提升临床效率、保障手术安全及赋能医保控费方面的显著优势，为全球手术记录范式变革提供“中国方案”，并完成从科研创新到产业转化的关键跨越。

### 具体量化预期目标：

目标维度	预期成果指标	预期量化值
临床效能	图文报告生成耗时 (TRG)	试验组 $\leq 30$ 秒，对照组 $\approx 15-30$ 分钟，组间差异 $p < 0.001$
	病历结构化内涵评分	试验组评分显著优于对照组，关键解剖要素遗漏率 $< 3\%$
临床安全与质量	教科书式结局 (TO) 达成率	试验组 $\geq 90.0\%$ ，对照组 $\approx 76.7\%$ ，绝对提升 $13.3\%$
	术后 30 天严重并发症发生率	试验组相较对照组显著降低 (预期降幅约 $25\%$ )
卫生经济学	高值耗材账物相符率	试验组 $\geq 98\%$ ，对照组可能存在的模糊/差误被显著压缩
	DRG/DIP 审核一次性通过率	试验组较对照组实现显著提升，医保拒付率大幅降低
学术产出	论文发表	拟发表 SCI 论文 1-2 篇 (目标期刊如 <i>Annals of Surgery</i> )，中文核心 1 篇
	专利与软著	新增申请发明专利 1-2 项，软件著作权 1 项
产业转化	医疗器械注册	完成 NMPA 三类创新医疗器械注册申报，获得受理通知书

## 六、安全性论证和风险预案

### 临床研究的安全性论证

本项目所涉 SurClaw 系统属于独立于治疗器械之外的辅助监测与记录设备，其核心功能为**实时无损采集视频流并进行后台 AI 分析**，不参与或干预手术操作，不向患者体内发射任何能量或引入任何物质。因此，对受试者造成直接物理伤害的风险极低。

**数据安全与隐私保护论证：**系统采用“端-边-云”架构，核心运算均在手术室内的边缘一体机上完成。视频数据经脱敏后仅在本地进行结构化分析，**原始数据及图像“不出手术室、不上公有云”**，仅将生成的文本报告及脱敏后关键帧上传至 HIS，符合《个人信息保护法》及《健康医疗大数据标准、安全和服务管理办法》的严苛要求。

**设备电气与电磁兼容安全：**SurClaw 一体机硬件在进入临床前，已通过电气安全与电磁兼容性（EMC）检测，确保其不会对内窥镜、麻醉机、高频电刀等关键生命支持设备造成电磁干扰，具备在复杂手术室电磁环境中安全运行的能力。

**信息安全与防篡改：**系统生成的图文报告及高值耗材日志均通过加密管道传输，并加盖不可篡改的数字时间戳与哈希校验值，确保了医疗证据链的完整性与真实性，从技术上防止信息被篡改。

### 临床研究的风险预案

尽管系统本身风险极低，但作为一项临床研究，必须对研究全流程中的潜在风险进行充分预估并制定预案，将风险最小化。

风险类别	具体风险描述	发生概率	最小化措施与应急预案
数据隐私风险	手术视频或患者信息发生泄露。	极低	① 系统启动即自动对人脸、非手术区域进行模糊化脱敏。 ② 数据本地化存储，禁绝 USB 等物理拷贝路径，网络接口仅单向推送文本报告。

<p>设备故障 风险</p>	<p>术中出现死机、图像卡顿、采集中断等影响正常手术流程的情况。</p>	<p>低</p>	<p>③ 所有研究人员签署保密协议,违反者承担法律责任。</p> <p>① 系统设计为旁路非介入式连接,硬件采集有 bypass 直通功能,即使系统断电,内窥镜原画面不受任何影响。</p> <p>② 手术室常规备份传统手术录像机,确保原始影像不丢失。</p> <p>③ 术前制定科内排查清单,每台开机前需完成自检。</p> <p>④ 若发生故障,由巡回护士立即断开采集链路即可,试验组手术转为常规流程。</p>
<p>临床决策 误导风险</p>	<p>AI 生成的实时预警或术后报告存在“幻觉”或关键信息遗漏,导致临床判断错误或纠纷。</p>	<p>低</p>	<p>① “人机双审”闭环机制:所有 AI 预警仅为第二意见,最终决策权始终归于主刀医生。AI 报告生成后,必须由主刀医师进行一键审阅、修订并电子签名后方能生效。</p> <p>② 系统内置的 RAG 增强检索和不动点病灶</p>

受试者损伤风险 所有手术相关并发症、不良事件。 与常规手术相当

锚定,已在前期测试中将手术关键步骤识别精度稳定在 90%以上。

③ 发生争议时,以 AI 与人工记录双重核对为准,并备有原始视频可溯源。

① 根本原则: 患者接受的手术质量不会因是否入组研究而降低。试验组与对照组采用完全相同的标准外科治疗方案。

② 严格遵守“教科书式结局(TO)”的多项质控要求,术中强制进行关键安全视野(CVS)确认。

③ 不良事件(AE)监测: 从入组至术后 30 天,对所有 AE/SAE 进行逐级记录与上报。一旦发生 SAE,研究者在 24 小时内向 PI、伦理委员会及申办方报告,并按照 GCP 规范积极救治、跟踪直至事件解决或病情稳定。

④ 设立独立的数据与安全监察委员会 (IDMC)，定期审查安全性数据，若发现试验组并发症发生率异常增高，有权叫停研究。

特殊人群保护：

本研究纳入 18-75 岁成人，不涉及儿童、孕妇等弱势人群。对于高龄受试者（60-75 岁），会特别关注其术前 ASA 分级及基础疾病，严格执行纳排标准，术中加强监护，确保其安全。所有受试者均享有在任何阶段无条件退出研究的权利，其医疗待遇与权益不受任何影响。

## 七、申请者工作积累及合作基础

### 1. 申请者工作积累

项目负责人彭子洋，为西安交通大学临床医学（医工学方向）博士研究生，长期深耕于智能化手术导航体系与医工结合外科诊疗一体化管理领域，具备扎实的临床医学背景与前沿的人工智能交叉研究能力。在博士研究期间，系统掌握了多模态时空大模型、计算机视觉、边缘计算等关键技术，并深度参与导师吕毅教授主持的国家自然科学基金重大仪器专项等国家级课题，全程参与了 SurClaw 智能图文报告系统从临床需求定义、算法选型（如 Mamba 架构的引入）、数据集构建到多中心临床部署验证的全链条研发工作。对微创外科的临床痛点、手术流程及 AI 落地的工程瓶颈有深刻的理解。作为核心骨干，协调了项目组与全国 16 家顶尖三甲医院的合作对接，成功推动了 SurClaw 系统的跨区域规模化部署与临床数据采集工作，展现了出色的组织协调与项目管理能力。此前已作为核心完成人，协助团队斩获 2024 年国家数据局“数据要素×”大赛医疗健康赛道一等奖、中国国际大学生创新大赛国赛金奖，并获评西安交通大学“优秀研究生标兵”及“产教融合之星”等荣誉称号。个人已取得多项前期研究成果，在国内外高水平期刊发表多篇学术论文，并作为主要发明人之一申请了多项国家发明专利，具备独立承担本项目研究任务的坚实基础。

### 2. 合作基础及前期成果

本项目依托西安交通大学第一附属医院、未来技术学院及国内顶尖 AI 团队的深度合作，已形成“临床提需求、工科攻算法、企业做产品、医院做验证”的成熟闭环，合作基础极其深厚，前期成果丰硕。

**1. 核心技术壁垒已突破：**围绕微创腔镜图文报告生成、手术视频智能分析、关键帧自动抓取等核心技术，团队已累计获得 4 项国家发明专利授权（如：一种微创腔镜图文报告生成系统及方法，专利号：ZL 2024 1 1527344.8），并拥有 2 项核心软件著作权，构筑了坚实且完整的自主知识产权保护体系。团队于 2015 年及 2025 年先后两次牵头发布《倡用图文外科手术记录专家共识》，确立了该领域的全国行业规范，为本项目产出的图文报告提供了权威的专业背书与合规保障。

**2. 工程化系统已规模化验证：**已成功研发出基于“FPGA+GPU+CPU”工业级异构计算架构的 SurClaw 软硬件一体机，攻克了高算力下的散热与电磁兼容难题。系统能够兼容 Storz、Olympus 及国产唯精等主流内窥镜系统，单台部署成本极具竞争力。系统已在全

国 16 家顶尖三甲医院的 5 大核心外科成功部署，累计为超过 4 万名微创手术患者提供了安全稳定的 AI 实时辅助，并常态化生成超过 3000 份高质量图文病历报告，充分验证了系统的稳定性、普适性及临床可行性。

**3. 顶级团队保障与资源支持：**以西安交大一附院吕毅教授（中华医学会外科学分会常委、磁外科与微创领域领军人物）为核心的顶级外科团队，为本项目提供了无可比拟的临床病案资源、手术技术指导及学术前瞻性把握。联合西安交通大学未来技术学院骨干（如王志博博士等），在 Mamba 核心算法框架、多模态大模型微调及边缘 AI 推理部署方面提供硬核技术支持。项目前期已获得国家数据局顶级奖项认可及数千万元级别的研发资金与战略投资（如联影医疗等），为本研究的顺利实施及后续产业化转化提供了充足的资源保障。

### 3. 近五年论文代表作

1. 4 份授权专利（与本项目强相关）：



# 驼人-交大一共建医疗器械研究院



- (1) ZL 2024 1 1527344.8 一种微创腔镜图文报告生成系统
- (2) ZL 2025 1 0096041.3 一种实时术中腔镜影像分析装置及方法
- (3) ZL 2025 1 0896530.7 模型与器官配准以辅助导航的方法及系统
- (4) ZL 2025 1 0933524.4 一种喉部的体外磁靶标定位导航方法及系统

## 2. 论文:

4. 研究团队

项目组主要成员 (含申请者)	姓名	职称	工作单位	项目中的分工	签名
	彭子洋	住院医师	西安交通大学第一附属医院		
	王志博	博士研究生	西安交通大学		