

# 外科医师非手术临床技能评分标准中 导师评价的研究

秦骏<sup>1</sup> 史苑<sup>1</sup> 敬然<sup>1</sup> 俞旻皓<sup>1</sup> 陈建军<sup>1</sup> 张斌<sup>1</sup> 顾磊<sup>1</sup> 蒋春晖<sup>1</sup> 叶光耀<sup>1</sup>  
孙隆慈<sup>1</sup> Jay N Shah<sup>2</sup> 钟鸣<sup>1</sup>

<sup>1</sup>上海交通大学医学院附属仁济医院胃肠外科 200127; <sup>2</sup>帕坦医科大学附属帕坦医院  
44700, 尼泊尔

通信作者: 钟鸣, Email: drzhongming1966@163.com

**【摘要】** 目的 非手术技能(non-technical skills, NTS)是考察外科医师综合素质的必需条目之一。本研究提出 C-NTS 的概念,即基于情景化(如病史采集)和录像的非手术临床技能的评分标准,并验证其实际操作效果。**方法** 研究对象分为导师组和学员组;导师组含三甲医院胃肠外科高年资主治医师 4 人,均为男性,博士学位;学员组随机选择 2018 年同科室 4 名外科基地轮转医生,男性 2 人,女性 2 人。导师在培训讨论前后分别参照标准对同一匿名视频评分。应用 SPSS 22.0,采用单因素方差分析比较组间及组内评分差异,用 Kendall 协和系数检验评分一致性。**结果** 导师接受培训前后,其评分的 Kendall 协和系数由 0.425 增长为 0.853,后者同时提示差异有统计学意义( $P=0.017$ ),证实了培训的有效性。**结论** C-NTS 初步解决了非手术临床技能考核同质化的难点。若组织临床导师展开 C-NTS 相应讨论及接受培训,可能有利于临床实际教学和评价中节省时间与人力。

**【关键词】** 非手术临床技能; 导师培训; 情景化; 病史采集; 住院医师规范化培训

**【中图分类号】** R6

**基金项目:** 国家自然科学基金青年项目(81700532); 上海市科学技术委员会科研项目(19411966200)

DOI: 10.3760/cma.j.issn.2095-1485.2019.10.018

## Tutor evaluation on the scoring criterion for non-technical skills in surgeons

Qin Jun<sup>1</sup>, Shi Yuan<sup>1</sup>, Jing Ran<sup>1</sup>, Yu Minhao<sup>1</sup>, Chen Jianjun<sup>1</sup>, Zhang Bin<sup>1</sup>, Gu Lei<sup>1</sup>, Jiang Chunhui<sup>1</sup>, Ye Guangyao<sup>1</sup>, Sun Longci<sup>1</sup>, Jay N Shah<sup>2</sup>, Zhong Ming<sup>1</sup>

<sup>1</sup>Department of Gastrointestinal Surgery, Ren Ji Hospital, School of Medicine, Shanghai Jiao Tong University, Shanghai 200127, China; <sup>2</sup>Department of Surgery, Patan Hospital, Patan Academy of Health Sciences, Lalitpur 44700, Nepal

Corresponding author: Zhong Ming, Email: drzhongming1966@163.com

**【Abstract】 Objective** Non-technical skills (NTS) are necessary to evaluate the comprehensive quality of surgeons. In this study, we proposed the concept of C-NTS, a scoring criterion for NTS based on real scenarios (for example, history taking) and video recording, and verified its practical application effects. **Methods** Study objects were divided into the tutor group and the student group. The tutor group contained four senior attending physicians in gastrointestinal surgery department of one tertiary hospital (all male with doctor degree). The student group had four rotating surgeons who were randomly selected from the same department in 2018 (two males and two females). Before and after the training, the tutor rated the same anonymous video by C-NTS. One-factor analysis of variance was used to compare the differences between and within the groups, and Kendall concordant coefficient was used to test the consistency by SPSS 22.0. **Results** After the tutor receiving training, Kendall concordant coefficient was increased from 0.425 to 0.853 and the latter suggested the difference had statistical significance ( $P=0.017$ ), which confirmed the effectiveness of the training. **Conclusion** C-NTS has preliminarily solved the difficulty

of homogenization of NTS assessment. Relevant C-NTS discussions and trainings for clinical tutors may be beneficial to save time and manpower in clinical teaching and evaluation.

**【Key words】** Non-technical skills; Tutor training; Real scenarios; History taking; Standardized residents training

**Fund program:** National Natural Science Foundation Youth Project (81700532); Science and Research Project of Shanghai Science and Technology Committee (19411966200)

DOI: 10.3760/cma.j.issn.2095-1485.2019.10.018

非手术临床技能(non-technical skills, NTS)是外科医师用以辅助手术技能来获得更佳成果的认知、社会交往和个人资源的能力, 一般包括情境意识、决策、沟通、团队合作及领导力<sup>[1]</sup>。目前评价 NTS 的常用量表主要有两种: 一种是 NOTSS (nontechnical skills for surgeon), 以被评估者的操作是否最终影响患者的安全结局及其程度为评分标准; 另一种则是 OSANTS (objective structured assessment of non-technical skills), 着重描述操作本身, 两者在模拟及现实环境中均有较好应用<sup>[2-3]</sup>。就我国复杂的医疗环境和多元化的医师成长体系, 如何使量表融入实际临床环境、如何提高导师评分原则的一致性, 在目前的 NTS 教育研究中, 仍缺乏详尽准确的报道。

本研究在对外科基地医师的非手术临床技能培训中运用了基于情景化与录像模式的 NTS 评价方法(C-NTS), 通过对导师培训前后的评分比较, 探讨了 C-NTS 评分表的信效度及首次评估了导师培训效果的一致性。

## 1 对象与方法

### 1.1 研究对象

研究涉及对象分为导师组和学员组。导师组为某三甲医院胃肠外科工作超过 8 年的高年资主治医师 4 名, 均为男性, 博士学历。学员组则随机选择 2018 年 9 月至 2018 年 12 月同科室 4 名外科基地轮训医生(轮转第二年), 男性 2 人、女性 2 人。

### 1.2 研究方法

量表的设计采用 NOTSS 的结构框架和 OSANTS 的描述形式。结合我国现况, 为体现医患沟通技巧的重要性, 量表的设计中除了填充不同情景的特点, 还增加了同样适用于录像评分的 ORCA (operating room communication assessment) 量表, 将操作者与患者交谈时的语音、口齿清晰度、措辞及理解力和肢体

语言纳入考核范围, 组成 C-NOTSS 量表<sup>[4]</sup>。总体安全结局评价(global safety evaluation, GSE)的加入则重新引入了患者安全评估。通过整合归纳, 最终得到 6 大类共计 16 项的 C-NTS 量表。本次选用的情景模式为患者入院询问病史阶段(表 1)。

经充分向学员本人告知并征得患者同意后, 研究采用视频录像的形式全程记录学员独立询问病史、体格检查及口头交班(包含初步诊断、鉴别诊断、诊疗方案等情况)。由专人收集视频、匿名编号并打包发送至全部导师的邮箱, 导师根据经典 NOTSS 标准对每位学员进行评分。在首轮计分后, 导师组讨论并重新界定各条目, 修改定稿后形成新的 C-NTS 标准, 再分别对同一视频打分。

### 1.3 统计学方法

采用 SPSS 22.0 分析数据, 用单因素方差分析比较组间及组内评分差异, 用 Kendall 协和系数检验导师打分的一致性。检验水准  $\alpha=0.05$ 。

## 2 结果

导师组在培训讨论前后对学员的 C-NOTSS 打分结果汇总如图 1 所示, 横坐标为 4 名学员(前后两次评分), 纵坐标为总分, 每条曲线代表一位导师评分情况。直观可见, 4 位导师对同一学员的第二轮评分相较第一轮评分, 更为接近。导师组对同一学员评分汇总后显示(表 2), 4 名学员平均得分均有变化; 培训后, 学员 1、2、4 平均得分的标准差下降。

将单个导师给出的 C-NOTSS 分数归为一组比较组间差异, 培训讨论前导师评分组间差异有统计学意义( $P<0.01$ ), 培训讨论后评分组间差异无统计学意义( $P>0.05$ )。培训前后协和系数由 0.425 增长为 0.853, 后者同时提示差异有统计学意义( $P<0.05$ )。计算量表 Cronbach's  $\alpha=0.77$ , 可以认为条目间一致性较好(表 3)。

表 1 病史采集情景下的 C-NTS 量表示例

C-NOTSS项目与内涵		评分分级			
情境意识	收集信息	询问病史,体格检查,开具辅助检查项目	评分标准说明 <sup>a</sup>		
	理解信息	综合各项结果,判断是否符合预期	A标准	B标准	C标准
	预测未来的状态	初步诊断,判断紧急程度、最佳治疗手段和预后情况	准备充分,询问病史、体格检查及开具辅助检查完整了解患者情况,作出初步诊断,常规预测病情紧急程度,预估治疗手段和预后情况	准备充分,询问病史、体格检查及开具辅助检查,有时难以了解患者情况作出初步诊断,或难以预测病情紧急程度,预估治疗手段和预后情况	准备不足,无法了解患者情况作出初步诊断,反复遗漏对病情紧急程度、预估治疗手段和预后情况的预测,正在面临可预见的困难
决策	收集决策选择	鉴别诊断,征求团队其他成员的意见,如有需要,请求上级医师协助	征求团队其他成员意见,清晰迅速地明确诊断及依据,作出鉴别诊断,开具医嘱,在发现病情变化时回顾结局并当机立断地更正方案	征求团队其他成员意见,明确诊断及依据,作出鉴别诊断,开具医嘱,必要时能够回顾结局并更正方案,但偶尔表现出犹豫不决或不确定	无法明确诊断及依据或未征求团队其他成员意见,难以作出鉴别诊断或开具医嘱,当发现病情变化后不能回顾结局且维持原有方案
	确定决策并沟通	补充必要检查,明确诊断依据			
	实施和回顾决策	开具医嘱,根据病情变化或意外情况即时调整			
沟通和团队合作	交换信息	罗列所有治疗方案和各自优劣	罗列所有治疗方案,并积极鼓励患者讲述需求,确认最终治疗方案后完成术前谈话及手术签字,如有变动及时组织讨论,主动协助护士及患者家属完善术前准备事项	罗列所有治疗方案,听取患者需求,但没有积极鼓励这一行为,确认最终治疗方案后完成术前谈话及手术签字,但在变动时组织讨论存在延迟,愿意听从安排协助护士及患者家属完善术前准备事项	未罗列所有治疗方案或未听取患者需求,确认最终治疗方案后未完成术前谈话及手术签字,忽略变动时的双方讨论环节,拒绝护士及患者家属协助完善术前准备事项的请求
	建立共识	启动讨论,听取患者需求,确认最终满意方案			
	协调团队活动	协助护士及患者家属完善术前准备事项,若有变动及时组织讨论			
领导力	制定和维护规则	遵守诊疗指南和病房规范	严格遵守诊疗指南和病房规范,体现职业素养,树立榜样,帮助护士协调医嘱执行情况,总是对患者和团队负责,在压力面前保持冷静,利用一切可及资源完成目标	偶尔表现出犹豫、被动等对待上级指示,在面对压力时偶尔有不符合职业素养的态度或行为,偶尔由于未把握可及资源降低工作效率	有频繁的不尊重团队成员或患者的言行,在压力面前完全丧失职业素养,由于无法把握可及资源降低工作效率
	支持同伴	帮助护士协调医嘱执行情况			
	应对压力	对患者和团队负责,在压力面前保持冷静,利用一切可及资源完成目标			
交流技巧	语音及清晰度	语音适中,口齿清晰	总是通过维持相对背景噪音适度的语音、清晰度及理解性,使用通俗的语言,保证交流的高效,使用礼节性称谓、敬语,保持眼神交流	绝大多数时间可以保证交流的高效,偶尔难以维持适度的语音、清晰度及理解性,或使用不加解释的医学术语,缺少礼节性称谓、敬语的使用及眼神交流,导致沟通精准度部分下降	难以保证交流的高效,无法维持适度的语音、清晰度及理解性,使用不加解释的医学术语,缺少礼节性称谓、敬语的使用及眼神交流,语言含糊不清或者不连贯,以至于沟通缺乏精准度
	措辞及理解力	使用通俗的语言,并能够理解他人的话语,无需多次重复			
	肢体语言	礼貌友善,保持眼神交流			
GSE					
总体安全结局评价			评分标准说明 <sup>b</sup>		
			非常满意	中立	非常不满意

注:<sup>a</sup>,5分对应 A 标准;4分介于 A、B 标准之间;3分对应 B 标准;2分介于 B、C 标准之间;1分对应 C 标准;<sup>b</sup>,5分对应“非常满意”;4分介于“非常满意”与“中立”之间;3分对应“中立”;2分介于“中立”“非常不满意”之间;1分对应“非常不满意”

总体安全结局的两轮评价方面,培训讨论前后,导师对同一学员的评分(均数及标准差)无明显趋势变化(表 4)。将单个导师给出的总体安全结局评分归为一组,培训讨论前后,评分组间差异无统计学意义。协和系数由 0.731 降低为 0.600,后者未提示差

异有统计学意义(表 5)。

### 3 讨论

由于普遍意识缺乏、当前教育体系与卫生服务实际需要略脱节,非手术临床技能的不足导致了一

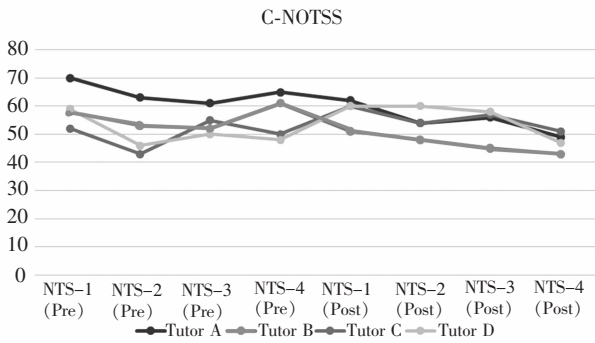


图 1 导师组培训前后 C-NOTSS 评分汇总

表 2 4 名学员所获两轮 C-NOTSS 评分 ( $\bar{x} \pm s$ )

学员	第一轮	第二轮
NTS-1	59.8 ± 7.5	58.3 ± 4.9
NTS-2	51.3 ± 8.9	54.0 ± 4.9
NTS-3	54.5 ± 4.8	54.0 ± 6.1
NTS-4	56.0 ± 8.3	47.5 ± 3.4

表 3 培训讨论前后导师给出的 C-NOTSS 评分单因素方差分析及 Kendall 协和系数检验

轮次	组别	SS值	df值	MS值	F值	P值	Kendall's W 值 <sup>a</sup>	P值
培训讨论前	组间	554.25	3	184.75	8.05	0.00	0.425	0.165
	组内	275.50	12	22.96				
	合计	829.75	15					
培训讨论后	组间	240.69	3	80.23	3.38	0.05	0.853	0.017
	组内	285.25	12	23.77				
	合计	525.94	15					

注:<sup>a</sup>, Kendall's coefficient of concordance

表 4 4 名学员所获两轮总体安全结局评分情况 ( $\bar{x} \pm s$ )

学员	第一轮	第二轮
NTS-1	4.50 ± 0.60	3.75 ± 0.50
NTS-2	3.25 ± 0.50	3.50 ± 0.60
NTS-3	3.50 ± 0.60	3.75 ± 0.60
NTS-4	3.75 ± 0.50	3.00 ± 0.00

表 5 培训讨论前后导师给出的总体安全结局评分单因素方差分析及 Kendall 协和系数检验

轮次	组别	SS值	df值	MS值	F值	P值	Kendall's W 值 <sup>a</sup>	P值
培训讨论前	组间	2.0	3	0.67	1.60	0.24	0.731	0.03
	组内	5.0	12	0.42				
	合计	7.0	15					
培训讨论后	组间	1.5	3	0.50	2.40	0.12	0.600	0.07
	组内	2.5	12	0.21				
	合计	4.0	15					

注:<sup>a</sup>, Kendall's coefficient of concordance

定不良医疗后果。张志峰等<sup>[5]</sup>发现,医疗投诉和纠纷案例事件发生的主要原因是医患沟通不足(42.2%)、服务态度欠缺(18.3%)。而 Uramatsu 等<sup>[6]</sup>统计的日本近年致死医疗事故中有近半数( $n=34, 46.6%$ )主要由非手术临床技能导致,其比例高于疾病自然进展(45.2%)及操作失误(5.5%)。因此,NTS 的培养已经逐渐被纳入住院医师培养指南中,成为考察外科医生整体素质的必需条目之一。

当前,我国住院医师培训中 NTS 的培养也面临一些问题:实践教学以教师理论授课为主、学员实践机会不足;进入科室轮训后,则更多依赖于学员对上级医师谈话签字等现实场景的主动观察,欠缺一个较为完整的理论体系<sup>[7]</sup>。临床环境中患者对医生使用的语言又往往有多元化的期待,更增加了对年轻医生沟通技能的考验<sup>[8]</sup>。

NTS 常用的评估方法(工具)有 NOTSS 及 OSANTS 量表。NOTSS 量表包含技能的各项内容,简洁明了,由于该量表要求打分者对患者的安全结局作预估判断,在实际操作时容易受到导师个人经验的左右。虽然有研究表明可以通过一定时长的培训使给分达到一定的可靠度,却也大大降低了量表的便利性和普适性<sup>[9]</sup>。另一常用量表 OSANTS 则在内容上移除了结局因素,更侧重于操作流程的完整性,缺失总体性的评判。存在学员虽能完成规定得分要点,但因为其他原因(语言、态度等)得不到患者配合导致诊疗失败,可能导致总体安全预估和 OSANTS 的评分结果产生较大偏差<sup>[3, 10]</sup>。由此不难发现,照搬国外的量表在我国并不可取。

同时,国外对外科医生 NTS 的评判常局限在特定的手术环境中。事实上,在我国临床住院医师及医学生的教学大纲中,除了各类实际操作训练外,门急诊接诊、入院病史采集、术前准备、术后康复等也占了重要的地位,是外科医生日常工作中必不可少的组成部分<sup>[11]</sup>。它们具有共性又各有侧重,在培训的过程中需逐个熟悉掌握,因此一系列适用于各类场景的模块化教学势在必行。从以上理念出发,本研究提出了 C-NTS 概念,即基于情景化和录像的非手术临床技能的评分标准。

C-NTS 主要融合以往 NOTSS、OSANTS、ORCA 量表的优点。第一,汲取了 NOTSS 的情境意识、决策、沟通和团队合作及领导力的框架,增加语言评判等符合我国国情的评分内容,并在此基础上根据不同情境深化拓展分项的具体评分内容,形成新的

C-NOTSS 体系。第二,设计总体安全结局评价(GES)来弥补 OSANTS 的缺点,计 1~5 分,反映导师对学员表现的整体印象及安全预估,从同理心角度间接体现在真实病房中患者对学员的直观感受和依从性。

对于 C-NOTSS 量表的评价,导师在培训讨论前打分的组间均分差异有统计学意义,反映学员得分情况区别较大,不具有良好的一致性。这可能是由于导师本身对非手术临床技能了解不够深入、对打分尺度的把握不足、容易受自身主观经验影响而导致。经过量表讨论及应用培训环节后,导师的打分已显示出接近趋势,各学员得分标准差降低、组间均分差异无统计学意义。同时,Kendall 协和系数的增长既印证了导师间评判原则同一性的提升,又提示导师对学员的评价并非有意保持一致,而是留存了显著的个性化差异,说明 C-NOTSS 的制定对于导师培训讨论后的评价同质化有明显帮助。

GSE 的高低主要取决于打分者结合自身临床经验来主观考察学员,更加贴近社会大众对学员的反馈和评价。培训讨论前后,导师对同一学员的打分差异均无统计学意义;说明导师对学员的总体印象和安全预估是相对恒定的,并不随考评标准变化而改变。推测原因为,同一平台相同资质的导师所拥有的临床经验相近,对学员操作带来的患者安全结局变动的认识相同。而 C-NOTSS 和 GSE 得分不同步改变的特性更突出显示,一个良好的 NTS 考察不仅要有详细的采分点,还需要有如 GSE 的独立评分因素,才能更加完善。

本研究创新性提出 C-NTS 概念,通过实验结论评估了其中导师培训(讨论)的有效性,结合录像的手段;既解决了导师培训考核同质化的难点,也有利于在今后培训时促使学员对自身技能的不足有直观而清晰的回顾。同时,导师经过 C-NTS 讨论及培训同质化后,对学员的表现仅需一位教师单独评分,从而在临床和带教实际中可节省时间和人力。由于本研究仅限于单个科室,样本量相对较少,相关结论仍需进一步讨论和验证。

**利益冲突** 所有作者均声明不存在利益冲突

**作者贡献声明** 秦骏、史苑:共同设计、撰写论文;敬然、俞旻皓、陈建军、张斌;参与课题设计及数据分析;顾磊、蒋春晖、叶光耀、孙隆慈;参与量表修正及结果讨论;Jay N Shah、钟鸣:总体把关、审订论文

## 参考文献

[1] Yule S, Flin R, Paterson-Brown S, et al. Development of a rating system for surgeons' non-technical skills [J]. Med Educ, 2006, 40:

1098-1104. DOI: 10.1111/j.1365-2929.2006.02610.x.

- [2] Jung JJ, Yule S, Boet S, et al. Nontechnical skill assessment of the collective surgical team using the non-technical skills for surgeons (NOTSS) system [J]. Ann Surg, 2019[Publish Ahead of Print]. DOI: 10.1097/SLA.0000000000003250.
- [3] Dedy NJ, Szasz P, Louridas M, et al. Objective structured assessment of non-technical skills: reliability of a global rating scale for the in-training assessment in the operating room [J]. Surgery, 2015, 157(6): 1002-1013. DOI: 10.1016/j.surg.2014.12.023.
- [4] Gardner AK, Russo MA, Jabbour II, et al. Frame-of-reference training for simulation-based intraoperative communication assessment [J]. Am J Surg, 2016, 212(3): 548-551.e2. DOI: 10.1016/j.amjsurg.2016.02.009.
- [5] 张志峰,袁岳鹏,杜兴旺.某院 10 年医疗投诉及纠纷分析[J].卫生职业教育,2018,36(22):123-125.  
Zhang ZF, Yuan YP, Du XW. Analysis of medical complaints and disputes in one hospital during the past decade [J]. Health Vocational Education, 2018, 36(22): 123-125.
- [6] Uramatsu M, Fujisawa Y, Mizuno S, et al. Do failures in non-technical skills contribute to fatal medical accidents in Japan? A review of the 2010-2013 national accident reports [J]. BMJ Open, 2017, 7(2): e013678. DOI: 10.1136/bmjopen-2016-013678.
- [7] 林建荣,李开为,金亚军,等.基于岗位胜任力的临床医学人才培养模式研究[J].湖北理工学院学报,2019,35(1):63-65,70. DOI: 10.3969/j.issn.2095-4565.2019.01.014.  
Lin JR, Li KW, Jin YJ, Lei Q. Research on training mode of clinical medical talents based occupational competency [J]. Journal of Hubei Polytechnic University, 2019, 35(1): 63-65, 70. DOI: 10.3969/j.issn.2095-4565.2019.01.014.
- [8] 龚余娟.医疗行业语言服务调查研究[D].扬州:扬州大学,2015. DOI: 10.7666/d.Y2908172.  
Gong YJ. Study on medical language services [D]. Yangzhou: Yangzhou University, 2015. DOI: 10.7666/d.Y2908172.
- [9] Jung JJ, Borkhoff CM, Jüni P, et al. Non-technical skills for surgeons (NOTSS): critical appraisal of its measurement properties [J]. Am J Surg, 2018, 216(5): 990-997. DOI: 10.1016/j.amjsurg.2018.02.021.
- [10] Dedy NJ, Bonrath EM, Ahmed N, et al. Structured training to improve nontechnical performance of junior surgical residents in the operating room: a randomized controlled trial [J]. Ann Surg, 2016, 263(1): 43-49. DOI: 10.1097/SLA.0000000000001186.
- [11] 彭春伟,熊斌.胃肠外科住院医师规范化培训体系建设与实践[J].医学教育管理,2018,4(5):356-359,369. DOI: 10.3969/j.issn.2096-045X.2018.05.002.  
Peng CW, Xiong B. Practice and discussion on the development of standardized training system of the gastrointestinal surgery residents [J]. Med Edu Mgt, 2018, 4(5): 356-359, 369. DOI: 10.3969/j.issn.2096-045X.2018.05.002.

(收稿日期:2019-04-12)

(本文编辑:蔡骏翔)