

·论著·

基于深度学习的内镜超声胆管扫查辅助分站系统构建

黄丽 张军 吴慧玲 姚理文 邓涛 于红刚

武汉大学人民医院消化内科 消化疾病湖北省重点实验室, 武汉 430060

通信作者: 于红刚, Email: yuhonggang@whu.edu.cn

【摘要】 目的 构建基于深度学习的人工智能内镜超声(endoscopic ultrasonography, EUS)胆管扫查辅助分站系统,以期辅助医师学习多站成像技术,提高操作水平。**方法** 回顾性收集武汉大学人民医院和武汉协和医院2016年5月—2020年10月522例EUS视频资料,基于视频截取图像,获得来自武汉大学人民医院的3 000张白光胃镜、31 003张超声胃镜图像和来自武汉协和医院的799张超声胃镜图像,用于EUS胆管扫查系统的模型训练和模型测试。模型包括:白光胃镜图像过滤模型,标准站图像与非标准站图像区分模型和EUS胆管扫查标准图像分站模型,用以将标准图像分为肝窗、胃窗、球窗、降窗。然后从测试集中随机抽取110张图像进行人机大赛,比较专家、高级内镜医师与人工智能模型对胆管扫查多站成像每个站点识别的准确度。**结果** 白光胃镜图像过滤模型准确率为100.00%(1 200/1 200),标准站图像与非标准站图像区分模型准确率为93.36%(2 938/3 147),EUS胆管扫查标准图像分站模型在内部测试集中各分类的准确率分别为肝窗97.23%(1 687/1 735),胃窗96.89%(1 681/1 735),球窗98.73%(1 713/1 735),降窗97.18%(1 686/1 735);外部测试集中准确率分别为肝窗89.61%(716/799),胃窗92.74%(741/799),球窗90.11%(720/799),降窗92.24%(737/799)。人机大赛中,模型分站的正确率为89.09%(98/110),高于内镜医师[85.45%(94/110),74.55%(82/110),85.45%(94/110)],接近专家水平[92.73%(102/110),90.00%(99/110)]。**结论** 本研究构建了一种基于深度学习的EUS胆管扫查系统,可以较为准确地实时辅助内镜医师进行标准多站扫查,提高EUS完整性及操作质量。

【关键词】 人工智能; 深度学习; 内镜超声检查; 胆管; 多站成像技术

基金项目:湖北省重大科技创新项目(2018-916-000-008);湖北省消化疾病微创诊治医学临床研究中心项目(2018BCC337)

Development of auxiliary substation system for endoscopic ultrasound bile duct scanning based on deep learning

Huang Li, Zhang Jun, Wu Huiling, Yao Liwen, Deng Tao, Yu Honggang

Department of Gastroenterology, Renmin Hospital of Wuhan University; Key Laboratory of Hubei Province for Digestive Diseases, Wuhan 430060, China

Corresponding author: Yu Honggang, Email: yuhonggang@whu.edu.cn

【Abstract】 Objective To construct a deep learning-based artificial intelligence endoscopic ultrasound (EUS) bile duct scanning substation system to assist endoscopists in learning multi-station imaging and improve their operation skills. **Methods** A total of 522 EUS videos in Renmin Hospital of Wuhan University and Wuhan Union Hospital from May 2016 to October 2020 were collected, and images were captured from these videos, including 3 000 white light images and 31 003 EUS images from Renmin Hospital of Wuhan University, and 799 EUS images from Wuhan Union Hospital. The pictures were divided into training set and test set in the EUS bile duct scanning system. The system included filtering model of

DOI: 10.3760/cma.j.cn321463-20210628-00007

收稿日期 2021-06-28 本文编辑 朱悦

引用本文:黄丽,张军,吴慧玲,等.基于深度学习的内镜超声胆管扫查辅助分站系统构建[J].中华消化内镜杂志,2022,39(4):295-300. DOI: 10.3760/cma.j.cn321463-20210628-00007.



white light gastroscopy images (model 1), distinguishing model of standard station images and non-standard station images (model 2) and substation model of EUS bile duct scanning standard images (model 3), which were used to classify the standard images into liver window, stomach window, duodenal bulb window, and duodenal descending window. Then 110 pictures were randomly selected from the test set for a man-machine competition to compare the accuracy of multi-station imaging by experts, advanced endoscopists and the artificial intelligence model. **Results** The accuracies of model 1 and model 2 were 100.00% (1 200/1 200) and 93.36% (2 938/3 147) respectively. Those of model 3 on the internal validation dataset in each classification were 97.23% (1 687/1 735) in liver window, 96.89% (1 681/1 735) in stomach window, 98.73% (1 713/1 735) in duodenal bulb window, and 97.18% (1 686/1 735) in duodenal descending window. And those on the external validation dataset were 89.61% (716/799) in liver window, 92.74% (741/799) in stomach window, 90.11% (720/799) in duodenal bulb window, and 92.24% (737/799) in duodenal descending window. In the man-machine competition, the accuracy of the substation model was 89.09% (98/110), which was higher than that of senior endoscopists [85.45% (94/110), 74.55% (82/110), and 85.45% (94/110)] and close to the level of experts [92.73% (102/110) and 90.00% (99/110)]. **Conclusion** The deep learning-based EUS bile duct scanning system constructed in the current study can assist endoscopists to perform standard multi-station scanning in real time more accurately and improve the completeness and quality of EUS.

【Key words】 Artificial intelligence; Deep learning; Endoscopic ultrasonography; Bile ducts; Multi-station imaging

Fund program: Hubei Major Science and Technology Innovation Project (2018-916-000-008); Project of Hubei Clinical Research Center for Digestive Disease Minimally Invasive Incision (2018BCC337)

内镜超声检查(endoscopic ultrasonography, EUS)是一种结合了消化内镜和超声成像技术的消化内镜诊疗方法^[1],可用于检查胆总管结石等胆管疾病。相较于经内镜逆行胰胆管造影术(endoscopic retrograde cholangiopancreatography, ERCP),EUS对发现胆总管结石^[2],尤其是小微型结石的敏感度和成功率更高^[3]。多站成像技术是EUS进行胆管检查的一种标准扫查技术,将EUS扫查胆管分为以下4站:胃底(简称肝窗)、胃体及胃窦(简称胃窗)、十二指肠球部(简称球窗)、十二指肠降部(简称降窗)^[4-6]。多站成像既可以定位胆管扫查位置,辅助实现EUS引导下细针穿刺抽吸术,又能保证胆管扫查的完整性,提高胆管扫查质量。

然而,EUS学习难度较高,对经验丰富的内镜医师来说,EUS的操作技巧和图像解析都具有较高的挑战性,初学者尤甚^[7]。操作者图像解析能力的提高依赖于重复的高强度学习,但是国内能提供足够数量EUS病例的教学医院数量有限,无法满足临床实际培训需求^[8-9]。因此,开发一种实时EUS辅助增强操作培训工具对初学者尤为有利。近年来,基于深度学习的人工智能技术广泛应用于消化内镜领域,带来了消化内镜技术的巨大变革^[10-12]。Wu等^[13]利用基于深度学习的图像分类技术实现了胃镜检查盲区实时检测,提高了胃镜检查质量。但目前利用深度学习进行EUS胆管扫查的研究依然空白。

本研究训练了一种基于深度学习的EUS胆管扫查辅助分站系统,通过实时对EUS进行分站,辅助操作医师识别EUS扫查图像并定位,保证胆管扫查的完整性,同时可以辅助培训初学者进行规范的EUS胆管扫查学习。

资料与方法

一、数据收集

回顾性收集2016年5月—2020年10月武汉大学人民医院消化内镜中心502个病例和武汉协和医院20个病例的EUS视频资料,基于视频截取图像,根据需要将图片分类。标准图像:包含4个站点的标准图像;肝窗标准图像:超声探头位于肝窗站点时的标准扫查图像;胃窗标准图像:胃窗站点的标准扫查图像;球窗标准图像:十二指肠球部站点的标准扫查图像;降窗标准图像:十二指肠降部站点的标准扫查图像;非标准图像:除标准图像外其他EUS扫查过程中的超声成像。EUS胆管扫查标准图像示例见图1。

二、模型训练

EUS胆管扫查共包含3个模型。模型1:白光胃镜图像过滤模型,用于区分白光胃镜图像与超声图像;模型2:标准站图像与非标准站图像区分模型,用以区分标准站点图像和非标准站点图像;模型3:EUS胆管扫查标准图像分站模型,将标准站点

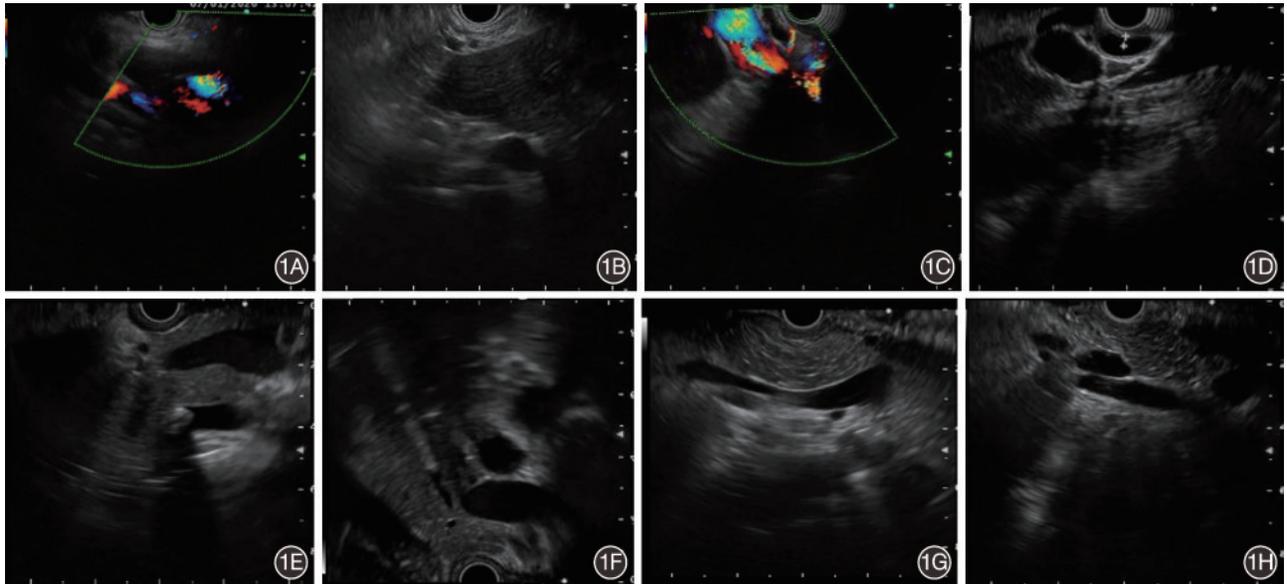


图1 内镜超声胆管扫查标准图像示例 1A、1B:肝窗标准图像示例图;1C、1D:球窗标准图像示例图;1E、1F:胃窗标准图像示例图;1G、1H:降窗标准图像示例图

图像分为肝窗、胃窗、球窗、降窗。模型的训练思路如图 2 所示。本研究采用残差神经网络 ResNet50 训练模型 1~3,所有输入图像统一转换为 512 像素×512 像素大小,进入残差神经网络 ResNet50 进行模型训练和验证。

三、图像资料

本研究共使用了来自武汉大学人民医院的 3 000 张白光胃镜图像,18 532 张非标准超声胃镜图像和 12 471 张 EUS 标准站图像进行模型的训练和内部验证。799 张来自武汉协和医院的图像进行外部测试。

1.模型 1:使用来自武汉大学人民医院的

3 000 张白光胃镜图像和 3 000 张非标准超声图像训练和测试模型 1,训练集和测试集的分配为 8:2,以过滤掉 EUS 过程中的白光胃镜部分。

2.模型 2:使用来自武汉大学人民医院的 14 120 张非标准图像和 10 736 张 EUS 标准站图像训练模型 2,1 412 张非标准图像和 1 735 张标准图像测试模型 2,以过滤掉非标准部分图像。

3.模型 3:使用来自武汉大学人民医院的 10 736 张 EUS 标准站图像(包括 1 063 张肝窗图像、5 875 张胃窗图像、1 430 张球窗图像和 2 368 张降窗图像)进行模型训练;1 735 张图像(包括 343 张肝窗图像、619 张胃窗图像、333 张球窗图像和

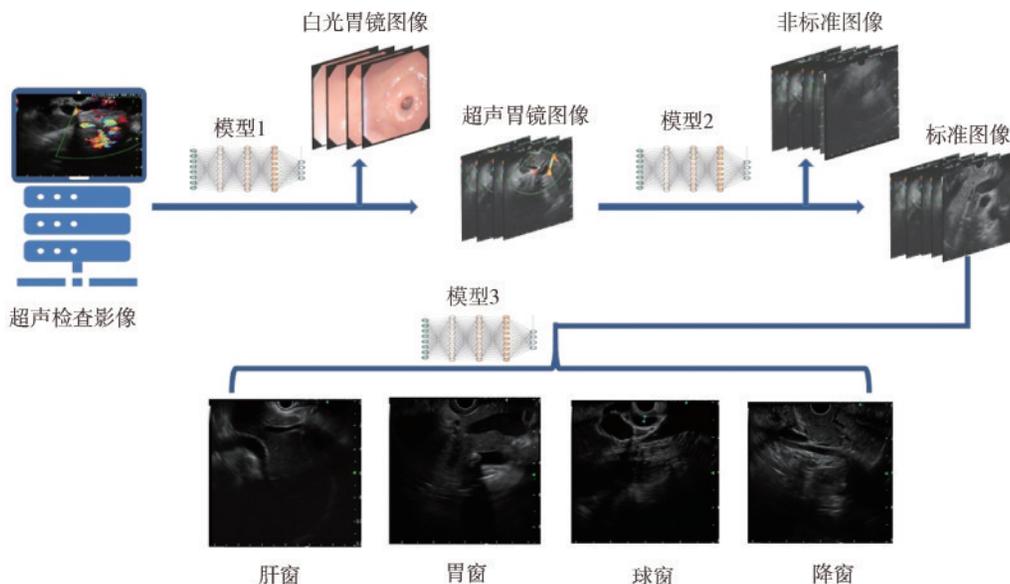


图2 胆管扫查分站模型构建示意图

440 张降窗图像)进行模型内部测试;799 张来自武汉协和医院的图像(包括 335 张肝窗图像、148 张胃窗图像、204 张球窗图像和 112 张降窗图像)进行外部测试。测试集和训练集相互独立,不源于同一个患者。

四、人机大赛

模型训练完成后,随机在模型 3 的测试集图像中抽取来自 44 例视频的 110 张 EUS 标准站图像,并邀请 2 位有 10 年以上 EUS 经验的专家和 3 位有 5~10 年 EUS 经验的高级内镜医师参与人机大赛,比较医师分站和机器分站的准确性。

五、数据统计

模型训练准确率=(真阳性图片数+真阴性图片数)/(真阳性图片数+真阴性图片数+假阳性图片数+假阴性图片数);特异度=真阴性图片数/(真阴性图片数+假阳性图片数);灵敏度=真阳性图片数/(真阳性图片数+假阴性图片数);人机大赛正确率=判断正确图片数/总图片数。

结 果

1. 白光胃镜图像过滤模型(模型 1)性能:模型 1 区分白光胃镜图像与超声胃镜图像的准确率为 100.00%(1 200/1 200),模型 1 能够完全正确区分两者,直接过滤掉白光胃镜图像。

2. 标准站图像与非标准站图像区分模型(模型 2)分类性能:模型 2 区分标准站与非标准站图像的准确率为 93.36%(2 938/3 147),说明模型性能能够筛选出标准站图像,进行站点分类。

3. 胆管分站模型(模型 3)性能:如表 1 所示,在内部测试集中,胆管分站模型对每一站的分类性能都表现出很好的效果,分站模型特异度高,均达到了 95% 以上,准确率均达到了 96% 以上。其中,球窗分类效果最好,准确率为 98.73%,特异度为

99.36%;胃窗分类效果最差,准确率为 96.89%,特异度为 95.43%。在外部测试集中,模型准确率、特异度均高于 89%,总体效果好,模型稳定,其中,胃窗准确率最高,为 92.74%,肝窗准确率最低,为 89.61%;肝窗特异度最高,为 99.14%,球窗特异度最低,为 89.58%。

4. 人机大赛结果:机器对于胆管超声扫查分站的正确率为 89.09%(98/110),低于 2 位高水平专家的结果[92.73%(102/110)和 90.00%(99/110)],高于 3 位有 5 年以上 EUS 操作经验的内镜医师[85.45%(94/110)、74.55%(82/110)和 85.45%(94/110)]。

5. 模型整合为系统:将模型进行封装整合,构建成为人工智能胆管扫查辅助分站系统,在进行 EUS 时,自动导入检查视频,可以实时提醒医师扫查站点覆盖情况(图 3)。

讨 论

EUS 是消化内镜不可或缺的检查技术之一,相对于 ERCP, EUS 发现胆管结石的敏感度更高,而且相对无创,并发症少,无辐射^[2]。欧洲胃肠病协会推荐对于中度怀疑有胆总管结石的患者首选 EUS^[14]。而 EUS 中的多站成像技术则可以通过分站扫查提高检查过程的完整性,保证检查质量。作为一种综合了内镜和超声的复杂内镜操作技术, EUS 的学习难度很高,除了常规的胃肠镜操作技术,操作者还需要掌握超声影像学的声学知识、影像分析等。由于 EUS 是探头在体内进行超声成像,操作者也需要具备良好的三维解剖认知能力。EUS 的诊疗过程强烈依赖内镜医师。所以,规范化的操作培训才能培养出合格的 EUS 医师,从而使 EUS 诊疗更为客观,这是 EUS 培训的立足之本^[15]。多站成像技术就是一种规范化的 EUS 操作技术,想

表 1 内镜超声胆管扫查标准图像分站模型分类准确度

测试集	站点	准确率(%)	灵敏度(%)	特异度(%)
内部测试集	肝窗	97.23(1 687/1 735)	90.67(311/343)	98.85(1 376/1 392)
	胃窗	96.89(1 681/1 735)	99.52(616/619)	95.43(1 065/1 116)
	球窗	98.73(1 713/1 735)	96.10(320/333)	99.36(1 393/1 402)
	降窗	97.18(1 686/1 735)	91.14(401/440)	99.23(1 285/1 295)
外部测试集	肝窗	89.61(716/799)	76.42(256/335)	99.14(460/464)
	胃窗	92.74(741/799)	88.51(131/148)	93.70(610/651)
	球窗	90.11(720/799)	91.67(187/204)	89.58(533/595)
	降窗	92.24(737/799)	75.00(84/112)	95.05(653/687)

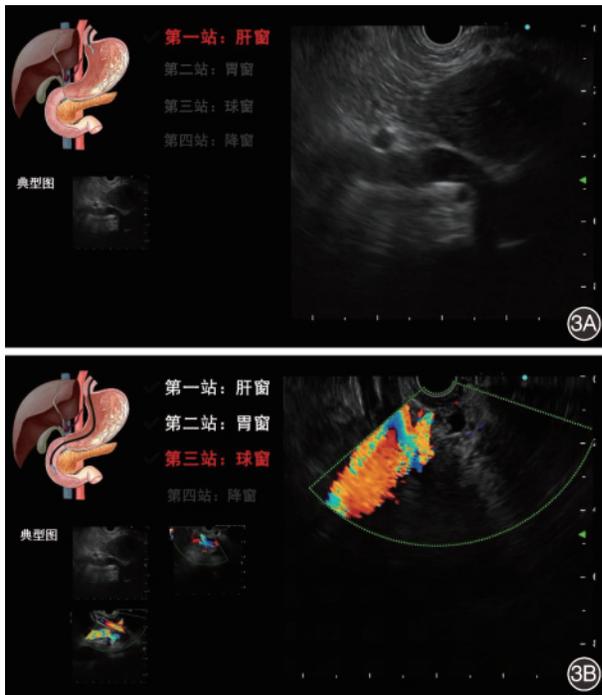


图3 内镜超声胆管扫查辅助分站系统自动分站界面展示
3A:当扫查到肝窗时,肝窗被点亮,左下角保留了1张肝窗典型图;3B:已完成肝窗、胃窗的扫查,正在扫查并点亮球窗,左下角分别保留了肝窗、胃窗和球窗的1张典型图片

要完成规范的EUS多站成像扫描,需要操作者准确理解并识别不同站点的EUS影像。

近年来,一些EUS培训着手于培训初学者多站成像技术。Wani等^[16-17]制定了基于多站成像技术的EUS培训4分制评分标准,用于评估操作者在每个站点的操作表现。研究发现,当操作者能被及时反馈位置和操作方式时,操作者得分会更高。但是在国内,具有足够病例数的EUS培训学习中心数量有限,为每一位操作者及时提供反馈的教学方式实行起来相对消耗人力和时间,对培训中心来说施行可能较为困难。

深度学习的提出带来了人工智能技术质的飞跃,并且临床实际应用证明了其辅助医师解决临床问题的能力^[18]。Zhang等^[19]基于深度学习开发了胰腺规范扫查质控系统;Bharti等^[20]提出了EUS影像中辅助鉴别慢性肝硬化与肝细胞癌的深度学习模型,其区分两者的准确率高达96.6%。

本研究中,我们构建了基于深度学习的人工智能EUS胆管扫查辅助分站系统,着眼于胆管系统的分站扫查,可以实时提示医师EUS胆管扫查分站覆盖情况,辅助医师提高胆管系统检查的完整性,提高EUS质量。同时,实时分站系统还能辅助医师明确EUS定位,从而有利于引导操作者进行EUS引导下细针穿刺抽吸术。组成此系统的3个模型准确

率高于90%,同时,在人机大赛中,胆管分站系统的准确率高于一般的高级内镜医师,接近专家水平,提示应用本系统可较好地培训EUS初学者,提高其影像识别准确率。

然而,本研究还有一些局限性。首先,本研究中胆管分站模型在外部测试集的准确度低于内部测试集,两者有一定差距。不同医院EUS的基础参数设置不同是导致这种差异最主要的原因。在后续研究中,我们要纳入尽可能多的EUS影像,囊括尽可能多的仪器和参数设置,来降低这种差异带来的影响,提高本系统的临床适用性,扩大应用范围。其次,虽然本研究提出了胆管辅助分站系统,帮助操作者进行标准胆管扫查,但是我们还未实时标注胆管,下一步研究中我们将在分站的基础上实现胆管的实时标注。

综上所述,本研究提出了一种基于深度学习的EUS胆管扫查系统,可以实时辅助内镜医师进行标准多站扫查,在未来的临床应用中可能提高初学者的EUS认知能力水平,加速学习过程,并提高其操作EUS的完整性及质量。

利益冲突 所有作者声明不存在利益冲突

作者贡献声明 黄丽:数据统计,模型训练,撰写论文;张军、邓涛、于红刚:设计试验,图片整理;吴慧玲、姚理文:收集并初步整理图片

参 考 文 献

- [1] Fickling WE, Wallace MB. Endoscopic ultrasound and upper gastrointestinal disorders[J]. J Clin Gastroenterol, 2003,36(2): 103-110. DOI: 10.1097/00004836-200302000-00004.
- [2] Kondo S, Isayama H, Akahane M, et al. Detection of common bile duct stones: comparison between endoscopic ultrasonography, magnetic resonance cholangiography, and helical-computed-tomographic cholangiography[J]. Eur J Radiol, 2005, 54(2): 271-275. DOI: 10.1016/j.ejrad.2004.07.007.
- [3] Artifon EL, Kumar A, Eloubeidi MA, et al. Prospective randomized trial of EUS versus ERCP-guided common bile duct stone removal: an interim report (with video) [J]. Gastrointest Endosc, 2009, 69(2): 238-243. DOI: 10.1016/j.gie.2008.05.020.
- [4] Irisawa A, Yamao K. Curved linear array EUS technique in the pancreas and biliary tree: focusing on the stations[J]. Gastrointest Endosc, 2009,69(2 Suppl):S84-89. DOI: 10.1016/j.gie.2008.12.006.
- [5] Sharma M, Pathak A, Shoukat A, et al. Imaging of common bile duct by linear endoscopic ultrasound[J]. World J Gastrointest Endosc, 2015, 7(15): 1170-1180. DOI: 10.4253/wjge.v7.i15.1170.
- [6] EUS-FNA Standardization Committee, Yamao K, Irisawa A, et al. Standard Imaging techniques of endoscopic

- ultrasound-guided fine-needle aspiration using a curved linear array echoendoscope[J]. Digestive Endoscopy, 2007, 19(s1): S180-S205. DOI: 10.1111/j.1443-1661.2007.00742.x.
- [7] Cho CM. Training in endoscopy: endoscopic ultrasound[J]. Clin Endosc, 2017,50(4):340-344. DOI: 10.5946/ce.2017.067.
- [8] Kim GH, Bang SJ, Hwang JH. Learning models for endoscopic ultrasonography in gastrointestinal endoscopy[J]. World J Gastroenterol, 2015,21(17):5176-5182. DOI: 10.3748/wjg.v21.i17.5176.
- [9] 王文海, 李鹏, 王拥军, 等. 规范化消化超声内镜培训方法的实践[J]. 中华医学教育探索杂志, 2020, 19(3): 344-347. DOI: 10.3760/cma.j.cn116021-20190518-00081.
- [10] 陈佳敏, 陈肖, 蔡建庭, 等. 人工智能在消化内镜中的临床应用现状[J]. 中华消化内镜杂志, 2020, 37(5): 305-308. DOI: 10.3760/cma.j.cn321463-20191222-00223.
- [11] 于红刚. 消化内镜人工智能的现状与展望[J]. 中华消化内镜杂志, 2019, 36(4): 229-232. DOI: 10.3760/cma.j.issn.1007-5232.2019.04.001.
- [12] Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks[J]. Commun ACM, 2017,60(6):84-90. DOI: 10.1145/3065386.
- [13] Wu L, Zhou W, Wan X, et al. A deep neural network improves endoscopic detection of early gastric cancer without blind spots[J]. Endoscopy, 2019, 51(6): 522-531. DOI: 10.1055/a-0855-3532.
- [14] Manes G, Paspatis G, Aabakken L, et al. Endoscopic management of common bile duct stones: European Society of Gastrointestinal Endoscopy (ESGE) guideline[J]. Endoscopy, 2019,51(5):472-491. DOI: 10.1055/a-0862-0346.
- [15] 诸琦. 规范化培训是内镜超声检查的立足之本[J]. 诊断学理论与实践, 2008, 7(6): 581-583.
- [16] Wani S, Han S, Simon V, et al. Setting minimum standards for training in EUS and ERCP: results from a prospective multicenter study evaluating learning curves and competence among advanced endoscopy trainees[J]. Gastrointest Endosc, 2019,89(6):1160-1168. DOI: 10.1016/j.gie.2019.01.030.
- [17] Wani S, Hall M, Keswani RN, et al. Variation in aptitude of trainees in endoscopic ultrasonography, based on cumulative sum analysis[J]. Clin Gastroenterol Hepatol, 2015, 13(7): 1318-1325. DOI: 10.1016/j.cgh.2014.11.008.
- [18] Gong D, Wu L, Zhang J, et al. Detection of colorectal adenomas with a real-time computer-aided system (ENDOANGEL): a randomised controlled study[J]. Lancet Gastroenterol Hepatol, 2020, 5(4): 352-361. DOI: 10.1016/S2468-1253(19)30413-3.
- [19] Zhang J, Zhu L, Yao L, et al. Deep learning-based pancreas segmentation and station recognition system in EUS: development and validation of a useful training tool (with video) [J]. Gastrointest Endosc, 2020, 92(4): 874-885. DOI: 10.1016/j.gie.2020.04.071.
- [20] Bharti P, Mittal D, Ananthasivan R. Preliminary study of chronic liver classification on ultrasound images using an ensemble model[J]. Ultrason Imaging, 2018, 40(6): 357-379. DOI: 10.1177/0161734618787447.

• 读者 • 作者 • 编者 •

《中华消化内镜杂志》对来稿中统计学处理的有关要求

1. 统计研究设计: 应交代统计研究设计的名称和主要做法。如调查设计(分为前瞻性、回顾性或横断面调查研究); 实验设计(应交代具体的设计类型, 如自身配对设计、成组设计、交叉设计、析因设计、正交设计等); 临床试验设计(应交代属于第几期临床试验, 采用了何种盲法措施等)。主要做法应围绕 4 个基本原则(随机、对照、重复、均衡)概要说明, 尤其要交代如何控制重要非试验因素的干扰和影响。

2. 资料的表达与描述: 用 $\bar{x} \pm s$ 表达近似服从正态分布的定量资料, 用 $M(Q_1, Q_3)$ 或 $M(IQR)$ 表达呈偏态分布的定量资料; 用统计表时, 要合理安排纵横标目, 并将数据的含义表达清楚; 用统计图时, 所用统计图的类型应与资料性质相匹配, 并使数轴上刻度值的标法符合数学原则; 用相对数时, 分母不宜小于 20, 要注意区分百分率与百分比。

3. 统计学分析方法的选择: 对于定量资料, 应根据所采用的设计类型、资料所具备的条件和分析目的, 选用合适的统计学分析方法, 不应盲目套用 t 检验和单因素方差分析; 对于定性资料, 应根据所采用的设计类型、定性变量的性质和频数所具备的条件以及分析目的, 选用合适的统计学分析方法, 不应盲目套用 χ^2 检验。对于回归分析, 应结合专业知识和散布图, 选用合适的回归类型, 不应盲目套用简单直线回归分析, 对具有重复实验数据的回归分析资料, 不应简单化处理; 对于多因素、多指标资料, 要在一元分析的基础上, 尽可能运用多元统计学分析方法, 以便对因素之间的交互作用和多指标之间的内在联系进行全面、合理的解释和评价。

4. 统计结果的解释和表达: 当 $P < 0.05$ (或 $P < 0.01$) 时, 应说明对比组之间的差异有统计学意义, 而不应说对比组之间具有显著性(或非常显著性)的差别; 应写明所用统计学分析方法的具体名称(如: 成组设计资料的 t 检验、两因素析因设计资料的方差分析、多个均数之间两两比较的 q 检验等), 统计量的具体值(如 t 值, χ^2 值, F 值等)应尽可能给出具体的 P 值; 当涉及到总体参数(如总体均数、总体率等)时, 在给出显著性检验结果的同时, 再给出 95% 可信区间。