·教育评价·

基于“分层—反馈—迭代”模型的肝胆外科规培生核心能力评价指标体系构建

黄思敏[1]　王绮雯[2]　杨伟帮[1]　魏宇轩[1]　何龙光[3]　罗稀楷[4]　蒋小峰[1]

[1]广州医科大学附属第二医院肝胆外科，广州 510260；[2]广州医科大学附属第二医院外科实验室，广州 510260；[3]南方医科大学珠江医院肝胆二科，广州 510282；[4]广东医科大学附属医院肛肠外科，湛江 524002

通信作者：蒋小峰，Email：jiangxiaofeng008@163.com

【摘要】 **目的**　构建基于“分层—反馈—迭代”三维模型的肝胆外科规培生核心能力评价指标体系，以提升规培培养的系统性与前瞻性。**方法**　成立研究小组，通过系统文献回顾和半结构访谈，结合“分层—反馈—迭代”三维模型框架，初步拟定评价指标体系。采用德尔菲专家函询法，对来自全国多个省市的18名专家进行2轮函询，并运用层次分析法计算各级指标权重，最终确立科学、共识的评价体系。**结果**　2轮专家函询问卷的有效回收率均为100.00%。专家权威系数（Cr）分别为0.911和0.915，总指标的肯德尔协调系数分别为0.342和0.328（χ^2=261.09，246.33；均P<0.05），表明专家意见协调性好、结果可靠。最终构建的评价体系包含4项一级指标、12项二级指标、48项三级指标，并确定了各指标权重。**结论**　本研究成功构建了一套科学、系统且具备临床前沿适应性的肝胆外科规培生核心能力评价指标体系。该体系通过“分层—反馈—迭代”模型的有机整合，对精准评估规培生能力、动态优化培养方案、全面提升其岗位胜任力具有重要的现实意义与推广价值。

【关键词】 “分层—反馈—迭代”模型；　德尔菲法；　住院医师规范化培训；　核心能力评价

【中图分类号】 R-05

基金项目：广州医科大学本科教学质量与教学改革工程建设项目（2024ZLGC13）

DOI: 10.3760/cma.j.cn116021-20250415-02216

Construction of a core competency evaluation indicator system for hepatobiliary surgery residents based on the stratification-feedback-iteration model

Huang Simin[1], Wang Qiwen[2], Yang Weibang[1], Wei Yuxuan[1], He Longguang[3], Luo Xikai[4], Jiang Xiaofeng[1]

[1]Department of Hepatobiliary Surgery, The Second Affiliated Hospital, Guangzhou Medical University, Guangzhou 510260, China; [2]Surgery Laboratory, The Second Affiliated Hospital, Guangzhou Medical University, Guangzhou 510260, China; [3]Department of Hepatobiliary Surgery II, Zhujiang Hospital, Southern Medical University, Guangzhou 510282, China; [4]Department of Anorectal Surgery, The Affiliated Hospital of Guangdong Medical University, Zhanjiang 524002, China

Corresponding author: Jiang Xiaofeng, Email: jiangxiaofeng008@163.com

【**Abstract**】 **Objective**　To construct a core competency evaluation indicator system for hepatobiliary surgery residents based on the stratification-feedback-iteration three-dimensional model, and enhance the systematicity and foresight of residency training. **Methods**　A research group was formed. Through systematic literature review and semi-structured interviews, combined with the stratification-feedback-iteration three-dimensional model framework, the evaluation indicator system was preliminarily constructed. The Delphi expert consultation method was adopted to conduct two rounds of consultation with 18 experts from multiple provinces and cities nationwide. Analytic hierarchy process was used to calculate the weights of indicators at all levels, ultimately establishing a scientific and consensus-based evaluation system. **Results**

The effective recovery rates of the two rounds of expert consultation questionnaires were both 100.00%. The expert authority coefficients were 0.911 and 0.915, respectively. The Kendall coordination coefficients for total indicators were 0.342 and 0.328, respectively (χ^2 =261.09, 246.33; both P<0.05), indicating good coordination of expert opinions and reliable results. The final evaluation system included 4 primary indicators, 12 secondary indicators, and 48 tertiary indicators, with weights determined for each indicator. **Conclusions** This study successfully constructed a scientific, systematic core competency evaluation indicator system for hepatobiliary surgery residents with adaptability to clinical advancements. Through organic integration of the stratification-feedback-iteration model, this system has important practical significance and promotional value for accurately assessing resident competencies, dynamically optimizing training programs, and comprehensively enhancing their job competency.

【**Key words**】 Stratification-feedback-iteration model; Delphi method; Standardized residency training; Core competency evaluation

Fund program: Guangzhou Medical University Undergraduate Teaching Quality and Teaching Reform Engineering Construction Project (2024ZLGC13)

DOI: 10.3760/cma.j.cn116021-20250415-02216

肝胆外科手术因其复杂度高、风险大，对参与规范化培训的住院医师（简称“规培生”）的核心能力提出了多维度的严格要求。规培生的核心能力不仅是其职业发展的基础，也是医疗质量与患者安全的重要保障[1]。然而，传统的终结性评价模式多局限于操作技能的考核，往往忽视临床决策能力及医学人文素养等关键维度，且反馈周期较长，不利于培训过程中的实时改进。此外，由于评价标准缺乏统一性，评价结果的可比性也受到限制[2-3]。为提升规培教学质量，有必要构建一套以“分层—反馈—迭代”为核心理念的培养体系，以适应现代医学教育的发展趋势，有效促进规培生核心能力向临床胜任力转化。

“分层”培养强调在统一管理的基础上，结合规培生个体差异设计个性化培养方案，实施分层递进式带教，并设定阶段化、层次化的培训目标[4-5]。“反馈”机制则注重多维度、动态化的信息收集与回应，帮助培训基地实时掌握规培生的学习进展，进而对课程内容与教学策略进行持续调整[6]。“迭代”优化是指借助德尔菲法与层次分析法构建具有动态响应能力的指标权重体系，以适应医学教育内涵与外延的不断演进[7]。

基于上述理念，本研究旨在构建基于“分层—反馈—迭代”三维模型的肝胆外科规培生核心能力评价指标体系，以实现对规培生核心能力全面、客观、系统且科学地评估，为现代医学教育提供新型教学模式与评价参考。

1 对象与方法

1.1 成立研究小组

为确保研究的专业性与科学性，成立了由 13 人构成的研究核心组。成员包括本院肝胆外科主任 1 人、副主任 3 人、主治医师 4 人及研究生 5 人；其中博士 3 人、硕士 5 人、在读研究生 5 人；高级职称 4 人、中级职称 4 人。研究核心组负责主导研究的整体设计与执行，具体任务包括：初步构建基于“分层—反馈—迭代”模型的评价指标框架；遴选德尔菲专家咨询的函询专家；编制、发放与回收专家咨询问卷；并对函询结果进行汇总、分析与讨论，最终确立指标体系的具体内容。

1.2 “分层—反馈—迭代”评价指标体系初步构建

1.2.1 理论基础

本研究基于系统化教学理论和形成性评价理念，整合现有研究成果，构建了“分层—反馈—迭代”三维模型，作为构建评价指标体系的核心理论框架。该框架并非直接引用某个单一现成模型，而是对现代医学教育中 3 个关键维度的有机整合与概念化提升。

“分层”维度：借鉴了阶梯式教育理念[8-9]，其核心在于承认并尊重规培生个体的能力差异与发展规律。打破“一刀切”的培养模式，根据规培生的知识基础、技能水平和学习进度，设计个性化、阶段

化、层次化的培训目标与路径。例如，在手术技能训练中，从基础操作到复杂手术配合的递进，即是“分层”理念的体现。

“反馈”维度：源于形成性评价理论[10-11]，强调评价的动态性与指导性。它指建立一种多源、实时、双向的信息回馈机制。通过来自带教教师、同行、护理团队乃至患者的多角度评价，以及基于病例讨论、操作复盘等场景的即时反馈，帮助规培生准确识别自身优势与不足，为后续学习提供明确导向。

“迭代”维度：取自质量管理中的“迭代开发”思想，并应用于教育评价体系的持续改进[12-13]。评价体系本身应具备自我进化、动态调整的能力。通过定期收集数据、分析评估效果，并运用德尔菲法、层次分析法等工具，对指标内容和权重进行周期性优化，使之能够适应医学技术飞速发展和临床需求的变化。

这3个维度构成了一个紧密相连、循环促进的有机整体。“分层”是基础，设定了培养的起点和路径；“反馈”是驱动者，在既定路径上进行实时调整与能力提升；“迭代”是保障，确保整个培养与评价体系能够与时俱进，长期有效。

1.2.2　指标体系构建的原则

本研究在构建肝胆外科规培生核心能力评价指标体系时，严格遵循以下4项基本原则，以确保其科学、实用且能有效服务于“分层—反馈—迭代”的培养模型。①系统性原则：指标体系应全面覆盖肝胆外科规培生所需的核心能力维度，不仅包括医学知识、临床技能等硬性要求，更需整合临床决策、沟通协作及职业素养等软性能力，使各项指标间既相对独立又有机联系，共同构成一个层次分明、逻辑清晰的综合评价系统，从而全方位映射“分层—反馈—迭代”模型的多维要求。②科学性原则：指标的筛选与确立必须基于坚实的医学教育理论和充分的循证依据。本研究将通过系统的文献回顾和严格的德尔菲专家咨询法，确保每一项指标都有其明确的定义、理论支撑和实践意义，从而保证指标体系在内容上的科学性与权威性，为客观、准确地评价规培生能力奠定坚实基础。③导向性原则：指标体系不仅用于终结性评价，更重在形成性引导。其设计需紧密结合临床实践与未来职业发展需求，旨在引导规培生明确各阶段（分层）的学习目标与发展方向。同时，通过指标评价结果的动态“反馈”，帮助带教教师与规培生自身识别短板，明确改进路径，最终实现教学相长与能力的持续“迭代”优化。④可操作性原则：所有指标必须是具体、可观察且可衡量的，评价标准应清晰明确，数据来源应易于获取。这确保了该体系在繁忙的临床教学环境中能够被方便、高效地应用，从而为持续的动态反馈与周期性的迭代评估提供稳定可靠的数据支持，避免体系因过于理想化而流于形式。

1.2.3　文献回顾

为确保评价指标体系构建的科学性与前沿性，本研究进行了系统、全面的文献与政策调研。检索时间范围为2015年1月至2024年11月，力求覆盖近10年该领域的研究进展。在中文文献方面，以“分层培养”“反馈机制”“迭代优化”“肝胆外科”“规培生”“核心能力”“评价指标体系”等为关键词，系统检索了中国知网（CNKI）、万方数据库、维普期刊（VIP）和中华医学期刊网（CMJN）数据库。在英文文献方面，以“Hierarchical training”“Feedback mechanism”“Iterative optimization”“Hepatobiliary Surgery”“Resident”“Core competency”“Evaluation indicator system”等为检索词[14-18]，全面检索了PubMed、Web of Science、EMbase和ScienceDirect等权威数据库。

此外，为紧密对接临床实践与政策要求，本研究还系统追溯了重要机构发布的相关文件，包括《住院医师规范化培训基地标准（2022年版）》[19]《住院医师规范化培训内容与标准（2022年版）》[20]《住院医师规范化培训基地临床技能培训中心工作指南（2023年版）》[21]等文件。通过多途径、多维度的文献调研，旨在全面掌握该领域的研究前沿与实践动态，为本研究初步构建基于“分层—反馈—迭代”模型的肝胆外科规培生核心能力评价指标体系提供坚实的理论、政策与实践依据。

1.2.4　半结构访谈

为确保访谈信息能够全面、多角度地反映肝胆外科规培的实际情况，本研究采用目的性抽样法，严格界定“关键知情人”为在规培生培养链条中扮演重要角色、并拥有丰富直接经验的人员。其遴选标准如下：①角色多样性：需覆盖规培管理的决策者（科室主任）、执行者（导师、主治医师）和体验者

（规培生）。②经验丰富性：带教教师需具备5年以上教学经验；规培生需已完成核心轮转，对培训全过程有深刻体会。③地域与机构代表性：为避免单一机构的局限性，受访者分别来自广州市内3家不同的三级甲等教学医院（广州医科大学附属第一医院、中山大学附属第三医院、广东省人民医院），这些医院在肝胆外科的规模和教学特色上具有一定差异。最终纳入的10名受访者具体信息见表1。

为深入探究“分层—反馈—迭代”模型在肝胆外科规培生核心能力评价中的具体内涵，本研究采用目的抽样法，于2024年8月至2024年10月期间，选取肝胆外科领域不同角色的关键知情人进行半结构式访谈。受访者包括3名科室主任（均具副高级及以上职称）、2名具有5年以上教学管理经验的规培生导师、2名中级及以上职称的主治医师，以及3名已完成至少1年规范化培训的规培生，共计10人。

访谈提纲围绕“分层—反馈—迭代”模型的核心要素设计，主要涵盖以下问题：①您认为肝胆外科规培生需具备哪些核心能力以胜任临床工作？②在分层培养中，应如何根据规培生能力差异制定个性化培训目标？③现有反馈机制在规培生能力提升中的作用如何？有何改进建议？④在迭代优化评价指标时，需重点关注的动态调整维度有哪些？⑤如何将临床决策能力、医学人文素养等隐性能力有效纳入评价体系？

访谈前均获取参与者知情同意，并对过程全程录音。每次访谈持续30~60 min，录音资料于访谈结束后24 h内由研究小组成员转录为文本，并进行初步整理。研究小组运用主题分析法对文本资料进行编码与提炼，通过3轮小组讨论，整合访谈内容与前期文献回顾结果，初步归纳出肝胆外科规培生核心能力的关键要素。例如，受访者普遍认为，“分层培养应结合规培生临床经验差异，设计阶梯式技能训练路径”；反馈机制应“覆盖操作技能、病例分析、团队协作等多维度内容”；迭代优化需“根据肝胆外科高风险特点，动态调整如急救技能等关键指标的权重”。

访谈资料分析方法与过程：为确保从访谈资料中提炼出的核心能力要素科学、可信，本研究采用质性研究中的主题分析法（thematic analysis），并遵循Braun和Clarke[22]提出的六阶段分析框架进行系统分析。具体过程如下：①熟悉资料：在将访谈录音逐字转录为文本后，研究小组全体成员反复阅读文本，以沉浸其中，获取整体印象，并记录下初步的想法。②生成初始编码：研究小组使用NVivo 12.0质性分析软件作为辅助工具，对10份访谈文本进行开放式编码。由2名研究人员背对背独立进行编码，将文本中有意义的内容单元（如一句话、一个段落）标记并赋予一个初步代码。例如，将“我们会根据规培生第1年的考核成绩，分配不同难度的腹腔镜模拟训练”编码为“按考核成绩分层培训”；将“每次大手术后，主刀医生都会立刻和我们复盘手术视频，指出可以改进的细节”编码为“术后即时视频复盘反馈”。③搜寻主题：在生成大量初始代码的基础上，研究人员通过多次小组讨论，将意义相关、模式相近的代码进行归类和聚合，形成潜在的主题。例如，将“按考核成绩分层培训”“个性化轮转计划”“阶梯式手术参与”等代码聚合，初步形成“分层培养路径”这一主题。④审阅主题：此阶段旨在检验主题与原始资料及整体编码的契合度。通过不断调整与优化，确保每个主题内部具有一致性且不同主题之间具有清晰的界线。例如，最初“团队协作”被视为一个独立主题，但在审阅中发现其与“沟通反馈”和“职业素养”紧密交织，因此最终将其整合为“隐性胜任力”主题下的一个次级要素。⑤定义和命名主题：对每个确定的主题进行精确定义，明确其核心内涵与边界。经过多轮迭代，最终精炼出4个高度概括且相互关联的核心主题（即一级指标雏形）：分层培养能力（包括基础操作技能分

表1 半结构访谈受访者基本信息（n=10）

角色	职称/年级	教学/规培经验	机构
科室主任	正高级职称	均>10年	广州医科大学附属第一医院1人、中山大学附属第三医院1人、广东省人民医院1人
规培生导师	副高级职称	5~10年	广州医科大学附属第一医院1人、中山大学附属第三医院1人
主治医师	中级职称	均>5年	中山大学附属第三医院1人、广东省人民医院1人
规培生	第2~3年规培	无	广州医科大学附属第一医院1人、中山大学附属第三医院1人、广东省人民医院1人

层训练、复杂手术分阶段参与、个性化学习路径设计）、动态反馈能力（涵盖实时操作反馈、病例讨论反馈、多导师联合评价）、迭代优化能力（涉及科研思维培养、新技术适应能力、评价指标动态权重调整）和隐性胜任力（如医患沟通能力、伦理决策能力、职业抗压能力）。⑥撰写报告：即本论文的撰写过程，选择最具代表性的、生动的原始陈述作为例证，来支撑和阐释每个确定的主题，最终形成研究发现。

为确保编码过程的可信度，由 2 名研究人员对同一份文本进行编码后，计算其编码者间信度，Cohen's Kappa 系数为 0.82，表明编码结果具有良好的一致性。对于有分歧的编码，通过第 3 名研究员的介入与小组讨论达成共识。研究小组结合“分层—反馈—迭代”三维模型框架，将上述要素整合为包含 4 项一级指标、12 项二级指标和 36 项三级指标的评价指标体系初稿，为后续德尔菲专家函询奠定基础。

1.3　编制专家函询问卷

基于前期文献回顾与半结构访谈结果所形成的指标体系初稿，编制了德尔菲专家函询问卷。问卷具体包括 3 个部分。①问卷说明：本部分旨在向专家清晰阐述研究背景与目的。同时，详细说明了问卷的填写方法、预计耗时、回收截止日期，并明确告知专家其参与是基于自愿原则。②专家基本情况调查表：内容涵盖专家的性别、年龄、工作年限、最高学历、专业技术职称、是否为博/硕士研究生导师、主要专业领域及研究方向。此外，特别设置了专家对本次函询内容的“判断依据”（通常分为理论分析、实践经验、同行了解和直观感受 4 个维度）与“熟悉程度”的自评项目。③指标评价函询表：采用 Likert 5 级评分法（1=“非常不重要”，2=“不重要”，3=“一般”，4=“重要”，5=“非常重要”），请专家对各级指标的重要性进行评判。为充分收集定性意见，在每个指标条目后均设有“修改意见栏”，专家可对指标的名称、表述或结构提出具体修改、删减或增补的建议。

1.4　遴选函询专家

为保证德尔菲专家函询的权威性与代表性，本研究采用目的性抽样法遴选函询专家。专家的纳入标准为：①在全国三级甲等医院肝胆外科从事临床或教学管理工作；②具有本科及以上学历；③具备副高级及以上专业技术职称；④知情同意并承诺能积极参与完成全部轮次的专家函询。最终共邀请来自全国 7 个省市（包括北京、上海、广东、湖北、四川、陕西、江苏）的 18 名专家参与本研究。这些地区在医疗资源和发展水平上具有一定梯度，旨在保证专家样本的地域代表性。所有专家均来自国家卫生健康委认定的住院医师规范化培训基地医院，以确保其对规培工作有深入的理解和实践。

专家群体构成如下：年龄范围为 43~61 岁，平均（53.41 ± 5.10）岁；临床或教学相关工作年限为 10~27 年，平均（21.61 ± 3.11）年；学历方面，本科 4 人（22.22%），硕士 9 人（50.00%），博士 5 人（27.78%）；职称方面，副高级职称 8 人（44.44%），正高级职称 10 人（55.56%）；其中，14 人（77.78%）为硕士研究生导师，7 人（38.89%）为博士研究生导师。在地域分布上，华东地区 5 人（27.78%），华北地区 4 人（22.22%），华南地区 4 人（22.22%），华中地区 2 人（11.11%），西南地区 2 人（11.11%），西北地区 1 人（5.56%）。在机构类型上，15 人（83.33%）来自大学附属医院，3 人（16.67%）来自省级重点专科医院。

1.5　实施专家咨询

本研究于 2024 年 12 月至 2025 年 2 月期间，共组织实施了 2 轮德尔菲专家函询。问卷主要通过电子邮件与专业的在线问卷平台进行发放与回收。在第 1 轮函询结束后，研究小组对回收的问卷数据进行整理与分析。指标的筛选标准如下：当某一指标的评分均值<3.50，或变异系数（CV）>0.25 时，予以删除。同时，研究小组对各指标后附的专家文字意见进行了归纳与整合，并据此对指标体系进行了修订，包括对指标表述的修改、条目的删除或新增，最终形成第 2 轮专家函询问卷。第 2 轮函询结束后，数据分析显示各指标评分的变异系数显著减小，专家意见已趋于一致，认为函询结果达到收敛标准，故结束本轮函询。

1.6　统计学方法

计量资料采用均数 ± 标准差表示，计数资料采用频数和百分比表示。专家积极性以问卷回收率表示，专家意见集中程度采用满分率和重要性评分表示；专家可靠性以专家权威系数（Cr）表示，专家权威系数通过判断系数（Ca）与熟悉程度（Cs）来计算；

专家意见协调程度以变异系数及肯德尔协调系数判断；采用层次分析法确定各级指标的权重。检验水准α=0.05。

2　结果

2.1　专家积极系数和权威程度

本研究共进行2轮专家函询。每轮均向18名专家发放问卷，2轮均回收有效问卷18份，问卷有效回收率均为100.00%，表明专家对本研究关注度与参与度很高，积极系数理想。专家权威程度（Cr）由判断依据（Ca）与熟悉程度（Cs）2个维度的平均值计算得出。2轮函询中专家的Ca、Cs及Cr值均处于较高水平。其中，第1轮函询的Ca、Cs、Cr分别为0.890、0.901、0.911；第2轮函询分别为0.902、0.897、0.915。2轮函询的Cr值均大于0.70的可接受标准，表明参与本研究的专家在整个过程中均具有较高的权威性，函询结果可靠。

2.2　专家意见协调程度

专家意见协调程度通过变异系数与肯德尔协调系数共同评估。在本研究的2轮函询中，各级指标的变异系数分别介于0.017~0.238与0.022~0.231之间，且所有指标变异系数均小于0.25，表明专家对各项指标的评价意见趋于一致，分歧较小。此外，经肯德尔协调系数检验，第1轮与第2轮函询的总指标协调系数分别为0.342与0.328（χ^2=261.09，246.33；均P<0.05），说明专家在整个函询过程中对指标重要性的评价具有良好且稳定的共识，群体意见协调程度较高。

2轮函询中，所有指标在2轮函询中的变异系数均低于0.25的临界标准，且第2轮函询的变异系数普遍较第1轮有所降低，表明经过意见征询与指标修订，专家共识度进一步提高，意见协调性良好（表2）。

表2　2轮德尔菲专家函询各级指标变异系数（CV）比较

指标层级	指标数量	第1轮函询变异系数（CV）范围	第2轮函询变异系数（CV）范围
一级指标	4	0.031~0.075	0.031~0.075
二级指标	12	0.029~0.074	0.022~0.074
三级指标	48	0.022~0.238	0.024~0.231
全体指标	64	0.017~0.238	0.022~0.231

注：变异系数（CV）=标准差/均值。CV值越小，代表专家对该指标重要性的意见越集中，共识度越高。通常认为CV<0.25即表示专家意见协调性好。

2.3　专家对构建评价指标体系问卷的函询结果

在2轮专家函询中，专家对各项指标的重要性评分显示出较高的认可度。第1轮函询中，各项指标的重要性评分均值为4.20~4.78，变异系数为0.017~0.238，满分率为24%~100%。

根据第一轮专家反馈的定性意见，研究小组对指标体系进行了修订，包括：①指标表述的精确化：针对专家指出原一级指标“隐性核心能力”一词过于宽泛、学术气过浓的问题，为更贴合临床胜任力模型的实际语境，将其修改为“隐性胜任力”。②指标内容的细化与场景化：对多个三级指标的描述进行了补充，以明确其具体内涵。例如，原“肝脏切除术阶段性配合”因专家建议“应明确配合的具体环节以利评价”，被细化为“肝部分切除术阶段性配合（解剖游离、止血、创面处理）”。原“手术模拟训练”因专家提出“应区分基础与高阶模拟，突出肝胆外科特色”，被修改为“高难度手术模拟训练（虚拟仿真系统、动物实验）”。③指标的新增：为回应专家关于体系应覆盖前沿技术与全周期照护的建议，新增了8项三级指标。例如，针对专家强调的“术中影像判读是肝胆外科关键能力”，新增“术中影像学判读能力（超声、CT图像实时分析）”。多位专家关于纳入尖端手术技术的共识，新增“机器人辅助肝胆手术基础操作技能”及“肝移植手术分阶段参与（供肝修整、血管吻合、免疫管理）”。

第2轮专家函询中，各级指标的重要性评分均值为4.21~4.89，变异系数为0.022~0.231，满分率为34%~100%，所有指标均满足纳入标准（重要性均值>3.50，变异系数<0.25）。专家意见更为集中，主要建议补充人文关怀与复杂病种处理能力。研究小组据此新增4项三级指标：“临终关怀实践能力（舒缓医疗、家属心理支持）”（回应专家对人文素养落地的建议）。“肝胆肿瘤分子分型与精准治疗策略应用能力”（回应专家对紧跟精准医学发展的建议）。“肝门部胆管癌根治术的规范化步骤参与”与“高龄/高危患者损伤控制外科理念的应用”（回应专家对覆盖高难度、特殊人群手术规范的建议）。

经2轮函询与多轮小组讨论，最终构建形成基于“分层—反馈—迭代”三维模型的肝胆外科规培生核心能力评价指标体系，其中包括4项一级指标、12项二级指标、48项三级指标（表3）。

表 3　基于“分层—反馈—迭代”三维指标模型的肝胆外科规培生核心能力评价指标体系

指标	重要性评分（$\bar{x} \pm s$）	变异系数	权重	满分率（%）
1 分层培养维度	4.82 ± 0.15	0.031	0.310	96
1.1 基础操作技能分层训练	4.75 ± 0.14	0.029	0.350	88
1.1.1 腹腔镜基础操作技能（扶镜、分离、缝合）	4.55 ± 0.20	0.044	0.080	86
1.1.2 开腹手术缝合技能（单纯缝合、间断缝合、皮内缝合）	4.62 ± 0.15	0.032	0.080	92
1.1.3 超声引导下穿刺技能（肝脏活检、胆道引流）	4.55 ± 0.22	0.048	0.064	87
1.1.4 围手术期加速康复（ERAS）理念的执行与管理能力	4.70 ± 0.19	0.040	0.080	78
1.1.5 术中影像学判读能力（超声、CT图像实时分析）	4.72 ± 0.25	0.053	0.080	82
1.2 复杂手术分阶段参与	4.50 ± 0.30	0.067	0.400	72
1.2.1 肝部分切除术阶段性配合（解剖游离、止血、创面处理）	4.52 ± 0.48	0.106	0.100	82
1.2.2 胰十二指肠切除术分阶段参与（吻合技术、术后并发症管理）	4.38 ± 0.28	0.064	0.100	71
1.2.3 胆道重建手术分阶段训练（胆肠吻合、T管置入）	4.53 ± 0.13	0.029	0.080	66
1.2.4 肝门部胆管癌根治术的规范化步骤参与	4.58 ± 0.29	0.063	0.080	77
1.2.5 肝移植手术分阶段参与（供肝修整、血管吻合、免疫管理）	4.56 ± 0.30	0.066	0.080	76
1.2.6 高龄/高危患者损伤控制外科理念的应用	4.45 ± 0.33	0.074	0.080	70
1.3 个性化学习路径设计	4.60 ± 0.25	0.054	0.250	65
1.3.1 基于规培生能力评估的个体化培训计划制定	4.27 ± 0.38	0.089	0.050	76
1.3.2 高难度手术模拟训练（虚拟仿真系统、动物实验）	4.50 ± 0.10	0.022	0.050	81
2 动态反馈维度	4.88 ± 0.20	0.041	0.320	92
2.1 实时操作反馈	4.92 ± 0.18	0.037	0.350	91
2.1.1 术中操作即时评价（精准度、效率、安全性）	4.68 ± 0.15	0.032	0.100	96
2.1.2 术后病例复盘与改进建议（手术视频分析、错误总结）	4.80 ± 0.29	0.060	0.100	91
2.1.3 床旁操作技能反馈（换药、引流管管理）	4.55 ± 0.16	0.035	0.100	86
2.1.4 急救操作实时评分（心肺复苏、大出血处理）	4.42 ± 0.39	0.088	0.100	81
2.2 病例讨论多维反馈	4.49 ± 0.32	0.071	0.350	90
2.2.1 疑难病例分析能力反馈（诊断思路、治疗方案选择）	4.73 ± 0.75	0.159	0.100	81
2.2.2 MDT讨论中的角色扮演与贡献评价（含AI辅助决策应用）	4.61 ± 0.65	0.141	0.100	83
2.2.3 术后并发症处理能力反馈（早期识别、干预措施）	4.65 ± 0.60	0.129	0.080	81
2.2.4 临床路径依从性评价（指南规范执行情况）	4.79 ± 0.31	0.065	0.080	76
2.3 多导师联合评价	4.75 ± 0.35	0.074	0.300	89
2.3.1 主刀医师对手术配合能力的评价	4.81 ± 0.18	0.037	0.080	76
2.3.2 护理团队对围手术期管理能力的评价	4.54 ± 0.32	0.070	0.080	81
2.3.3 患者及家属对沟通效果的满意度反馈	4.38 ± 0.20	0.046	0.080	71
3 迭代优化维度	4.85 ± 0.22	0.045	0.300	88
3.1 科研思维与创新能力	4.72 ± 0.19	0.040	0.400	82
3.1.1 临床问题转化为科研选题的能力	4.43 ± 0.20	0.045	0.100	83
3.1.2 机器人辅助肝胆手术基础操作技能	4.71 ± 0.23	0.049	0.100	88
3.1.3 循证医学实践能力（指南解读、证据整合）	4.61 ± 0.39	0.085	0.100	85
3.1.4 教学相长与知识传承能力	4.52 ± 0.26	0.058	0.080	77
3.1.5 临床科研能力分层培养	4.60 ± 0.29	0.063	0.080	75
3.1.6 终身学习与资源整合能力	4.67 ± 0.24	0.051	0.080	80
3.1.7 肝胆肿瘤分子分型与精准治疗策略应用能力	4.58 ± 0.31	0.068	0.080	76
3.2 评价指标动态权重调整	4.86 ± 0.24	0.049	0.400	85
3.2.1 急救技能权重动态匹配肝胆外科高风险特点	4.63 ± 0.27	0.058	0.100	82
3.2.2 技术更新对指标体系的适应性调整	4.75 ± 0.35	0.074	0.100	78
3.2.3 基于临床需求的指标优先级迭代	4.73 ± 0.29	0.061	0.080	73
3.2.4 区域医疗差异适应性调整	3.97 ± 0.52	0.131	0.080	62

续表 3

3.3 职业发展持续优化	4.70 ± 0.25	0.053	0.200	75
3.3.1 个人职业规划与科室发展目标的契合度	3.98 ± 0.26	0.065	0.100	63
3.3.2 参与继续教育与学术会议的主动性	4.55 ± 0.19	0.042	0.080	63
4 隐性胜任力	4.38 ± 0.33	0.075	0.070	83
4.1 医患沟通与人文素养	4.85 ± 0.18	0.037	0.350	78
4.1.1 复杂病情告知能力（如与晚期肝癌患者沟通）	4.37 ± 0.58	0.133	0.100	82
4.1.2 精准医学时代的医患共享决策沟通	4.45 ± 0.55	0.124	0.100	79
4.1.3 跨文化沟通能力（少数民族、外籍患者）	4.07 ± 0.24	0.059	0.100	63
4.1.4 临终关怀实践能力（舒缓医疗、家属心理支持）	4.49 ± 0.11	0.024	0.100	77
4.2 伦理决策与法律意识	4.74 ± 0.31	0.065	0.350	86
4.2.1 终末期患者治疗决策的伦理考量	4.32 ± 0.85	0.197	0.080	77
4.2.2 医疗纠纷预防与应对能力	4.42 ± 0.83	0.188	0.080	77
4.2.3 患者隐私保护与数据安全管理	3.66 ± 0.39	0.107	0.080	72
4.3 团队协作与抗压能力	4.24 ± 0.21	0.050	0.300	65
4.3.1 多学科协作中的角色定位与执行力	4.27 ± 0.14	0.033	0.100	67
4.3.2 高强度手术中的专注力与应变能力	4.52 ± 0.18	0.040	0.100	52
4.3.3 职业倦怠与心理压力的自我调节能力	4.52 ± 0.15	0.033	0.080	47
4.3.4 临床工作任务管理与效率优化能力	3.85 ± 0.12	0.031	0.080	48

3 讨论

3.1 该指标体系的可靠性和科学性

本研究通过德尔菲专家函询法与层次分析法构建指标体系，确保了方法的科学性和结果的可靠性。2 轮专家函询的有效回收率均为 100.00%，专家权威系数（Cr 值分别为 0.911 和 0.915）均高于 0.70 的可接受标准，表明参与专家的意见具有高度权威性；总指标的肯德尔协调系数分别为 0.342 和 0.328（均 $P<0.05$），说明专家小组的意见协调性良好。

在指标权重确定上，本研究采用层次分析法，将专家的主观经验判断与数学模型的客观计算相结合，有效避免了主观随意性带来的偏倚。计算得出的各维度权重（如分层培养维度权重 0.310、动态反馈维度权重 0.320、迭代优化维度权重 0.300）具有较好的结构合理性。

此外，所有三级指标的变异系数均控制在 0.25 以下，进一步验证了专家对指标重要性评价的高度一致性。整个研究过程遵循严格的构建流程，通过系统的文献回顾、深入一线的半结构访谈以及多轮专家意见的迭代优化，确保了最终形成的指标体系不仅理论依据充分，而且紧密贴合临床实践，全面覆盖了肝胆外科规培生核心能力培养的全流程需求，体现了严谨、科学的研究设计。

3.2 该评价指标体系的内容特点及创新亮点

本指标体系突破了传统评价模式中“重技能、轻素养”“重结果、轻过程”的局限，通过“分层—反馈—迭代”三维模型的深度整合，实现了以下创新性突破：①分层递进性。构建阶梯式与前沿性并重的培养路径：一级指标逻辑清晰，二级指标细化培养路径，三级指标精准对应基础与前沿临床场景，体现从基础到高阶的能力进阶[23-24]。一级指标“分层培养维度”通过“基础操作技能分层训练”“复杂手术分阶段参与”等二级指标，将肝胆外科手术技能分解为递进式模块。例如，三级指标不仅包含“肝部分切除术阶段性配合”等经典项目，更新增了“肝门部胆管癌根治术的规范化步骤参与”和“高龄/高危患者损伤控制外科理念的应用”，确保了培训内容与最新临床指南和外科理念同步。这种分层设计不仅契合《住院医师规范化培训标准》中对分阶段能力的要求，还能显著降低复杂及高风险手术的初期失误率。②动态适应性。推动指标体系与技术前沿协同进化：传统评价体系固化滞后，本研究通过“迭代优化维度”模块实现动态权重调整与技术迭代[25-26]。例如，三级指标“技术更新对指标体系的适应性调整”明确要求将“机器人辅助肝胆手术”等新技术纳入评价标准；而新增的“肝胆肿瘤分

子分型与精准治疗策略应用能力”则直接响应了精准医学的临床进展。这一机制使指标体系具备“自适应性”，能够持续响应医疗技术革新与临床需求的快速变化。③隐性能力显性化。破解人文与职业素养评价难题：“隐性胜任力”模块将伦理决策、抗压能力及前沿医患沟通等抽象概念转化为可量化指标。例如，新增的“精准医学时代的医患共享决策沟通”聚焦于向患者解释复杂新疗法（如免疫、靶向治疗）这一现实挑战，使沟通能力的评价更具时代性。同时，“职业倦怠自我调节能力”等条目则引入心理量表与导师观察记录的双维度评分，为人文素养与职业精神的评价提供了可操作的解决方案。④多源反馈网络。提升评价客观性与全面性：通过整合主刀医师、护理团队、患者及跨学科导师的多维度评价，构建全景式反馈网络[27-28]。优化后的指标“MDT讨论中的角色扮演与贡献评价（含AI辅助决策应用）”不仅考核规培生的专业知识与协作能力，更强调了在现代化、多学科诊疗环境中的实践贡献。这种多源数据融合机制有效避免了单一评价主体的主观偏倚，为规培生能力画像提供了立体、客观的支撑。

本体系以“三维模型”为框架，实现了分层培养、动态反馈与技术迭代在肝胆外科规培领域的系统性融合，并通过将隐性能力量化和多源数据整合，构建了“能力发展—反馈改进—动态优化”的闭环评价生态。其核心创新在于从静态考核转向持续性发展评价、从单一技能导向转为综合胜任力导向，为肝胆外科乃至外科规培体系的科学化、精细化发展提供了兼具前瞻性与实用性的可推广范式。

3.3　指标体系的应用价值、挑战与未来展望

本研究构建的指标体系不仅是一套评价工具，更是一个融入“分层—反馈—迭代”理念的规培生能力发展管理系统。其应用与推广将为肝胆外科乃至外科规培教学带来多维度的价值，同时也面临相应挑战，并孕育着与新技术深度融合的广阔前景。

3.3.1　应用价值与潜在挑战

本体系的核心应用价值在于实现“从静态考核到动态发展”的范式转变。①临床教学管理层面，它能为培训基地提供精细化的管理工具。通过“分层培养维度”，教学管理者可以依据指标为不同阶段的规培生定制个性化轮转计划；借助“动态反馈维度”的多源数据，可实现形成性评价，及时干预和辅导；利用“迭代优化维度”的机制，定期审视并更新培训内容与标准，确保其与前沿技术（如机器人手术、精准医疗）同步。②规培生个人发展层面，该体系如同一位“职业导航师”。它通过清晰的指标层级，为规培生指明了从入门到精通的成长路径。持续的、多维的反馈能帮助其获得准确的自我认知，明确改进方向。而体系对“隐性胜任力”和“迭代能力”的强调，则引导规培生注重医学人文、科研思维和终身学习习惯的养成，为其长远职业发展奠定坚实基础。

然而，体系的全面推广也面临挑战：①实施成本较高，依赖于带教师资投入大量时间进行细致评价与反馈，在繁忙的临床环境中可能遭遇阻力；②对数据采集与处理能力要求高，如何高效地整合手术视频、病历记录、多源评价等海量数据是一大难题；③评价者间信度问题，需对参与评价的各级人员（主刀医师、护理团队等）进行统一培训，以确保评价标准的一致性。

3.3.2　未来展望

未来，本指标体系可与前沿信息技术深度融合，以应对上述挑战，并开拓更智慧的应用场景。如开发智能规培管理平台，将指标体系嵌入信息化平台，实现指标分配、数据采集、结果计算与报告生成的全流程自动化。这能极大降低管理成本，提升效率。也可以探索基于大数据的自动化评价：利用自然语言处理技术，自动分析规培生书写的病历、手术记录，评估其临床思维；通过视频分析技术，对其手术操作视频进行自动化评估，提供客观、量化的技能反馈。这将部分解放带教教师，并实现更密集的过程性评价。

本研究为现代医学教育提供了一个兼具科学性、前瞻性与可操作性的评价范式，不仅对提升肝胆外科规培质量具有直接指导意义，其核心框架也为其他专科的规培评价体系改革提供了可借鉴的路径，对保障医疗质量与患者安全具有深远影响。

利益冲突　所有作者声明不存在利益冲突

作者贡献声明　黄思敏：确定研究题目、设计研究方案、撰写及修改论文；王绮雯、魏宇轩：资料查找、询问卷设计；杨伟帮、何龙光、罗稀楷：询问卷发放、资料收集；蒋小峰：提供研究建议、审校论文

参考文献

[1]　顾妍鑫，刘丽，胡惠惠，等. 以核心能力为导向的麻醉护理硕

士专业学位研究生临床实践培养方案的构建[J]. 护理学报, 2025, 32(1): 7-12. DOI: 10.16460/j.issn1008-9969.2025.01.007.

[2] 杨晨晨, 徐茜. 基于“雨课堂”智慧教学模式的组织学与胚胎学形成性评价体系研究[J]. 中国组织化学与细胞化学杂志, 2021, 30(1): 107-114. DOI: 10.16705/j. cnki. 1004-1850.2021.01.018.

[3] 马桂芝, 常军民, 李改茹, 等. 基于金课标准的《药物分析学》形成性评价体系的构建及效果分析[J]. 新疆医科大学学报, 2020, 43(12): 1646-1649. DOI: 10.3639/j.issn.1009-5551.2020.12.026.

[4] 张佳慧, 丁欣, 崔娜. 分层培养模式在重症医学科进修医师教学实践中的探索[J]. 医学研究杂志, 2023, 52(4): 185-188. DOI: 10.11969/j.issn.1673-548X.2023.04.040.

[5] 夏迪, 吴昊, 阮侠, 等. 非麻醉专业住院医师麻醉科规范化分层培养方案探索[J]. 协和医学杂志, 2022, 13(5): 895-900. DOI: 10.12290/xhyxzz.2022-0290.

[6] 谢郁芳, 张先翠. 基于反馈机制的案例导入式竞赛活动在临床护理本科生带教中的应用效果[J]. 护理实践与研究, 2023, 20(1): 128-131. DOI: 10.3969/j.issn.1672-9676.2023.01.026.

[7] 李钰婕, 高伟, 孔媛, 等. 基于改进德尔菲法与层次分析法构建寒区冷习服评价指标体系[J]. 临床军医杂志, 2024, 52(3): 257-261, 266. DOI: 10.16680/j.1671-3826.2024.03.10.

[8] 唐继卫, 徐立臻. 阶梯式教育模式在爱尔兰[J]. 中国高等教育, 2005(2): 46-47.

[9] 郭凌波, 常广明, 王永晨. 全科医学基层实践基地师资综合评价指标体系的构建[J]. 中华医学教育探索杂志, 2023, 22(10): 1483-1488. DOI: 10.3760/cma.j.cn116021-20211210-01445.

[10] 王烁, 宗序连. 形成性评价的理论内涵与实践反思[J]. 教学与管理, 2020(15): 1-4.

[11] 曹飞. 形成性评价前沿动向及借鉴[J]. 湖南师范大学教育科学学报, 2024, 23(2): 87-96, 116. DOI: 10.19503/j.cnki.1671-6124.2024.02.011.

[12] 姜大源. 论迭代、加速、平衡: 新时代继续教育的理论视角[J]. 职教论坛, 2024, 40(1): 77-85. DOI: 10.3969/j.issn.1001-7518.2024.1.zjlt202401011.

[13] 梅鲁海, 王耀军. 新《职业教育法》下工匠精神培养的“四维进阶”理论模型和“四重迭代”运作路径[J]. 宁波职业技术学院学报, 2024, 28(1): 25-30. DOI: 10.3969/j.issn.1671-2153.2024.01.005.

[14] Deng Y, Zhu W, Gong Q, et al. Building a hierarchical and classified "two-way" training system and cultivating a dual-teacher innovative teaching team through school-enterprise cooperation [J]. Erd, 2024, 6(6): 64-69. DOI: 10.26689/erd.v6i6.7552.

[15] Tang H, Ming JS, Chen Y. Application of PBL can improve the teaching quality of professional degree graduate students in hepatobiliary surgery [J]. Creat Educ, 2020, 11(12): 2609-2616. DOI: 10.4236/ce.2020.1112193.

[16] Stratman EJD, Vogel CA, Reck SJ, et al. Analysis of dermatology resident self-reported successful learning styles and implications for core competency curriculum development [J]. Med Teach, 2008, 30(4): 420-425. DOI: 10.1080/01421590801946988.

[17] Xia W. Research on the performance evaluation indicator system of university teachers' ethics in the New Era based on AHP model [C]//2021 4th International Conference on Information Systems and Computer Aided Education. Dalian China. ACM, 2021. DOI: 10.1145/3482632.3483043.

[18] Esan OT, Ogunoye BT, Sodimu TO, et al. Extent of research involvement and influencing factors among resident doctors at a tertiary health facility in Osun State, Nigeria [J]. Niger Med J, 2024, 65(5): 691-704. DOI: 10.60787/nmj-v65i3.536.

[19] 中国医师协会. 住院医师规范化培训基地标准(2022年版)[EB/OL].(2022-08-01)[2025-01-01]. https://www.pkufh.com/Html/News/Articles/52634.html.

[20] 中国医师协会. 住院医师规范化培训内容与标准(2022年版)[EB/OL]. (2022-08-01)[2025-01-01]. https://www.pkufh.com/Html/News/Articles/52635.html.

[21] 中国医师协会.《住院医师规范化培训基地临床技能培训中心工作指南(2023年版)》[EB/OL]. (2022-08-01)[2025-01-01]. https://www.pkufh.com/Html/News/Articles/52641.html.

[22] Braun V, Clarke V. Using thematic analysis in psychology [J]. Qual Res Psychol, 2006, 3(2): 77-101. DOI: 10.1191/1478088706qp063oa.

[23] 阎红琳, 严丹丹, 卢章洪, 等. 基于“扶—牵—引—放”四步教学法的病理学住院医师科研能力分层培养[J]. 诊断病理学杂志, 2023, 30(5): 517-518. DOI: 10.3969/j.issn.1007-8096.2023.05.034.

[24] 崔敏, 马颖, 丁宏, 等. 公共事业管理(卫生事业管理方向)专业本科毕业生职业基准胜任力评价指标研究[J]. 中华医学教育探索杂志, 2023, 22(10): 1476-1482. DOI: 10.3760/cma.j.cn116021-20211120-01444.

[25] 杨能瑞, 阳孟君, 周浩, 等. 腹腔镜模拟教学效果评价指标体系的构建和实证研究[J]. 中华医学教育探索杂志, 2022, 21(3): 272-276. DOI: 10.3760/cma.j.cn116021-20200506-00782.

[26] 张海英, 钱风华. 住院医师规范化培训职业精神和人文素养评价体系建设实践与探索[J]. 中国医院管理, 2020, 40(10): 94-96.

[27] 唐宝春, 文小玲, 黄丽, 等. 基于CE的评价指标体系对本科学业导师带教效果的干预[J]. 昆明医科大学学报, 2024, 45(11): 166-173. DOI: 10.12259/j.issn.2095-610X.S20241124.

[28] 石景芬, 王宇婵, 刘志会, 等. 区县级医院学科发展竞争力指标体系研究[J]. 中国医院, 2024, 28(5): 37-40. DOI: 10.19660/j.issn.1671-0592.2024.5.09.

(收稿日期:2025-04-15)

(本文编辑:唐宗顺)