

· 临床教学 ·

机器人手术教学培训中 GEARS 评分与客观操作指标相关性的初步探索

肖行远, 刘振宇, 杜远卓, 杨 琨, 李 兵

(武汉大学中南医院泌尿外科, 湖北武汉 430071)

Preliminary exploration of the correlation between GEARS score and objective operative metrics in robot-assisted surgical training

XIAO Xingyuan, LIU Zhenyu, DU Yuanzhuo, YANG Kun, LI Bing

(Department of Urology, Zhongnan Hospital of Wuhan University, Wuhan 430071, China)

ABSTRACT: **Objective** To analyze correlations between intraoperative digital operative metrics and GEARS scores in robotic urological surgery training, and screen objective skill assessment indicators, so as to provide references for standardized robotic surgical training. **Methods** A total of 15 untrained urologists receiving robotic training were included. Instrument position, out-of-view, clutch, camera and switching metrics were extracted via video recognition and trajectory tracking. Seven experts conducted double-blind GEARS scoring. Spearman correlation and partial correlation analyses adjusted for age and total laparoscopic surgical volume were carried out. **Results** Ten metrics were significantly correlated with GEARS scores without adjustment. The minimum distance to the right central point was strongly positively correlated ($r_s = 0.760, Q = 0.010$), whereas metrics of instrument out-of-view, clutch, camera and switching frequency/duration showed negative correlations. After age adjustment, four metrics retained independent associations, and average switching time had the strongest negative correlation ($r_s = -0.749, Q = 0.020$). No significant correlations remained after adjusting for total laparoscopic volume alone or both confounders (all $Q > 0.05$). **Conclusion** Metrics of instrument positioning, field maintenance, clutch and arm switching can supplement subjective GEARS scoring. Previous laparoscopic experience is a key confounder. This small-sample exploratory study requires large multicenter verification; the objective metrics offer targeted training feedback to identify trainees' operational weaknesses.

KEY WORDS: robot-assisted surgery; surgical training; GEARS score; operative metrics; correlation analysis

摘要: **目的** 探讨泌尿外科机器人辅助手术培训场景下,术中数字化操作指标与 GEARS 评分的相关性,筛选可用于手术技能评价的客观指标,为机器人手术规范化教学与技能评估提供参考。**方法** 纳入 15 名初次接受机器人手术培训的泌尿外科医师,通过视频图像识别与轨迹追踪技术提取器械空间位置、器械脱离视野、离合、镜头及切换操作等量化指标;采用 7 名专家双盲法完成 GEARS 评分。运用 Spearman 秩相关分析各指标与 GEARS 评分的关联,并进一步校正年龄、总腹腔镜手术量混杂因素开展偏相关分析。**结果** 未校正混杂因素时,共 10 项指标与 GEARS 评分显著相关;其中距离右侧中心点最近距离呈强正相关($r_s = 0.760, Q = 0.010$),器械脱离视野、离合、镜头、器械切换相关频次、时长指标均呈显著负相关。仅校正年龄后,距离右侧中心点最近距离、器械脱离视野总次数、切换使用总时长、切换使用平均时间仍存在独立统计学关联,切换使用平均时间负相关性最强($r_s = -0.749, Q = 0.020$);单独校正总腹腔镜手术量、联合校正两项混杂因素后,全部指标校正 Q 值均 > 0.05 ,原有相关关系消失。**结论** 器械空间定位稳定性、术野维持、离合与器械切换相关数字化指标可作为 GEARS 主观评分的补充客观评价依据;既往腹腔镜手术经验是二者关联的重要混杂因素。本研究为小样本初步探索,结论有待大样本、多中心研究进一步验证,提取的客观指标可用于手术培训精细化反馈,精准定位学员操作缺陷。

关键词: 机器人辅助手术;手术培训;GEARS 评分;操作指标;相关性分析

中图分类号:G642.0

文献标志码:A

DOI:10.12483/j.issn.1009-8291.202606051

CSTR:32408.14.j.CN61-1374.202606051

机器人辅助手术已在泌尿外科多种复杂手术中

得到应用,规范化培训和技能评价是其安全推广的重要环节^[1-2]。

收稿日期:2026-06-15

修回日期:2026-08-08

通信作者:李兵,主任医师,教授,博士生导师。

E-mail:bingli2023@whu.edu.cn

作者简介:肖行远,主任医师,副教授,博士生导师。研究方向:尿路修复重建与泌尿系肿瘤。

E-mail:xiaoxingyuan@whu.edu.cn

与传统腹腔镜手术相比,机器人手术在视野呈现、器械活动自由度、主从式控制、离合使用及双手协同等方面具有不同特点。既往微创手术经验虽可为机器人手术学习提供基础,但并不能完全反映术者在

机器人平台下的具体操作表现。因此,在培训过程中,除评价术者总体手术技能外,进一步观察其具体操作行为,对于了解机器人手术学习特点具有一定意义。

机器人技能综合评估量表(Global Evaluative Assessment of Robotic Skills, GEARS)是目前机器人手术技能评价中较为常用的工具,可从深度感知、双手灵活性、效率、力量敏感性、自主性和机器人控制等方面评价术者整体表现^[3-4]。但 GEARS 评分主要基于评价者对操作过程的综合判断,难以进一步区分评分差异与哪些具体操作行为有关。既往研究尝试通过手术录像或机器人系统数据提取客观指标,用于描述器械运动、镜头控制、操作效率及特定步骤中的技术差异^[5-6]。

本研究基于泌尿外科机器人手术培训录像,以 GEARS 评分作为参照,提取器械空间位置、器械脱离视野、离合使用、镜头使用及切换操作等数字化指标,分析其与 GEARS 评分之间的相关性。本研究旨在初步观察培训场景下不同数字化指标与 GEARS 评分的关系,筛选可能与操作评分相关的行为特征,为后续开展标准化训练任务和较大样本研究提供依据。

1 方法

1.1 培训方案 每期培训周期为 7 d。课程内容按照阶梯式递进原则设置,依次包括:模拟器(基础)、模拟器(进阶)、实机 Dry-lab、实机 Wet-lab、团队 Wet-lab、团队动物实验及团队动物考核。前 4 d 为单人课程,由担任主刀医生的学员独立完成。自第 5 d 上午起,团队课程中的第一助手与主刀医生加入培训,共同完成第 6~7 d 的动物操作及手术模拟操作。动物采用 42 kg 广西巴马小型猪。动物手术考核术式为肾门分离术及输尿管离断-吻合术中的一种或两种。本研究已通过武汉大学中南医院医学伦理委员会(编号:2023114K)和湖北省预防医学科学院湖北省疾病预防控制中心实验动物管理与使用委员会(编号:202660041)审批。

1.2 研究对象 2023 年 5 月—2026 年 1 月共有 15 名泌尿外科医生在武汉大学手术机器人培训中心进行了系统化培训。

1.3 GEARS 评分 本研究采用国际通用标准化机器人手术技能评估量表 GEARS 完成操作评分,量表包含深度感知、双手灵活性、操作效率、力量敏感性、操作自主性、机器人控制,共 6 个评价维度,每个维度分值 1~5 分,总分范围 6~30 分,得分越高代表机器

人手术综合操作水平越好。培训结束后,从 12 位专家中随机挑选 7 位,对所有动物手术录像进行双盲评估。

①专家赴指定地点,单独进行评估;

②评估过程中关闭手机,由专门工作人员播放操作录像,每位学员的评审时间不少于 20 min;

③评估量表使用标准 GEARS 量表,总分为 30 分;

④评估结束后,去掉一个最高分和一个最低分,统计剩余 5 位专家的评估成绩,取平均值得到最终的成绩。

1.4 操作指标提取 本研究首先对机器人手术视频进行标准化处理,将连续术野视频按时间序列转换为可分析的图像帧,并统一帧率(60 fps)、分辨率(1920×1080)及图像尺度,以减少不同来源视频在采集参数上的差异。随后,基于训练好的器械识别模型[基于 Real-Time Detection Transformer(RT-DETR)架构;平均精度均值(mean average precision, MAP): 0.829, Precision: 0.813, Recall: 0.827, dice: 0.820),预处理后的手术视频画面被拆分为左、右机械臂对应的独立术野分区,分别以各分区画面像素几何中心为基准,对每一帧图像中的目标器械区域进行自动识别与定位,获得器械在术野中的空间位置。为进一步量化器械运动特征,研究从识别结果中提取器械关键部位的二维坐标,并按照视频时间顺序连续记录,形成器械运动轨迹。基于该轨迹数据,可进一步计算器械位移、运动速度、运动范围、路径稳定性及操作节律等指标,用于描述机器人手术过程中器械操作行为的时空特征,并为后续手术技能评估、操作模式分析及人机协同研究提供量化依据(图 1)。最终,本研究提取器械运动轨迹、器械脱离视野、离合、镜头与器械切换等相关数字化指标用于后续研究。

1.5 数据分析 采用 R 4.4.3 软件进行统计学分析。计数资料以频率和百分比表示,计量资料以均值±标准差表示。结合本研究样本量偏小($n=15$)、部分术中操作量化指标分布不符合正态分布的特点,统一选用 Spearman 非参数相关体系开展关联性分析进行探索。首先采用 Spearman 秩相关分析评价 GEARS 评分与各项数字化操作指标之间的关联;随后分别控制年龄、总腹腔镜手术量,以及同时控制年龄和总腹腔镜手术量,进行偏相关分析。各模型内对多重检验的 P 值分别采用 Benjamini-Hochberg 法进行错误发现率校正,以 $Q<0.05$ 为差异有统计学意义。

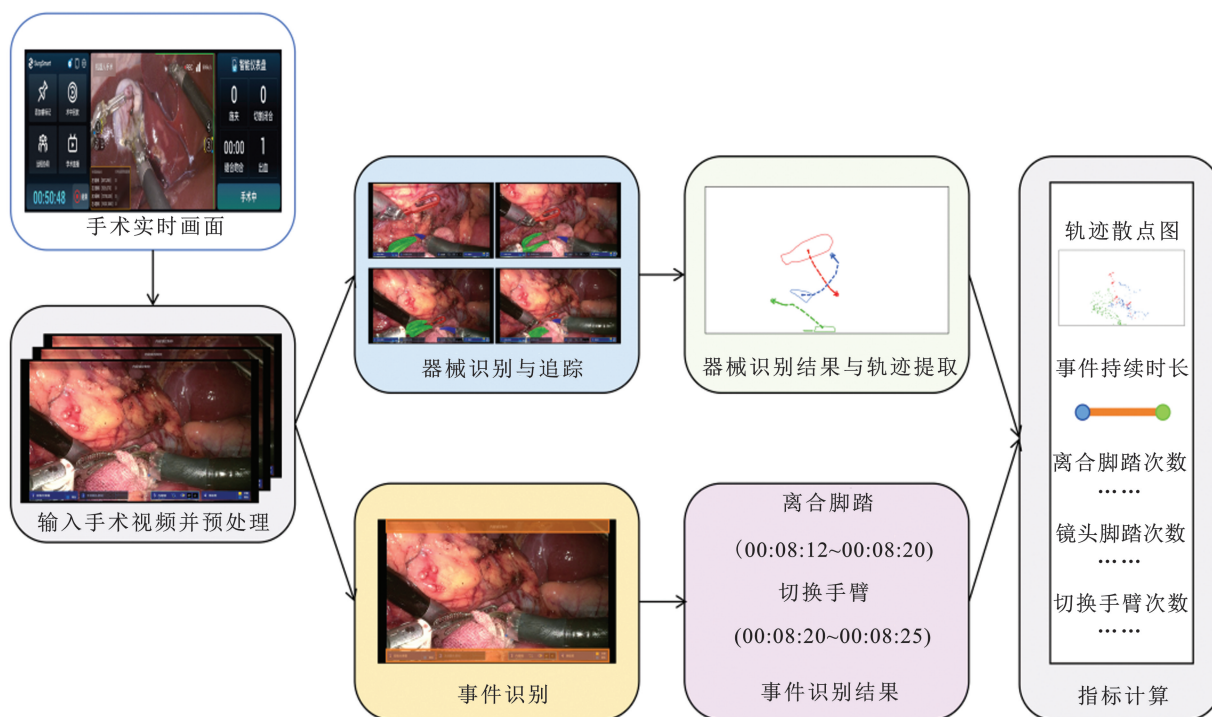


图1 操作指标提取流程图

2 结果

2.1 参加培训人员的基线资料 本研究共纳入 15 名泌尿外科医师,均为男性,年龄 33~55 岁,均值

(43.27±6.58)岁;腹腔镜手术开展年份 2~23 年,均值(11.27±5.04)年;年腹腔镜手术量 5~300 例,均值(149±87)例;总腹腔镜手术量 50~4800 例,均值(1813±1519)例(表 1)。

表1 参与培训人员的基线资料

[n=15,例(%)]

变量	数值	变量	数值
年龄*	43.27±6.58	腹腔镜手术年数*	11.27±5.04
性别		年腹腔镜手术数量*	149±87
男	15(100)	总腹腔镜手术数量*	1813±1519
女	0(0)	术式	
职称		肾门分离	2(13.3)
主治医师	4(26.7)	肾门分离+输尿管吻合	12(80.0)
副主任医师	7(46.7)	输尿管吻合	1(6.7)
主任医师	4(26.7)	手术时长/s*	4 362.53±3 245.51

*: $\bar{x} \pm s$ 。

2.2 GEARS 评分与数字化操作指标的相关性分析

2.2.1 GEARS 评分与手术操作指标的 Spearman 秩相关分析 本研究首先采用 Spearman 秩相关分析,分析 GEARS 评分与机器人手术各项操作指标的相关性,结果见表 2。经过错误发现率(false discovery rate,FDR)校正后,空间距离指标方面,仅距离右侧中心点最近距离呈强正相关($r_s = 0.760, Q = 0.010$),其余距离类指标均无统计学相关性。器械操作指标与 GEARS 评分关联更为突出。器械脱离视

野总次数($r_s = -0.735, Q = 0.010$)、离合使用总次数($r_s = -0.736, Q = 0.010$)及总时长($r_s = -0.735, Q = 0.010$)均呈强负相关;镜头使用总次数($r_s = -0.662, Q = 0.023$)、总时长($r_s = -0.692, Q = 0.017$)也表现为显著负相关。除了切换每分钟使用平均时长,切换操作其余指标均存在统计学关联,其中切换使用总时长相关性最强($r_s = -0.744, Q = 0.010$)。离合、镜头对应的部分平均频次、平均时长指标则未表现出显著相关性。

表 2 GEARS 评分与各操作指标相关性分析结果

操作指标	模型 1 简单相关				模型 2 校正年龄				模型 3 校正腹腔镜手术量				模型 4 校正腹腔镜手术量及年龄			
	样本统计量		模型 1 简单相关		样本统计量		模型 2 校正年龄		样本统计量		模型 3 校正腹腔镜手术量		样本统计量		模型 4 校正腹腔镜手术量及年龄	
	(r_s)	P 值	Q 值		(r_s)	P 值	Q 值		(r_s)	P 值	Q 值		(r_s)	P 值	Q 值	
距离左侧中心点最远距离	-0.269	0.333	0.400													
距离左侧中心点最近距离	0.386	0.155	0.310													
距离左侧中心点左方最远距离	-0.358	0.190	0.342													
距离左侧中心点右方最远距离	-0.583	0.023	0.058													
距离左侧中心点上方最远距离	-0.339	0.216	0.360													
距离左侧中心点下方最远距离	-0.115	0.684	0.708													
距离左侧中心点平均距离	-0.147	0.601	0.644													
距离右侧中心点最远距离	-0.545	0.036	0.083													
距离右侧中心点最近距离	0.760	0.001	0.010		0.646	0.013	0.038		0.412	0.144	0.245		0.408	0.167	0.369	
距离右侧中心点左方最远距离	-0.227	0.416	0.480													
距离右侧中心点右方最远距离	-0.305	0.269	0.391													
距离右侧中心点上方最远距离	-0.308	0.264	0.391													
距离右侧中心点下方最远距离	-0.528	0.043	0.092													
距离右侧中心点平均距离	0.072	0.800	0.800													
器械脱离视野的总次数	-0.735	0.002	0.010		-0.634	0.015	0.038		-0.435	0.120	0.245		-0.476	0.100	0.369	
离合使用总次数	-0.736	0.002	0.010		-0.561	0.037	0.056		-0.409	0.147	0.245		-0.329	0.272	0.369	
离合每分钟使用平均次数	-0.292	0.290	0.391													
离合使用总时长	-0.735	0.002	0.010		-0.556	0.039	0.056		-0.418	0.137	0.245		-0.347	0.246	0.369	
离合每分钟使用平均时长	-0.290	0.294	0.391													
离合使用平均时间	-0.355	0.194	0.342													
镜头使用总次数	-0.662	0.007	0.023		-0.445	0.111	0.111		-0.124	0.673	0.673		-0.254	0.402	0.447	
镜头每分钟使用平均次数	-0.147	0.601	0.644													
镜头使用总时长	-0.692	0.004	0.017		-0.525	0.054	0.068		-0.202	0.489	0.543		-0.334	0.265	0.369	
镜头每分钟使用平均时长	-0.287	0.300	0.391													
镜头使用平均时间	-0.273	0.326	0.400													
切换使用总次数	-0.736	0.002	0.010		-0.574	0.032	0.056		-0.300	0.297	0.424		-0.315	0.295	0.369	
切换每分钟使用平均次数	-0.614	0.015	0.045		-0.446	0.110	0.111		-0.227	0.435	0.543		-0.205	0.501	0.501	
切换使用总时长	-0.744	0.001	0.010		-0.658	0.011	0.038		-0.461	0.097	0.245		-0.440	0.132	0.369	
切换每分钟使用平均时长	-0.596	0.019	0.052													
切换使用平均时间	-0.682	0.005	0.019		-0.749	0.002	0.020		-0.657	0.011	0.110		-0.615	0.025	0.250	

2.2.2 校正年龄的偏相关分析 为排除年龄的混杂效应,将上述分析中 $Q < 0.05$ 的指标纳入研究,开展校正年龄的偏相关分析,结果见表 2。空间距离指标中,距离右侧中心点最近距离仍维持显著正相关($r_s = 0.646, Q = 0.038$),为独立相关指标。器械操作指标中,校正年龄后共 3 项指标与 GEARS 评分存在独立显著负相关。其中切换使用平均时间相关性最强($r_s = -0.749, Q = 0.020$),相关强度较初次分析有所提升。器械脱离视野总次数、切换使用总时长,均保持显著的独立负相关。离合使用总次数、离合使用总时长、镜头使用总次数及总时长、切换使用总次数以及切换每分钟平均使用次数在校正后无统计学意义,表明其原有相关关系受年龄混杂影响。

2.2.3 校正总腹腔镜手术量的偏相关分析 仅校正总腹腔镜手术量单一协变量后,所有指标经 FDR 多重比较校正后 Q 值均 > 0.05 ,无任何数字化操作指标与 GEARS 评分存在稳定显著关联。其中切换使用平均时间原始 $P = 0.011$,但校正后 $Q = 0.110$,未达统计学显著性标准。

2.2.4 同时校正年龄与总腹腔镜手术量的偏相关分析 同时纳入年龄与总腹腔镜手术量作为双协变量校正后,切换使用平均时间原始 $P = 0.025$,但经 FDR 校正后 $Q = 0.250$,其余指标校正后 Q 值均 > 0.05 ,均未达统计学显著标准。

3 讨论

机器人辅助手术已成为泌尿外科微创治疗的重要方式之一,随之而来的培训质量和技能评价也受到越来越多关注^[9-10]。GEARS 量表作为目前机器人手术技能评价中应用较多的工具,能够从深度感知、双手灵活性、效率、力量敏感性、自主性及机器人控制等方面对术者进行整体评价^[4,11-12]。但在实际培训中,单纯依靠综合评分往往难以进一步说明术者在哪些具体操作环节存在不足。本研究基于手术视频提取的客观操作指标,通过无校正粗相关、分别校正年龄、总腹腔镜手术量及双协变量联合校正四组分析,并采用 FDR 多重比较校正控制假阳性,系统量化器械运动、器械脱离视野、离合、镜头、器械切换指标与机器人手术技能 GEARS 评分的关联规律,结果呈现清晰的分层差异。

未校正混杂因素时,经 FDR 校正后,共 10 项数字化操作指标与 GEARS 评分存在统计学相关性,主要表现为 GEARS 评分较高者器械脱离视野、离合、

镜头调节及器械切换的总次数或总时长较少。校正年龄后,距离右侧中心点最近距离、器械脱离视野总次数、切换使用总时长及切换使用平均时间仍与 GEARS 评分存在统计学相关性,提示这些指标与技能评分之间的关系可能不仅由年龄差异所解释。其中,切换使用平均时间的相关系数绝对值较大,但由于本研究样本量有限,尚不能据此认定其为稳定或具有独立预测价值的评价指标。在进一步校正总腹腔镜手术量后,上述指标的 FDR 校正 Q 值均未达到统计学显著性;同时校正年龄和总腹腔镜手术量后亦得到相似结果。这提示既往腹腔镜手术经验可能参与数字化操作指标与 GEARS 评分之间的关联。不过,显著性消失也可能与样本量较小、年龄与手术经验之间的相关性以及双协变量模型统计效能不足有关。因此,本研究尚不能确定总腹腔镜手术量是导致相关性变化的唯一或主要混杂因素。

相比 GEARS 评分,这类数字化指标更容易被培训学员理解,也更容易转化为训练反馈。例如,器械频繁脱离视野、反复离合或频繁调整镜头,往往意味着操作过程被多次打断,操作连续性和空间控制能力可能不足。既往研究表明,基于手术录像的技术分析能够提示机器人手术中的操作差异,并为培训和质量改进提供依据^[5,15]。此外,视频回顾和视频教学在机器人手术培训中的作用也受到关注,相关研究显示,其可用于促进术者自我反思、技能改进和培训过程评价^[16-18]。本研究结果与上述观点基本一致,即数字化指标并不是替代专家评价,而是为专家评分提供更细化的操作指标。

器械脱离视野是本研究中较值得关注的指标之一。机器人手术依赖三维视野下的器械定位和目标识别,既往研究显示,立体视觉有助于提高机器人辅助手术任务完成的效率和准确性^[19]。若器械反复离开视野范围,可能影响操作的连续性,也可能增加术者重新定位器械和目标组织的时间,近年来,基于手术器械追踪的研究为此类操作行为的量化分析提供了技术基础^[6-8,20-21]。本研究中,校正年龄因素后,器械脱离视野次数仍与 GEARS 评分存在相关性,提示该指标可作为观察机器人操作稳定性的潜在指标。不过,器械脱离视野并不一定都代表操作失误,也可能与主动调整操作角度、扩大牵拉范围或更换操作区域有关。因此,后续研究有必要结合具体手术步骤和视频标注,对该指标进行更详细的解释。

离合和切换操作也是机器人手术中较有代表性

的功能性动作。离合是机器人手术中调整主手位置和维持合适操作姿态的重要动作,已有研究尝试基于离合使用情况评价机器人手术技能^[22]。因此,本研究提示,在相同或相近任务条件下,频繁离合或切换可能反映术者在初始操作规划、器械活动范围判断和双手协调方面仍不够熟练。本研究中,在未校正模型中,离合使用总次数和总时长与 GEARS 评分呈负相关;但校正年龄、总腹腔镜手术量并进行 FDR 校正后,两项指标均未达到统计学显著性,提示其关联可能受到年龄、总腹腔镜手术量或其他未测量因素影响。切换使用总时长和平均时间在年龄校正后仍保持统计学相关,但在进一步校正总腹腔镜手术量后不再显著。CHEN 等^[6]和 HUNG 等^[7]的研究显示,自动化指标可用于描述机器人手术训练或特定操作步骤中的技术差异;其他研究进一步提示,不同专业水平术者、专家与高水平专家之间,也可在部分自动化指标上表现出差异^[23-25]。但这类指标的解释必须依赖具体任务背景,不能脱离操作内容单独判断术者水平。

镜头使用相关指标与 GEARS 评分的关系相对复杂。镜头控制是机器人手术的重要组成部分,良好的镜头控制有助于维持清晰、稳定和适宜的操作视野。JARC 等^[8]指出,内镜控制相关指标可用于评价机器人手术操作表现,但其意义会受到观察视角和任务场景的影响。本研究中,镜头使用总次数和总时长在未校正模型中与 GEARS 评分呈负相关;但校正年龄、总腹腔镜手术量并进行 FDR 校正后,相关性均未达到统计学显著性。因此,本研究尚不能确认镜头使用频率或时长与 GEARS 评分之间存在独立关联。但镜头调整并非单纯的无效动作,在复杂解剖区域或精细操作过程中,主动调整镜头有时是为了获得更安全的视角。因此,后续应进一步分析镜头移动幅度、目标区域居中情况、有效视野维持时间等内容,可能有助于判断镜头控制能力与机器人操作水平之间的关系。

本研究的意义在于初步提示,数字化操作指标可以从另一个层面对机器人手术技能进行描述,为 GEARS 评分提供补充信息。对于机器人手术培训而言,综合评分可以说明总体表现如何,而数字化指标则可能进一步提示问题出在哪里,如器械控制、视野维持、离合使用或操作连续性等方面。需要指出的是,本研究仍存在一定局限。首先,样本量较小,结果稳定性需要进一步验证;其次,研究基于培训录像,尚

不能直接应用于临床手术;再次,本研究主要分析相关性,不能说明数字化指标与 GEARS 评分之间存在因果关系;此外,术者既往腹腔镜经验和操作时长等因素均可能影响结果。未来应在标准化训练任务和更大样本基础上,进一步验证相关指标的重复性及其与临床手术表现之间的关系。

参考文献:

- [1] 周君,江爱民,王林辉,等.泌尿外科机器人辅助手术培训的思考与展望[J]. 机器人外科学杂志(中英文),2026,7(3):498-502.
- [2] VOLPE A, AHMED K, DASGUPTA P, et al. Pilot validation study of the European Association of Urology robotic training curriculum[J]. Eur Urol, 2015, 68(2):292-299.
- [3] 梁渝靖,崔皓鑫,李阳辉,等.基于达芬奇手术模拟器评分的机器人技能全面评价预测模型研究[J]. 机器人外科学杂志(中英文),2025,6(6):1031-1035.
- [4] GOH A C, GOLDFARB D W, SANDER J C, et al. Global evaluative assessment of robotic skills: validation of a clinical assessment tool to measure robotic surgical skills[J]. J Urol, 2012, 187(1):247-252.
- [5] WU R C, PREBAY Z J, PATEL P, et al. Using video review to understand the technical variation of robot-assisted radical prostatectomy in a statewide surgical collaborative[J]. World J Urol, 2020, 38(7):1607-1613.
- [6] CHEN J, OH P J, CHENG N, et al. Use of Automated performance metrics to measure surgeon performance during robotic vesicourethral anastomosis and methodical development of a training tutorial[J]. J Urol, 2018, 200(4):895-902.
- [7] HUNG A J, CHEN J, JARC A, et al. Development and validation of objective performance metrics for robot-assisted radical prostatectomy: a pilot study[J]. J Urol, 2018, 199(1):296-304.
- [8] JARC A M, CURET M J. Viewpoint matters: objective performance metrics for surgeon endoscope control during robot-assisted surgery[J]. Surg Endosc, 2017, 31(3):1192-1202.
- [9] MAZZONE E, MISTRETTA F A, KNIPPER S, et al. Contemporary National Assessment of Robot-Assisted Surgery Rates and Total Hospital Charges for Major Surgical Uro-Oncological Procedures in the United States[J]. J Endourol, 2019, 33(6):438-447.
- [10] MOTTRIE A, MAZZONE E, WIKLUND P, et al. Objective assessment of intraoperative skills for robot-assisted radical prostatectomy (RARP): results from the ERUS Scientific and Educational Working Groups Metrics Initiative[J]. BJU Int, 2021, 128(1):103-111.
- [11] AGHAZADEH M A, JAYARATNA I S, HUNG A J, et al. External validation of Global Evaluative Assessment of Robotic Skills (GEARS)[J]. Surg Endosc, 2015, 29(11):3261-3266.

- [12] SANCHEZ R, RODRIGUEZ O, ROSCIANO J, et al. Robotic surgery training: construct validity of Global Evaluative Assessment of Robotic Skills (GEARS)[J]. J Robot Surg, 2016, 10(3):227-231.
- [13] WANG Y, WILDER S, HIJAZI M, et al. Surgeon skill and perioperative outcomes in robot-assisted partial nephrectomy [J]. JAMA Netw Open, 2024, 7(7):e2421696.
- [14] OH P J, CHEN J, HATCHER D, et al. Crowdsourced versus expert evaluations of the vesico-urethral anastomosis in the robotic radical prostatectomy: is one superior at discriminating differences in automated performance metrics? [J]. J Robot Surg, 2018, 12(4):705-711.
- [15] FAINBERG J, VANDEN BERG R N W, CHESNUT G, et al. A novel expert coaching model in urology, aimed at accelerating the learning curve in robotic prostatectomy[J]. J Surg Educ, 2022, 79(6):1480-1488.
- [16] VAN DER LEUN J A, SIEM G, MEIJER R P, et al. Improving robotic skills by video review [J]. J Endourol, 2022, 36(8):1126-1135.
- [17] SOLIMAN M M, SOLIMAN M K. How expert surgeons review robotic videos: a grounded theory study[J]. Am J Surg, 2023, 226(5):709-716.
- [18] LARKINS K, KHAN M, MOHAN H, et al. A systematic review of video-based educational interventions in robotic surgical training[J]. J Robot Surg, 2023, 17(4):1329-1339.
- [19] MUNZ Y, MOORTHY K, DOSIS A, et al. The benefits of stereoscopic vision in robotic-assisted performance on bench models[J]. Surg Endosc, 2004, 18(4):611-616.
- [20] LEE D, YU H W, KWON H, et al. Evaluation of surgical skills during robotic surgery by deep learning-based multiple surgical instrument tracking in training and actual operations[J]. J Clin Med, 2020, 9(6):1964.
- [21] CHEN J, CHENG N, CACCIAMANI G, et al. Objective assessment of robotic surgical technical skill: a systematic review [J]. J Urol, 2019, 201(3):461-469.
- [22] LI L, CHEN Z, ZAW T H H, et al. Skill assessment based on clutch use in cross-platform robot-assisted surgery [J]. Surg Endosc, 2024, 38(8):4336-4343.
- [23] HUNG A J, OH P J, CHEN J, et al. Experts *vs.* super-experts: differences in automated performance metrics and clinical outcomes for robot-assisted radical prostatectomy[J]. BJU Int, 2019, 123(5):861-868.
- [24] CHEN A, GHODOUSSIPOUR S, TITUS M B, et al. Comparison of clinical outcomes and automated performance metrics in robot-assisted radical prostatectomy with and without trainee involvement[J]. World J Urol, 2020, 38(7):1615-1621.
- [25] NGUYEN J H, CHEN J, MARSHALL S P, et al. Using objective robotic automated performance metrics and task-evoked pupillary response to distinguish surgeon expertise[J]. World J Urol, 2020, 38(7):1599-1605.

(编辑 王 玮)