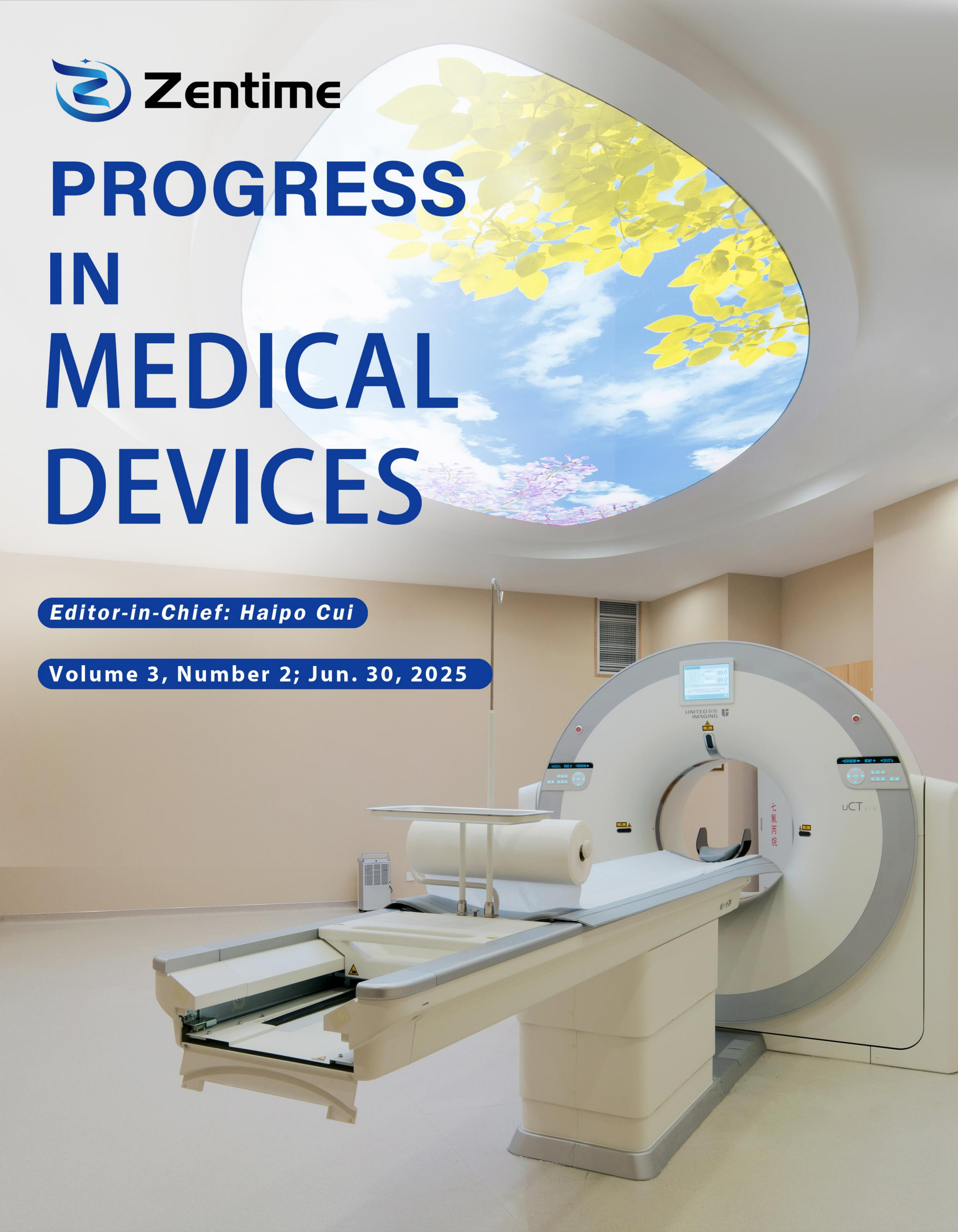
Zentime

PROGRESS IN MEDICAL DEVICES



Editor-in-Chief: Haipo Cui

Volume 3, Number 2; Jun. 30, 2025

ISSN: 2957-5478

Published by ZENTIME PUBLISHING CORPORATION

Progress in Medical Devices
ISSN: 2957-5478

Zentime

Review Article https://doi.org/10.61189/489352mmqzpk

Research progress on fixation methods for mandibular defect reconstruction

Xiaohan Yu[1], Haipo Cui[1], Hui Zhong[2]

[1]Shanghai Institute for Minimally Invasive Therapy, University of Shanghai for Science and Technology, Shanghai 200093, China. [2]Medical Engineering Department of Northern, Jiangsu People's Hospital, Yangzhou 225001, Jiangsu Province, China.

Corresponding authors: Haipo Cui and Hui Zhong.

Declaration of conflict of interest: None.

Received November 13, 2024; Accepted March 06, 2025; Published June 30, 2025

Highlights

- Autologous bone grafting remains the primary method for mandibular reconstruction.
- The integration of computer-aided design and 3D printing enhances surgical precision and safety.
- Distraction osteogenesis is a promising alternative, reducing donor site morbidity.
- The mortise-and-tenon structure is a novel fixation method.

Abstract

The mandible plays a crucial role in facial structure and function. Mandibular defects, often caused by tumor resection, trauma, or congenital conditions, severely impact patients' physiological functions and quality of life. Reconstruction methods for mandibular defects primarily include autologous bone grafting, computer-aided reconstruction, and distraction osteogenesis. Among these, autologous bone grafting remains the most widely used technique. As medical technologies advance, the integration of computer-aided design and 3D printing has provided new technical support for personalized mandibular reconstruction, improving both surgical precision and safety. Moreover, distraction osteogenesis has emerged as a promising alternative in recent years. This paper reviews research on fixation methods for V mandibular defect reconstruction to offer clinical guidance for effective mandibular repair.

Keywords: Mandibular defect, reconstruction, autologous bone grafting, 3D printing, distraction osteogenesis

Introduction

The mandible is in the lower third of the face and plays a crucial role in maintaining facial structure and key physiological activities such as chewing and speech. Due to its unique anatomical structure, it is central to oral and maxillofacial surgery. Mandibular defects may arise from tumor resection, trauma, or congenital deformities, leading to varying degrees of impairment in both facial aesthetics and function. The location and size of the defect are key factors in determining its impact on the patient, making the classification of mandibular defects crucial for selecting appropriate treatment strategies and evaluating postoperative functional recovery [1].

A widely used classification system is the HLC classification, proposed by Jewer et al., which categorizes mandibular defects based on location and extent. In this system, "H" refers to unilateral mandibular defects that do not extend beyond the midline, often leading to joint dysfunction and impaired mandibular movement [2]. "C" denotes defects in the anterior mandibular region, between the canines, primarily affecting aesthetics and occlusal function. "L" represents defects in the lateral mandible (excluding the condyle), potentially impairing masticatory muscle function and mandibular mobility. Boyd et al. later refined this classification by adding subdivisions for soft tissue defects, enhancing its clinical relevance in mandibular reconstruction [3].

Address correspondence to: Haipo Cui, Shanghai Institute for Minimally Invasive Therapy, University of Shanghai for Science and Technology, No. 580, Gongjun Road, Yangpu District, Shanghai 200093, China. Email: h_b_cui@163.com; **Hui Zhong**, Medical Engineering Department of Northern, Jiangsu People's Hospital, No. 98, Nantong West Road, Guangling District, Yangzhou 225001, Jiangsu Province, China. Email: yzzhonghui@126.com.

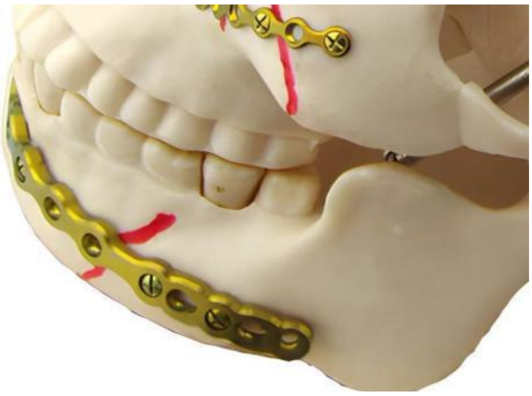


Figure 1. Titanium Plate Fixation.

Mandibular reconstruction is essential for restoring both appearance and functionality. The ideal reconstruction should achieve three primary goals: (i) restoring facial appearance as close to normal as possible, (ii) achieving normal dental alignment and occlusion, and (iii) preserving or approximating normal chewing and speech functions. Mandibular reconstruction has long been a focal point of medical research, with oral and maxillofacial surgeons constantly exploring more effective methods to restore both the form and function of the mandible [4].

Methods for mandibular defect reconstruction

Modern mandibular reconstruction surgery began to emerge in the early 20th century. Over the past century, significant advancement in research have driven rapid progress in the techniques and materials used for mandibular defect repair. Currently, various reconstruction methods are available, with the specific choice depending on the defect's location, cause, and severity [5]. While multiple methods have been applied to mandibular defect reconstruction, autologous bone grafting remains the cornerstone of the reconstruction process. Recent advances in computer technology have significantly enhanced mandibular repair techniques. The integration of computer-assisted technology and personalized reconstruction offers substantial technical support for functional mandibular restoration. For instance, Yoonjoo introduced a case of prosthetic reconstruction of maxillary defects in an elderly patient caused by a traumatic fall, demonstrating the application value of prosthetic technology in jaw reconstruction [6]. Regarding autologous bone grafting, a 2020 study by Samy et al. proposed a comprehensive digital and prosthetic-guided protocol for inserting dental implants into autogenous free bone grafts following segmental mandibular defect reconstruction [7]. Furthermore, Nobis et al. developed a template tool to assist fibular osteotomy in mandibular defect reconstruction through digital analysis of the human mandible [8]. These studies highlight the potential of digital technology to enhance the precision and outcomes of mandibular reconstruction surgery. Additionally, the application of distraction osteogenesis has opened new pathways for mandibular reconstruction, further expanding the available options for reconstruction [9].

Traditional titanium plate internal fixation

In mandibular reconstruction using autologous bone grafts, achieving stable fixation is crucial. The quality of fixation directly influences the success of fibular grafts, as inadequate fixation often leads to graft infection or necrosis. Due to the wide range of mandibular movements and the complex load-bearing environment, as shown in **Figure 1**, rigid internal fixation using titanium plates remains the most common fixation method. This technique stabilizes the graft by securing titanium or titanium alloy plates and screws at both ends of the fracture line, promoting bone healing. A 2024 prospective comparative study by Sugumar et al. evaluated the functional and radiographic outcomes of open reduction internal fixation and minimally invasive plate fixation for humeral shaft fractures, highlighting that traditional titanium plate fixation is perpendicular to the occlusal and muscular force trajectories of the mandible, which is inconsistent with the mandibular biomechanical load transfer mechanism [10]. Additionally, titanium implants may interfere with radiotherapy, potentially increasing the risk of complications like osteoradionecrosis.

Despite its widespread use, rigid internal fixation with titanium plates has notable limitations. The perpendicular alignment of the fixation method to the occlusal and muscular force trajectories contrasts with the mandible's natural biomechanical load transfer pathway. Furthermore, titanium implants may interfere with radiotherapy, heightening the risk of complications such as osteoradionecrosis.

With advancements in medical technology, the concept of semi-rigid fixation has recently gained attention. Semi-rigid fixation implants, such as sliding plates, allow for controlled micro-movements, which reduce the forces applied to the bone-metal interface during strain and promote localized bone healing. This approach offers advantages in stress distribution and biomechanics, but further research and validation are required to establish its applica-

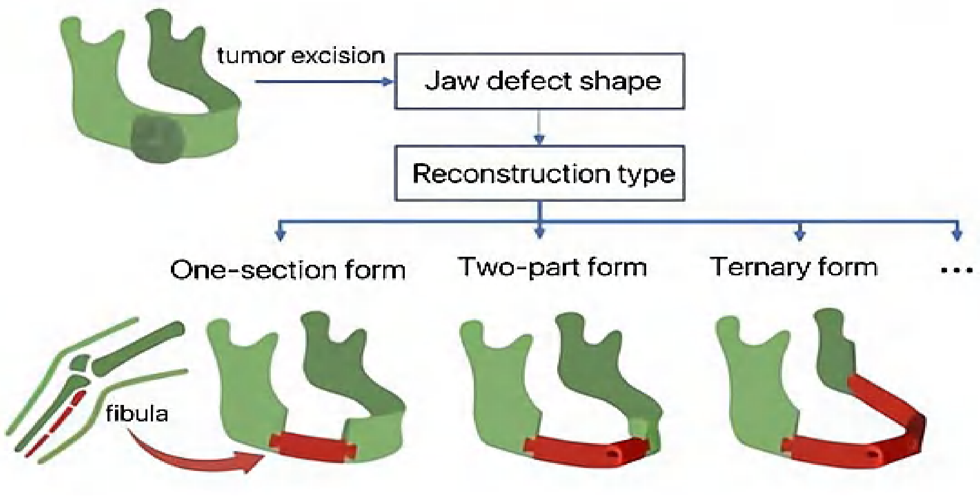


Figure 2. Mandibular Reconstruction Theory Based on Mortise and Tenon Architectural Technique (Selection of Mortise and Tenon Structure).

tion. In a 2024 study, Wang et al. introduced a specially contoured pelvic rim reconstruction titanium plate combined with cross-plate support screws (quadrangular screws) for treating acetabular fractures involving the quadrilateral plate through an anterior ilioinguinal approach [11]. The semi-rigid fixation concept demonstrated in this study offers new insights into its potential application in mandibular reconstruction. Additionally, Cui et al. introduced a novel base-fixation system and evaluated it using three-dimensional reconstruction and finite element analysis [12]. The results showed that compared to conventional buccal fixation, the base-fixation system exhibited superior fixation ability with a safety factor of 3.8, making it suitable for clinical applications. This method also reduced the volume of fixation titanium plates, shortened surgery duration, and minimized the risk of wound infection, presenting a promising alternative for mandibular defect reconstruction. Furthermore, Chen et al. conducted a finite element analysis comparing the biomechanical performance of polyetheretherketone (PEEK) and titanium alloys in mandibular defect reconstruction [13]. The study found that glass fiber-reinforced PEEK (GFR-PEEK) generated excessive stress beyond the yield strength of the mandible when loaded in the anterior region, making it unsuitable for defect repair. However, carbon fiber-reinforced PEEK (CFR-PEEK) exhibited more favorable stress distribution, positioning it as a more stable candidate for fixation systems.

Mortise and tenon structure-based fixation method

Inspired by the traditional Chinese “mortise and tenon” architectural technique, the authors propose a novel fibula-mandible semi-rigid fixation theory based on this structure, combined with tension screws, as shown in **Figure 2**. The mortise and tenon structure, with a long history in China, is renowned for its precise structural design and mechanical stability. Particularly, the dovetail mortise structure, consisting of a head tenon and tail tenon, offers enhanced structural strength and tensile resistance through precise proportion control. Finite element analysis shows that this fixation, combined with tension screws, offers superior biomechanical distribution in the mandible compared to traditional rigid internal fixation.

Inspired by this concept, researchers may explore its application in mandibular reconstruction by simulating a mortise and tenon structure between the fibula and mandible, achieving a fixation effect that is more consistent with the mandible’s biomechanical properties. Xin et al. conducted a preclinical trial assessing the biomechanical performance and osteogenic capability of 3D-printed polymeric scaffold implants for repairing critical-sized mandibular defects [14]. The study incorporated in silico finite element modeling, in vivo implantation in sheep, and in vitro mechanical testing. Results showed that the selective laser sintering-printed PEEK scaffolds exhibited excellent bone-scaffold integration and new bone formation at the scaffold-bone interface, further validating the potential of 3D-printed polymeric biomaterials for mandibular defect reconstruction.

In a 2023 study, Yu et al. explored terahertz reconfigurable multifunctional metamaterials based on 3D-printed mortise and tenon structures [15]. Their in-depth analysis and applica-

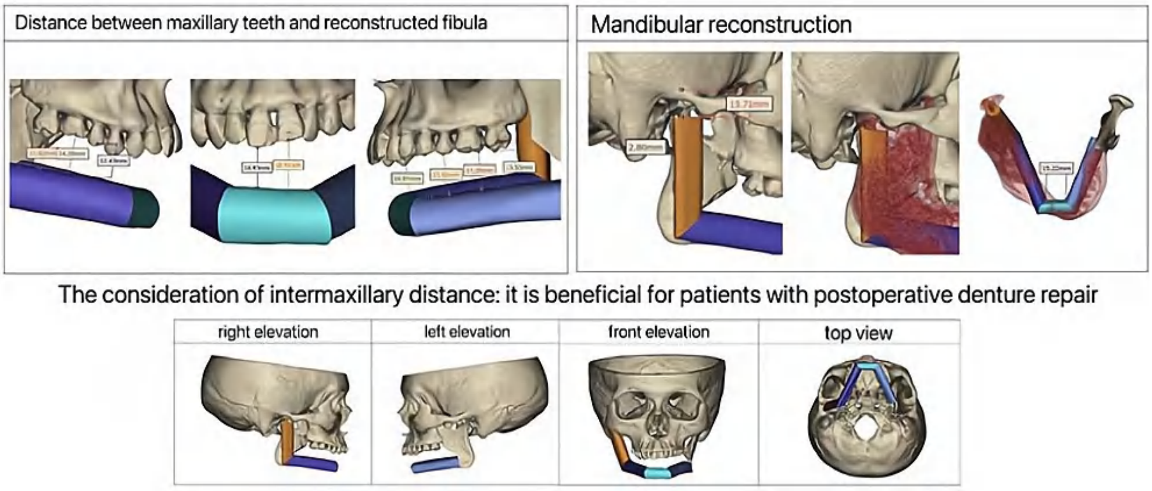


Figure 3. 3D Simulation of Mandibular Reconstruction.

tion of the mortise and tenon structure provide new inspiration for mandibular reconstruction. Finite element analysis indicates that the dovetail mortise and tenon structure, combined with tension screws, yields superior biomechanical distribution in the mandible compared to traditional rigid internal fixation. Zhang et al. investigated the use of customized bioceramic scaffolds and metal meshes for large-size mandibular defect regeneration and repair [16]. The study demonstrated that magnesium-doped calcium silicate and α-tricalcium phosphate scaffolds significantly promoted osteogenesis compared to β- tricalcium phosphate. Moreover, magnesium-doped calcium silicate exhibited superior mechanical properties during the early stage of bone regeneration, suggesting that combining high-strength bioactive scaffolds with titanium meshes could be a promising strategy for repairing large, load-bearing mandibular defects.

The core advantage of the mortise and tenon structure lies in its natural stress-dispersion capability. While traditional titanium plate fixation provides sufficient rigidity, its excessive rigidity often causes stress concentration, particularly under the complex mechanical forces generated during chewing and speaking. In contrast, the mortise and tenon structure, with its unique interlocking design, allows for micro-movements that naturally disperse external forces under load. This effectively reduces mechanical friction between the grafted segment and surrounding tissues, lowering the incidence of complications and decreasing the risk of postoperative bone resorption. Furthermore, the mortise and tenon structure offers improved shear and tensile resistance, ensuring the stability of the grafted bone segment.

Although the biomechanical advantages of the mortise and tenon structure have been preliminarily validated, its clinical application remains in an exploratory phase. Current research mainly relies on finite element simulation models, and the precise design of the mortise and tenon structure based on the type and size of mandibular defects in clinical practice remains an area requiring in-depth investigation. For instance, design parameters such as the angle, size, and fit of the mortise and tenon structure directly affects its mechanical performance. Future research should leverage personalized three-dimensional imaging and simulation technology to further optimize the design of the mortise and tenon structure, thereby enhancing its clinical applicability.

Computer-aided mandibular reconstruction

In recent years, with the continuous development of medical imaging technology and computer science, digital surgical techniques have gained widespread recognition and application in maxillofacial surgery. To achieve personalized restoration of mandibular appearance and function with optimal reconstruction accuracy, scholars worldwide have applied digital surgical technologies to repair segmental mandibular defects. In 1983, Hemmy et al. first utilized 3D CT technology for the diagnosis and treatment of cranio-maxillofacial diseases, a groundbreaking advancement that opened new directions for maxillofacial surgery [17].

With ongoing advancements in computer technology, computer-aided design (CAD)/manufacturing technology based on 3D reconstruction and computer-assisted navigation systems have been widely applied in the maxillofacial field, as shown in **Figure 3**. This technology acquires 3D data of lesions using imaging techniques such as CT or MRI, generates 3D models through computer systems, and conducts simulation analyses to formulate scientifically sound treatment plans. It aids surgeons in preoperative planning, including defining resection boundaries, determining the shape and position of reconstruction materials, and designing the optimal surgical approach. This approach ensures intraoperative precision, predicts postoperative outcomes, and enables the production of customized prostheses, thereby supporting personalized medicine in clinical practice.

In a 2024 study, Bagheri et al. used finite element analysis to explore the application of a new intramedullary internal distractor in distraction osteogenesis [18]. Based on a patient's 3D model, 3D printing technology can be used to produce surgical guides, such as osteotomy guides and shaping guides, which assist surgeons in precise cutting and fixation of bone tissue during surgery. Additionally, 3D printing enables the creation of customized titanium plates, prostheses, or implants to ensure a perfect fit between the grafted bone and the mandibular defect area [19]. Xue et al. conducted a study involving 3D digital reconstruction for 26 patients who underwent mandibular resection from 2014 to 2018, examining the application value of 3D printing technology in preoperative planning and intraoperative guidance for mandibular reconstruction surgery [20]. The results indicated that combining digital reconstruction with 3D printing significantly enhanced the precision of mandibular repair surgery.

Yu et al. from the First Affiliated Hospital of Zhejiang University School of Medicine investigated the application of a novel surgical guide in achieving positional accuracy for mandibular reconstruction and implants, finding that the guide improved the accuracy of mandibular repair, increased both the fit of the prosthesis, and enhanced patient satisfaction [21]. Liang et al. developed a digital surgical guide technique, combined with the Combined Osteotomy and Reconstruction Pre-bent Plate Positioning technique, which reduced errors between the pre-bent titanium plate and reconstructed mandible, yielding satisfactory clinical results [22]. Studies have shown that 3D printing technology achieves submillimeter accuracy, as demonstrated in mandibular reconstruction procedures where bone displacement errors are maintained below 1.27 mm, ensuring patient-specific anatomical solutions. Furthermore, it significantly enhances clinical quality and efficiency by reducing the production time of facial prostheses from weeks to 2–3 days, while achieving 100% aesthetic approval rates in prosthetic rehabilitation. This advancement markedly improves patient satisfaction and reduces postoperative complications [23, 24].

Yan et al. studied the application of computer-assisted surgical systems combined with 3D printing technology in free fibula flap reconstruction for mandibular defects, outlining three main advantages of the technology [25].

Firstly, it allows for personalized customization. By obtaining precise anatomical information of the patient's mandible through CT scanning, this approach provides a foundation for tailored reconstruction. Using CAD technology, surgeons can precisely design and simulate a reconstruction plan based on the patient's mandibular defect, osteotomy range, and repair requirements, ensuring that the reconstruction closely matches the patient's unique anatomical structure. Subsequently, 3D printing technology converts the virtual design into a physical prosthesis, completing the entire personalized reconstruction process, which significantly enhances surgical outcomes and the potential for functional recovery.

Secondly, this technology improves surgical precision. By transforming virtual images into 3D models, computer-assisted systems and 3D printing enable a more accurate understanding of the patient's physiological structure, thereby increasing diagnostic precision. During surgery, the computer-assisted system provides precise navigation and visualization, aiding the surgeon in performing more accurate procedures and improving surgical success rates. Particularly for free fibula flap transfers, computer systems secure and pre-bend titanium plates, enhancing stability at the mandibular endpoint, which is critical for accurate temporomandibular joint reconstruction. Furthermore, designing the surgical plan with computer software and creating a guide ensures that osteotomy shaping perfectly aligns with the defect site, further enhancing the precision of mandibular reconstruction.

Lastly, technology enhances surgical safety. Surgeons utilize 3D reconstruction tools and surgical simulation systems to gain a comprehensive understanding of the patient's anatomy, accurately define the surgical resection boundaries, and ensure predictability and precision in surgery. This predictability enables surgeons to anticipate and manage potential intraoperative issues, reducing damage to blood vessels and surrounding normal tissues, thus significantly increasing surgical safety. Additionally, the technology broadens the surgical field of vision, enhancing operational flexibility and autonomy, allowing for minimized incisions while reducing functional damage to the reconstructed tissue structure to the lowest possible level [25]. Clinical studies indicate that this technology effectively reduces surgical time, decreases blood loss, accelerates postoperative recovery, and improves patient satisfaction with appearance, chewing, and phonation functions while reducing postoperative complications. Kang et al. developed a novel functional repair method combining a 3D-printed PEEK implant with a free vascularized fibula graft, assessing its performance under masticatory motion [26]. Finite element analysis indicated that the safety factor exceeded 2.3, and the deformation of the reconstructed model was lower than that of a normal mandible during most clenching tasks, ensuring primary stability. Additionally, this method demonstrated superior clinical outcomes compared to conventional fibular grafting, offering enhanced safety, stability, and aesthetic restoration.

Distraction osteogenesis

Distraction osteogenesis (DO) has become increasingly common in mandibular reconstruction, especially for treating mandibular defects and deformities, where it demonstrates significant advantages. Compared to traditional bone grafting methods, DO minimizes donor site damage and reduces the risk of immune rejection. By utilizing the principle of tension-stress bone regeneration, it facilitates simultaneous growth of bone and soft tissue, achieving a more natural repair outcome [27]. The core concept of DO involves using a distraction device to gradually widen the space between bone segments or osteotomized bones, at a specified rate and frequency, stimulating new bone formation and enabling bone lengthening or widening. This technique is particularly suitable for complex mandibular reconstructions, including segmental defects, condylar loss, and congenital or acquired craniofacial deformities [28].

In the 1990s, DO began to be used in oral and maxillofacial surgery, primarily for correcting maxillofacial deformities and repairing bone defects. In 1992, McCarthy et al. first attempted to use DO to treat hemifacial microsomia, which marked the beginning of the technique's widespread application. Su et al. performed tumor resection surgeries on six patients with ameloblastoma or odontogenic keratocyst and used DO to repair the mandibular defects, further validating the feasibility and clinical value of this technique for broader application [29].

Huang et al. investigated the clinical application of pedicled mandibular osteomuscular flaps in the reconstruction of acquired segmental mandibular defects, assessing postoperative speech, swallowing function, and aesthetic outcomes [30]. The study found that all flaps remained viable, with an average speech function score of 7.6 ± 0.6 and satisfactory masticatory efficiency. The Visual Analog Scale score for appearance was 7.8 ± 0.8. These results demonstrate that this method is an effective option for secondary repair, offering functional and aesthetic benefits.

Conclusion

The mandible is a central component of the maxillofacial region, and repairing its defects and deformities is crucial for patients. In recent years, fixation techniques for mandibular defect reconstruction have made significant advancements. Among the various methods, autologous bone grafting, particularly vascularized grafts, remains the preferred technique for defect repair. Notably, the application of digital technologies, such as CAD and 3D printing, has enabled personalized mandibular reconstruction, significantly enhancing surgical success rates, precision, and patient satisfaction postoperatively. Additionally, distraction osteogenesis is emerging as an effective alternative, offering advantages such as reduced donor site damage, avoidance of immune rejection, and the ability to simultaneously repair of soft tissue and bone structures.

Future research should focus on further optimizing these techniques and exploring their applications for different types of mandibular defect. The success of mandibular defect reconstruction depends not only on advanced technologies and materials but also on appropriate fixation methods to ensure graft stability and effective postoperative recovery. This review systematically examines existing mandibular fixation techniques, analyzes the strengths and weaknesses of various methods, and introduces a novel mortise-and-tenon-based fixation approach. While this approach shows clear biomechanical advantages, it is still in the research and development phase, necessitating further clinical trials and validation studies for widespread clinical adoption.

Author contributions: Xiaohan Yu: Conceptualization, Methodology, Investigation, Data Curation, Writing-Original Draft. Haipo Cui: Conceptualization, Methodology, Validation, Supervision, Writing-Review & Editing. Hui Zhong: Supervision, Writing-Review & Editing.

References

[1] Sclaroff A, Haughey B, Gay WD, et al. Immediate mandibular reconstruction and placement of dental implants. At the time of ablative surgery. Oral Surg Oral Med Oral Pathol 1994;78(6):711-717.

[2] Jewer DD, Boyd JB, Manktelow RT, et al. Orofacial and mandibular reconstruction with the iliac crest free flap: a review of 60 cases and a new method of classification. Plast Reconstr Surg 1989;84(3):391-395.

[3] Boyd JB, Gullane PJ, Rotstein LE, et al. Classification of mandibular defects. Plast Reconstr Surg 1993;92(7):1266-1275.

[4] Jia J, Zhang H, Deng Y. Research on methods of mandibular defect repair. J Luzhou Med Coll 2010;33(5):542-544.

[5] Gilbert RW. Reconstruction of the oral cavity; past, present and future. Oral Oncol 2020;108:104683.

[6] Yoonjoo B, Sunyoung C. Prosthetic reconstruction of maxillary defect resulting from a traumatic fall in an elderly patient: A case report. J Korean Acad Prosthodont 2019;57(1):75-80.

[7] Samy M, Mohamed M, Amr G. Full-staged digital and prosthetic guided protocol for the insertion of dental implants in autogenous free bone grafts after reconstruction of segmental mandibular defects. Oral Maxillofac Surg 2020;24(2):189-201.

[8] Nobis CP, Kesting MR, Wolff KD, et al. Development of a template tool for facilitating fibula osteotomy in reconstruction of mandibular defects by digital analysis of the human mandible. Clin Oral Investig 2020;24(9):3077-3083.

[9] Liu DX, Dou ZQ, Niu YM, et al. Application and research progress of distraction osteogenesis in oral and maxillofacial surgery. J Clin Stomatol 2014;30(7):442-444.

[10] Sugumar N, Sathiyaseelan N, Purushothaman JR, et al. Assessing functional and radiological outcomes: open reduction and internal fixation vs. minimally invasive plate osteosynthesis for humerus shaft fractures - a prospective comparative study. Int Orthop 2024;48(11):2979-2991.

[11] Wang W, Cai X, Liu X, et al. Special contoured pelvic brim reconstruction titanium plate combined with trans-plate buttress screws (quadrilateral screws) for acetabular fractures with quadrilateral plate involvement through the anterior ilioinguinal approach. Front Surg 2024;11:1438036.

[12] Cui H, Gao L, Han J, et al. Biomechanical analysis of mandibular defect reconstruction based on a new base-fixation system. Comput Methods Biomech Biomed Engin 2022;25(14):1618-1628.

[13] Chen L, Gao L, Cui H, et al. Finite element comparison of titanium and polyetheretherketone materials for mandibular defect reconstruction. Am J Transl Res 2024;16(10):6097-6105.

[14] Xin H, Ferguson B M, Wan B, et al. A Preclinical Trial Protocol Using an Ovine Model to Assess Scaffold Implant Biomaterials for Repair of Critical-Sized Mandibular Defects. ACS Biomater Sci Eng 2024;10(5):2863-2879.

[15] Yu B, Yin L, Wang P, et al. Terahertz reconfigurable multi-functional metamaterials based on 3D printed mortise-tenon structures. Virtual Phys Prototyp 2023;18:e2230468.

[16] Zhang B, Yin X, Zhang F, et al. Customized bioceramic scaffolds and metal meshes for challenging large-size mandibular bone defect regeneration and repair. Regen Biomater 2023;10:rbad057.

[17] Hemmy DC, David DJ, Herman GT. Three-dimensional reconstruction of craniofacial deformity using computed tomography. Neurosurgery 1983;13(5):534-541.

[18] Bagheri MA, Aubin CE, Nault ML, et al. Finite element analysis of distraction osteogenesis with a new extramedullary internal distractor. Comput Methods Biomech Biomed Engin 2024;1-15.

[19] Karakawa R, Yano T, Yoshimatsu H, et al. Computer-aided Design and Syringe-aided Manufacturing for Mandibular Reconstruction Using a Vascularized Fibula Flap. Plast Reconstr Surg Glob Open. 2020;8(5):e2819.

[20] Xue J, Qiu JZ, Yang SB, et al. Application of three-dimensional reconstruction and 3D printing in precise repair of mandibular defects with vascularized fibula grafts. J Precis Med 2018;33(1):45-50.

[21] Yu D, Huang JY, Yu CY, et al. Clinical application and accuracy analysis of modified mandibular functional reconstruction with occlusal guidance. Chin J Reconstr Surg 2020;34(11):1410-1416.

[22] Liang Y, Jiang C, Wu L, et al. Application of Combined Osteotomy and Reconstruction Pre-Bent Plate Position (CORPPP) Technology to Assist in the Precise Reconstruction of Segmental Mandibular Defects. J Oral Maxillofac Surg 2017;75(9):2026.e1-2026.e10.

[23] Peng X, Zhang WB. Application of digital surgical technology in the reconstruction of mandibular defects. Oral Dis Prev 2017;25(9):545-553.

[24] Ciocca L, Fantini M, Marchetti C, et al.

Immediate facial rehabilitation in cancer patients using CAD-CAM and rapid prototyping technology: a pilot study. Support Care Cancer 2010;18(6):723-728.
[25] Yan GQ, Wang X, Tan XX, et al. Application of SurgiCase software in guiding vascularized fibula flap reconstruction of mandibular defects. Chin J Reconstr Surg 2013;27(8):1006-1009.
[26] Kang JF, Zhang J, Zheng JB, et al. 3D-printed PEEK implant for mandibular defects repair-a new method. J Mech Behav Biomed Mater 2021;116:104335.
[27] McCarthy JG, Schreiber J, Karp N, et al. Lengthening the human mandible by gradual distraction. Plast Reconstr Surg 1992;89(1):1-8; discussion 9-10.
[28] Shen GF. Reconstruction of mandibular condyle defects using distraction osteogenesis. Chin J Pract Stomatol 2011;4(5):257-259.
[29] Su CL, Zhu SS, Li YF. Distraction osteogenesis with a transport disk to repair mandibular segmental defects. Chin J Tissue Eng 2021;25(35):5604-5609.
[30] Huang ZS, Liao JK, Chen WL, et al. Reconstruction of acquired segmental mandibular defects using pedicled mandibular muscle flap and evaluation of speech function and aesthetic outcomes. J Craniofac Surg 2023;34(2):494-497.

Progress in Medical Devices
ISSN: 2957-5478

Zentime

Review Article https://doi.org/10.61189/072185gbtgzi

Application of deep learning in the diagnosis of gastrointestinal diseases

Liying Pang[1], Xudong Guo[1], Qin Zhang[2]

[1]School of Health Science and Engineering, University of Shanghai for Science and Technology, Shanghai 200093, China. [2]Medical Engineering Department of Northern Jiangsu People's Hospital, Yangzhou 225001, Jiangsu Province, China.

Corresponding authors: Xudong Guo and Qin Zhang.

Declaration of conflict of interest: None.

Received November 27, 2024; Accepted January 15, 2025; Published June 30, 2025

Highlights

- This review presents the current status and challenges of gastrointestinal diseases.
- This review discusses the application of deep learning in diagnosing gastrointestinal diseases.
- This review explores the future development of deep learning in disease diagnosis.

Abstract

With the rapid development of artificial intelligence, deep learning technology has been widely applied across various fields. In the medical field, deep learning models, by analyzing medical images and clinical data, can automatically detect features of different types of lesions, such as polyps, ulcers, and cancers, thereby assisting physicians in early diagnosis of disease. This review provides an overview of recent progress in applying deep learning for disease diagnosis in various parts of the gastrointestinal tract, including the esophagus, stomach, small intestine, and colon. It also discusses the challenges and potential future directions for deep learning in this field.

Keywords: Deep learning, gastrointestinal endoscopy, convolutional neural network, computer-aided detection

Introduction

Gastrointestinal (GI) diseases are prevalent worldwide and have been steadily increasing due to lifestyle changes and dietary shifts in modern society, significantly compromising the quality of life of patients. In 2018, GI cancer patients in China accounted for 40% of the global total, with this trend continuing to rise [1]. In 2019, the global incidence of GI diseases reached 7.32 billion, resulting in 8 million deaths, and GI diseases accounted for over one-third of all epidemiological diseases [2].

Common GI diseases include gastroenteritis, ulcerative colitis, Crohn's disease (CD), gastric cancer, and esophageal cancer. Early diagnosis, precise treatment, and prognostic evaluation of these diseases remain key areas of medical research. However, the diagnostic process is complex due to the diverse symptoms, often requiring a combination of clinical manifestations, laboratory tests, endoscopic examinations, and imaging studies. Traditional diagnostic methods rely on the expertise of physicians, and the accuracy and efficiency of diagnosis are limited by manpower and time, highlighting the urgent need for automated and intelligent diagnostic tools.

In recent years, artificial intelligence (AI), especially deep learning (DL) technology, has shown great potential in medical image analysis and disease prediction. DL, a type of multi-layer artificial neural network model introduced in 2006, simulates the neural network of the human brain. Unlike traditional machine learning, DL can automatically extract features from data by building multi-layer nonlinear network structures, without relying on manually designed features (**Figure 1**) [3]. This method is well-suited for large-scale data processing and complex

Address correspondence to: Xudong Guo, School of Health Science and Engineering, University of Shanghai for Science and Technology, No.516 Jungong Road, YangPu District, Shanghai 200093, China. Tel: +86-021-55271115; E-mail: guoxd@usst.edu.cn. **Qin Zhang**, Medical Engineering Department of Northern Jiangsu People's Hospital, No.98 West Nantong Road, Yangzhou 225001, Jiangsu Province, China. Tel: +86-0514-87373012; E-mail: yzzqin@qq.com.

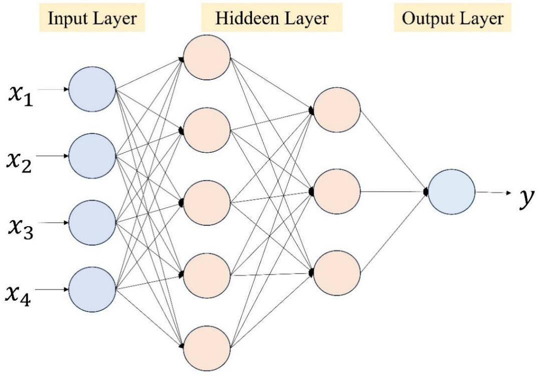


Figure 1. Schematic of deep learning neural network.

pattern recognition. Currently, DL is widely applied in speech recognition, natural language processing, semantic segmentation, and object detection [4]. In medical imaging, DL, such as convolutional neural networks (CNN) and recurrent neural networks, can efficiently process multi-dimensional data, like images, audio, and video, to extract subtle features for diagnostic interpretation. This aids doctors in making diagnostic decisions, improves diagnostic efficiency, reduces workload, and decreases the likelihood of missed or incorrect diagnoses [5]. Compared to traditional machine learning, DL methods are generally more accurate and robust, making them increasingly suitable for clinical practice [6].

This article reviews recent progress in DL technology for the diagnosis of GI diseases and discusses the challenges and future directions for intelligent-assisted diagnosis using deep learning technology.

Application of DL in the diagnosis of GI diseases

Esophageal carcinoma (ESCA)

ESCA is the eighth most common malignant tumor globally and the sixth leading cause of cancer-related death, with approximately 400,000 deaths annually [1]. Early symptoms of ESCA are often subtle, and by the time patients experience significant symptoms, such as dysphagia, weight loss, or other GI issues, the tumor is typically at an advanced stage, leading to a poor prognosis. According to statistics, the 5-year relative survival rate for ESCA from 2019 to 2021 is approximately 27.9% [7]. Early diagnosis of ESCA is crucial for improving patient survival rates. The main types of esophageal malignancies include esophageal adenocarcinoma (EAC) and esophageal squamous cell carcinoma (ESCC), which differ significantly in terms of epidemiology, pathogenesis, and prognosis.

EAC

EAC is one of the main subtypes of ESCA and predominantly occurs in North America and Europe [8]. Barrett's esophagus (BE) is a precursor lesion for the development of EAC, and early detection can effectively prevent BE from progressing to EAC. In recent years, researchers have increasingly utilized deep learning models to automate the detection and diagnosis of early-stage EAC and BE to improve the accuracy and efficiency of early screening.

The study by Ghatwary et al. in 2019 marked an important attempt at using deep learning for EAC recognition [9]. They trained various object detection models based on the VGG16 backbone network, including Regional-based CNN (R-CNN), Fast R-CNN, Faster R-CNN, and Single-Shot Multibox Detector (SSD), using 100 high-definition white light endoscopy images of the esophagus from 22 patients diagnosed with EAC and 17 patients with non-cancerous BE. The results demonstrated that the SSD and Faster R-CNN models performed well in detection, with the SSD model showing a sensitivity of 0.96, specificity of 0.92, and an F-measure of 0.94. The Faster R-CNN model, on the other hand, excelled in localizing the EAC region, achieving a recall rate of 0.83. Ebigbo et al. were the first to apply a DL system for real-time evaluation of early-stage EAC, further advancing the application of DL in EAC detection [10]. They used a total of 129 endoscopic images from the University Hospital Augsburg image database to train a real-time AI system based on deep CNNs and a residual network (ResNet) architecture with DeepLab V.3+, a state-of-the-art encoder–decoder network. The system achieved a sensitivity of 83.7%, specificity of 100%, and an accuracy of 89.9%.

de Groof et al. focused on distinguishing between images containing neoplasms or nondysplastic BE, developing a computer-aided diagnosis (CAD) system with a hybrid ResNet-UNet model [11]. Trained and tested on 1,704 early tumor images, the system showed excellent classification accuracy, sensitivity, and specificity (89%, 90% and 88%, respectively), accurately classifying images as containing neoplasms or nondysplastic BE. In addition, this system outperformed non-expert endoscopists, with comparable delineation performance. Hashimoto et al. combined the Xception-based CNN with the YOLO v2 algorithm for binary classification and detection of neoplasia images (**Figure 2**),

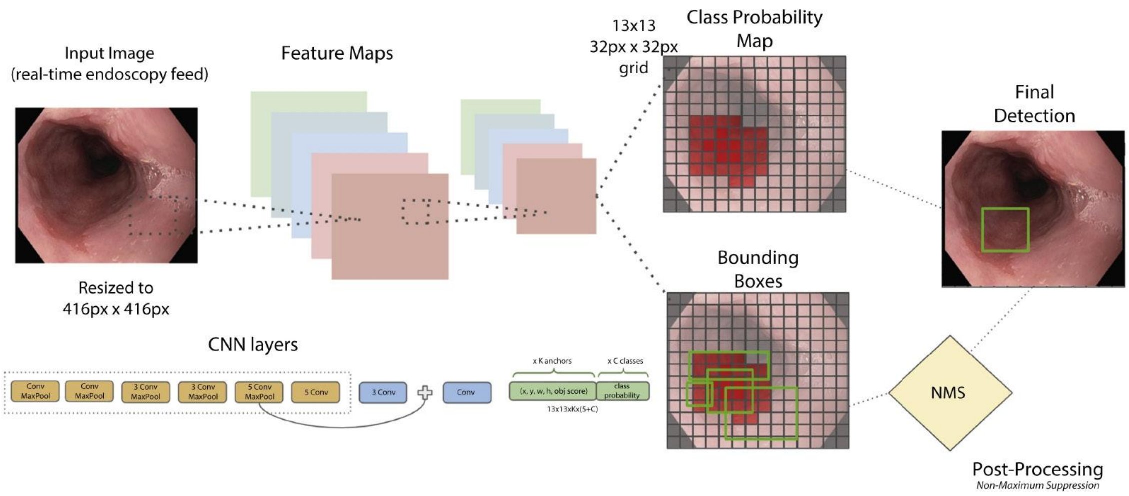


Figure 2. Architecture of the method proposed by Hashimoto et al. CNN, convolutional neural network; NMS, non-maximum suppression. This figure is cited from [12].

achieving a binary classification sensitivity of 96.4%, specificity of 94.2%, and accuracy of 95.4%, along with a mean average precision of 0.7533 [12]. Hussein et al. used the ResNet101 network to classify dysplastic and non-dysplastic BE images, achieving a mean area under the receiver operating characteristic curve (AUC) of 0.93, with sensitivity at 91% and specificity at 79% [13]. Additionally, they trained a model with the FCNResNet50 architecture for detecting dysplastic regions and predicting targeted biopsies, achieving a sensitivity of 97%. These studies demonstrate the exceptional performance of deep learning-based AI systems in diagnosing early EAC and BE, providing valuable auxiliary tools for clinical practice.

ESCC

ESCC is a common malignancy in regions such as Asia, Africa, and South America, with particularly high prevalence in China, where it accounts for over 90% of all esophageal cancer cases [8, 14]. As ESCC often presents with few symptoms in its early stages, many patients have progressed to the advanced stage at the time of diagnosis, leading to poor treatment outcomes. Therefore, early detection is crucial to improving treatment success and reducing ESCC mortality.

Cai et al. collected 2,428 esophagoscopic images from 746 patients during standard white light endoscopy, including 1,332 abnormal samples and 1,096 normal samples [15]. They developed a CAD system using deep neural networks (DNN) to localize and identify early-stage ESCC in routine white light imaging (WLI). Validation on 187 images resulted in a sensitivity of 97.8%, specificity of 85.4%, accuracy of 91.4%, positive predictive value (PPV) of 86.4%, and negative predictive value of 97.6%. With significant advances in optical technologies, tools such as narrow-band imaging (NBI) have further enhanced ESCC detection rate. However, the effectiveness of NBI in early ESCC screening largely depends on the operator's expertise. To address this, Everson et al. constructed an AI system based on CNN, using NBI images from magnified endoscopy to classify early squamous cell neoplasia (ESCN) and normal images under the abnormal intrapapillary capillary loop pattern [16]. This system achieved an accuracy of 93.7%, with sensitivity and specificity for abnormal intrapapillary capillary loop patterns reaching 89.3% and 98%, respectively. Additionally, Guo et al. used 6,473 NBI images, including those of precancerous lesions, early ESCC, and noncancerous lesions, to develop a highly efficient CAD system for intelligent diagnosis of precancerous lesions and ESCC [17]. This system achieved an impressive sensitivity of 98.04%, specificity of 95.03%, and an AUC value of 0.989, greatly enhancing the diagnostic capabilities for esophageal cancer. Li et al. also built a CAD-NBI system based on NBI images and CNN for non-magnified endoscopy to identify early ESCC [18]. The system performed excellently, with sensitivity, specificity, accuracy, positive predictive value, and negative predictive value of 91.0%, 96.7%, 94.3%, 95.3%, and 93.6%, respectively. Compared to the previously reported CAD-WLI system, this CAD-NBI system demonstrated superior specificity and accuracy [15].

In addition to the studies mentioned above, several research teams have applied deep learning models for intelligent recognition of esophageal and non-esophageal cancer images [19-22]. Horie et al. initially applied deep learning to the diagnosis of esophageal

Table 1. Research progress of deep learning in esophageal cancer

Ref.	Model	Lesions	Sensitivity (%)	Specificity (%)	Accuracy (%)	AUC
Ghatwary et al. [9]	Multiple CNNs	EAC	96.00	92.00	-	-
Ebigbo et al. [10]	ResNet	EAC	83.70	100.00	89.90	-
de Groof et al. [11]	ResNet-UNet	Neoplasia	90.00	88.00	89.00	-
Hashimoto et al. [12]	CNN-YOLOv2	Neoplasia	96.40	94.20	95.40	-
Hussein et al. [13]	ResNet101	Dysplasia	91.00	79.00	-	0.93
Cai et al. [15]	DNN	ESCC	97.80	85.40	91.40	-
Everson et al. [16]	CNN	ESCN	89.30	98.00	93.70	-
Guo et al. [17]	SegNet	ESCC	98.04	95.03	-	0.99
Li et al. [18]	FCN	ESCC	91.00	96.70	94.30	-
Horie et al. [19]	SSD	ESCA	98.00	-	-	-
Sui et al. [20]	VB-Net	ESCA	88.80	90.90	-	-
Takeuchi et al. [21]	VGG16	ESCA	71.70	90.00	84.20	-
Yasaka et al. [22]	CNN	ESCA	87.00	92.00	92.00	0.95

Note: FCN, fully convolutional network; EAC, esophageal adenocarcinoma; ESCC, esophageal squamous cell carcinoma; ESCA, esophageal carcinoma; DNN, deep neural networks; CNN, convolutional neural networks; SSD, Single-Shot Multibox Detector; ESCN, early squamous cell neoplasia.

cancer, training on 8,428 static images using the SSD network architecture within the Caffe deep learning framework [19]. They validated the model on independent test images from 97 patients, achieving a sensitivity of 98%, further demonstrating the powerful potential of deep learning for early diagnosis of ESCC. An overview of these studies is provided in **Table 1**.

Gastric diseases

Gastric diseases, particularly gastric cancer and gastric ulcers, are common GI tract diseases with high mortality. Gastric cancer ranks the fifth most common cancer worldwide and the fourth leading cause of cancer-related deaths. Due to its tendency to be diagnosed at advanced stages, gastric cancer has a very high mortality rate, with a 5-year survival rate of less than 40% [23, 24]. In 2018, approximately 780,000 people worldwide died from gastric cancer [1]. Traditional diagnostic methods, such as gastroscopy, CT scans, and magnetic resonance imaging, rely on the visual judgment and experience of doctors. While effective, these methods still have limitations, such as image noise interference, misdiagnosis, and errors due to fatigue. Recently, the introduction of DL has provided new solutions for the early diagnosis of gastric diseases. Researchers have developed various deep learning models, applying them to medical imaging data, such as endoscopic images, to enhance diagnostic accuracy.

Since 2018, a research team in Japan has conducted several studies on using AI for gastric cancer detection [25-28]. Hirasawa et al. [25] were the first to evaluate the ability of CNNs in detecting gastric cancer in endoscopic images. They used 13,584 endoscopic images of gastric cancer to train an SSD model and collected 2,296 independent images. In this study, the CNN system detected 92.2% of gastric cancer cases in the test set. Subsequently, the research team conducted a pilot study applying the CNN system to video images, which showed an accuracy of 94.1% [26]. Ikenoyama et al. compared the performance of this CNN system in diagnosing early gastric cancer with that of experienced endoscopists [27]. The results showed that the CNN had significantly higher sensitivity than the experienced endoscopists, with faster diagnostic speed. However, this CNN had lower PPV and specificity compared to endoscopists. This indicates that while CNNs may reduce the risk of missed cancer diagnoses, they may also increase the number of biopsies for non-cancerous lesions. In 2020, the team added 4,453 gastric ulcer images to their dataset and developed an "advanced CNN" (A-CNN) [28]. The A-CNN achieved a sensitivity of 99.0%, specificity of 93.3%, and PPV of 92.5% for gastric cancer classification, while the sensitivity, specificity, and PPV for gastric ulcer classification were 93.3%, 99.0%, and 99.1%, respectively. These results demonstrated that the new system could effectively classify gastric cancer and gastric ulcers.

Additionally, other research teams have also explored intelligent detection for gastric cancer. Noda et al. trained and validated the ResNet50 network using 906 cancer images and 717 non-cancer images from endocytoscopy (ECS), and the CNN showed a specificity of 90.9%, outperforming endoscopists in diagnosing early gastric cancer (EGC) [29]. Jin et al. used a mask-based convolutional neural network (Mask R-CNN, **Figure 3**) to automatically detect EGC in endoscopic images, with excellent performance [30]. Zhang et al. proposed an

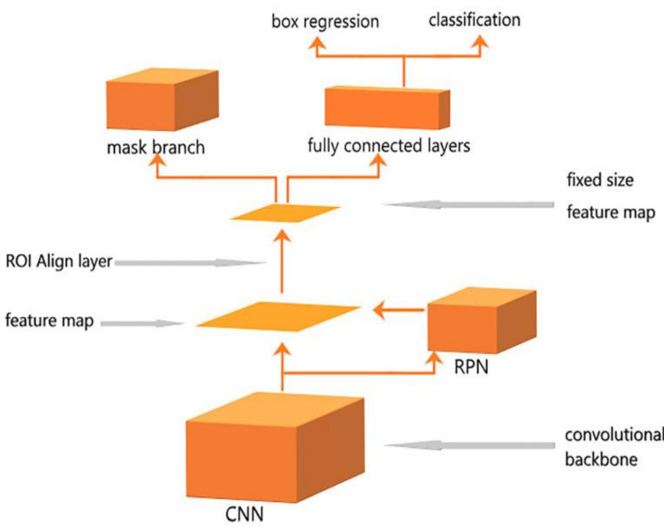


Figure 3. Mask region-based CNN architecture developed by Jin et al. CNN, convolutional neural networks; RPN, region proposal network; ROI, region of interest. This figure is cited from [30].

improved version of Mask R-CNN (IMR-CNN) to enhance the detection of EGC and segmental lesions in gastric endoscopic images [31]. Their model outperformed the original Mask R-CNN, significantly assisting doctors in diagnosing EGC from endoscopic images.

In addition to gastric cancer, DL techniques have also been widely applied in the recognition of other gastric lesions such as gastric ulcers. Zhang et al. implemented intelligent diagnostic system for peptic ulcers (PU), early gastric cancer (EGC), high-grade intraepithelial neoplasia (HGIN), advanced gastric cancer (AGC), and submucosal tumors (SMTs) in gastric endoscopic images using the ResNet34 model and the DeepLabv3 structure [32]. Xia et al. proposed a novel automatic gastric lesion detection system, employing the ResNet34 and Faster-Region-based CNN networks to differentiate between erosions, polyps, ulcers, submucosal tumors, xanthomas, and normal images in magnetic-controlled capsule endoscopy (MCE) images [33]. The system achieved an accuracy of 77.1% and an AUC value of 0.84, demonstrating good performance in diagnosing focal gastric lesions.

Fu et al. developed a new multi-scale model, StoHisNet, based on Transformer and CNN, for multi-class classification tasks of gastric pathological images [34]. The model achieved an accuracy of 94.69%, an F1 score of 94.96%, a recall rate of 94.95%, and a precision rate of 94.97% on the SEED public competition dataset of gastric pathological images, showing strong generalization ability across different datasets. Bhardwaj et al. evaluated the accuracy of various CNN models in identifying gastric diseases using images from the Kvasir dataset and deep transfer learning models [35]. Their research demonstrated the effectiveness of deep transfer learning techniques for the early prediction and classification of gastric cancer. **Table 2** presents the research progress of deep learning in gastric diseases.

Intestinal diseases

Small Intestine

The small intestine, a complex and highly convoluted structure connecting the stomach and colon, is challenging to fully examine with conventional imaging techniques due to its unique anatomical location. However, with the advent of advanced technologies such as small bowel endoscopy and capsule endoscopy, the diagnostic rate of small intestinal diseases has significantly improved. Despite these advancements, the interpretation of images still largely depends on the experience of clinicians, and diagnostic delays or misdiagnoses may occur due to image similarities. Deep learning-based medical image analysis offers promising potential for intelligent, auxiliary diagnosis of small intestine diseases.

Fan et al. utilized a Caffe framework to train the AlexNet architecture (**Figure 4**) on GPU and fine-tuned it to successfully distinguish small intestinal ulcers, erosions, and normal tissues in wireless capsule endoscopy (CE) images [36]. Their experimental results demonstrated high accuracy, sensitivity, and specificity in identifying ulcers and erosions, with AUC values exceeding 0.98. In subsequent studies, Aoki et al. applied the SSD architecture to detect small intestinal erosions and ulcers in CE images, achieving an accuracy of 90.8% and a specificity of 90.9%, further validating the effectiveness of DL in detecting small bowel disorders [37]. Ding et al. developed a CNN-based algorithm that efficiently differentiated abnormal from normal images in small bowel capsule endoscopy examinations [38]. Their model demonstrated a higher sensitivity (99.88% vs. 74.57%) and shorter reading time compared to gastroenterologists.

In the detection of specific lesions in the small intestine, Tsuboi et al. trained and validated an SSD-based deep CNN system with 2,237 CE images of vascular dilation [39]. The system achieved automatic detection of vasodilatation

Table 2. Research progress of deep learning in gastric diseases

Ref.	Model	Lesions	Sensitivity (%)	Specificity (%)	Accuracy (%)	AUC
Hirasawa et al. [25]	SSD	Gastric cancer	92.20	-	-	-
Namikawa et al. [28]	A-CNN	Gastric cancer	99.00	93.30	-	-
		Gastric ulcers	93.30	99.00	-	
Noda et al. [29]	ResNet50	EGC	82.10	90.90	86.10	0.93
Jin et al. [30]	Mask R-CNN	EGC-WLI	91.06	89.01	90.25	0.94
		EGC-NBI	97.59	89.01	95.12	0.81
Zhang et al. [31]	IMR-CNN	EGC	95.30	92.50	93.90	-
	ResNet34	EGC, HGIN	-	91.20	-	-
Xia et al. [33]	Faster RCNN	Gastric focal lesions	96.20	76.20	77.10	0.84
Fu et al. [34]	StoHisNet	Gastric pathological image	94.95	95.03	94.69	-

Note: Gastric focal lesions include erosion, polyp, ulcer, submucosal tumor, xanthoma, normal mucosa, and invalid images; gastric pathological image includes normal tissue, tubular adenocarcinoma, mucinous adenocarcinoma, and papillary adenocarcinoma. EGC-WLI, early gastric cancer under the white light image; EGC-NBI, early gastric cancer under the narrow-band imaging; HGIN, high-grade intraepithelial neoplasia; CNN, convolutional neural networks; SSD, Single-Shot Multibox Detector.

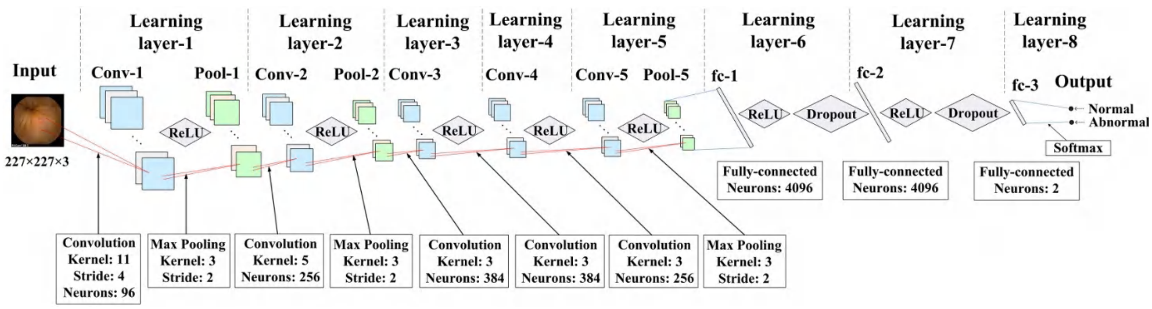


Figure 4. Illustration of CNN architecture diagram proposed by Fan et al. CNN, convolutional neural networks. This figure is cited from [36].

in CE images with sensitivities, specificities, PPV, and negative predictive value of 98.8%, 98.4%, 75.4%, and 99.9%, respectively. Klang et al. employed the Xception CNN to train on 17,640 CE images from 49 patients, achieving an accuracy ranging from 95.4% to 96.7% [40]. The model accurately distinguished between normal mucosa and mucosal ulcers, further demonstrating the practical applicability of CNN for automatic detection of mucosal ulcers in CD patients. Additionally, they tested the effectiveness of the EfficientNet-B5 network on 1,942 images of strictures, 14,266 images of normal mucosa, and 11,684 images of ulcers. The model achieved an average accuracy of 93.5% ± 6.7% in classifying strictures versus non-strictures, showcasing the powerful capabilities of DL in complex lesion classification [41]. Leenhardt et al. combined segmentation methods with CNN-based deep feature extraction to classify images of gastrointestinal angioectasia and normal images from the capsule endoscopy computer-assisted diagnosis for capsule endoscopy (CAD-CAP) database [42]. Their algorithm achieved a specificity of 96%, effectively detecting gastrointestinal angioectasia in static frames of small bowel capsule endoscopy and improving the diagnostic performance for small bowel diseases. Hosoe et al. used ERFNet to classify various lesions in small intestine capsule endoscopy images, including bleeding, vascular malformations, ulcers, and tumors [43]. Sensitivity and specificity were calculated from 271 detection results across 35 cases, with sensitivity at 93.4% and specificity at 97.8%.

Xie et al. applied the EfficientNet-B5 deep learning model for the automatic detection of CD lesions, including ulcers, non-inflammatory stenosis, and inflammatory stenosis, achieving high accuracy rates of 96.3%, 95.7%, and 96.7%, respectively [44]. Additionally, they performed grading of the ulcer surface, size, and depth, which also showed high accuracy, demonstrating the great potential of deep learning in both automatic detection and objective grading of small bowel CD lesions. Furthermore, Zhu et al. developed a real-time CAD system combining YOLO and ResNet50 for lesion detection and classification during double-balloon enteroscopy [45]. This system detected lesions such as

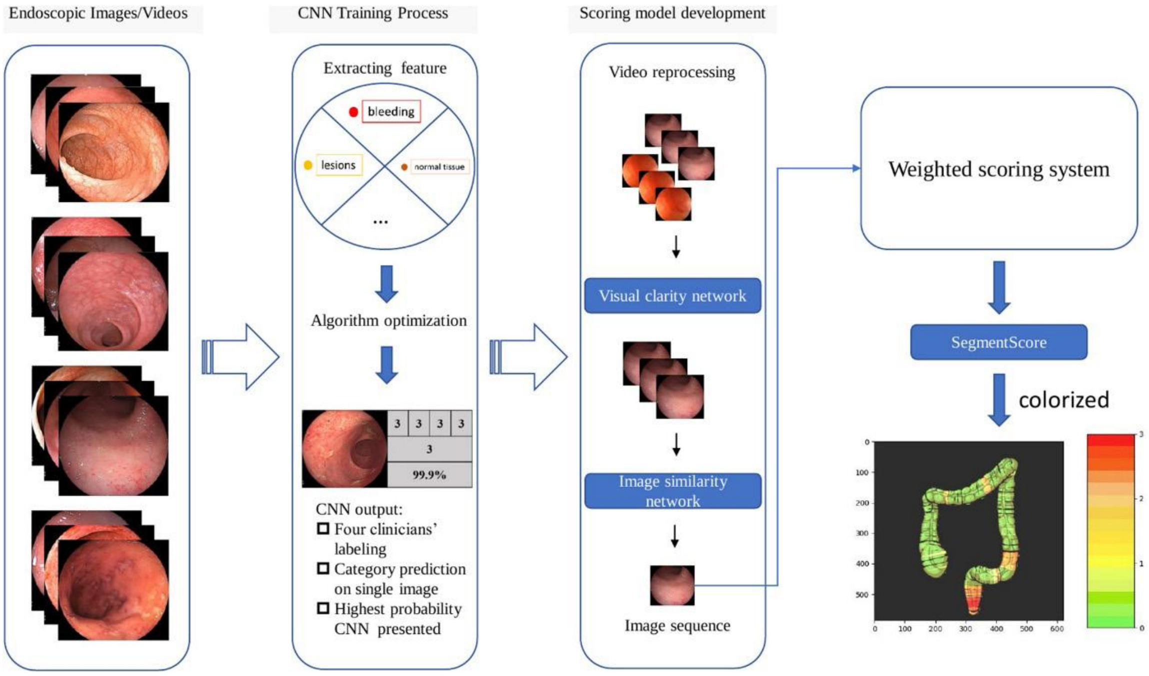


Figure 5. Workflow of CAD system designed by Fan et al. CAD, computer-aided diagnosis. This figure is cited from [51].

protrusions, diverticula, erosions, ulcers, and vascular dilation, achieving a sensitivity of 92% and an AUC of 0.947, while a classification accuracy of 86%, demonstrating excellent experimental performance.

Colon

The colon is a high-prevalence site for intestinal diseases, including colon cancer, colon polyps, and ulcerative colitis. Among these, colon cancer ranks as the third most common cancer worldwide and the second leading cause of cancer-related deaths [14]. The diagnostic accuracy of colonoscopy, a key tool in diagnosing colorectal diseases, is often limited by the experience of the physician and the quality of the images, resulting in variability in diagnostic results. Adenoma detection rate, a crucial metric for assessing colonoscopy quality, is closely associated with the risk of interval colon cancer. To enhance adenoma detection rate, the application of real-time automatic polyp detection systems has garnered significant attention.

Urban et al. trained a CNN model using 8,641 manually annotated colonoscopy images, achieving a polyp recognition accuracy of up to 96.4% and an AUC of 0.991, which effectively improves polyp detection and localization [46]. Wang et al. developed and validated a deep learning algorithm for high-precision detection and recognition of adenomatous polyps in colonoscopy images [47]. This algorithm not only demonstrated high sensitivity and specificity but also enabled real-time analysis of colonoscopy videos, helping to improve polyp detection performance by endoscopists and reduce variability in detection. In addition, Chen et al. analyzed narrow-band images of small colorectal polyps based on a DNN-CAD system, achieving a PPV of 89.6% and a negative predictive value of 91.5% on the test set, along with shorter detection times [48].

In terms of multi-class classification of colonoscopy images, Wang et al. proposed a deep learning-based method for classifying medical colonoscopy images into four categories: polyps, inflammation, tumors, and normal conditions [49]. By transferring learning from AlexNet, VGGNet, and ResNet networks, this method not only ensured classification accuracy but also improved polyp recognition rates. Additionally, Sarwinda et al. utilized ResNet-18 and ResNet-50 deep learning architectures for classifying colonic gland images, finding that ResNet-50 outperformed ResNet-18 in terms of accuracy, sensitivity, and specificity, further demonstrating the reliability and reproducibility of deep learning in biomedical image analysis [50].

Beyond that, Fan et al. developed a CAD system for scoring ulcerative colitis in various regions of intestinal endoscopy videos using deep learning techniques [51]. As shown in **Figure 5**, the system comprises a classification module, a pre-processing CNN model, and a weighted scoring system, capable of reliably completing

Table 3. Research progress of deep learning in intestinal diseases

Ref.	Model	Lesions	Sensitivity (%)	Specificity (%)	Accuracy (%)	AUC
Fan et al. [36]	AlexNet	Ulcer	96.80	94.79	95.16	0.989
		Erosion	93.67	95.98	95.34	0.986
Aoki et al. [37]	SSD	Ulcer and erosion	88.20	90.90	90.80	0.958
Ding et al. [38]	CNN	Multi-lesions	99.88	-	-	-
Tsuboi et al. [39]	SSD	Angioectasia	98.80	98.40	-	0.998
Klang et al. [40]	Xception CNN	CD	-	-	-	0.990
Klang et al. [41]	EfficientNet-B5	Stricture	-	97.20	93.50	0.962
Leenhardt et al. [42]	CNN	GIA	100.00	96.00	-	-
Hosoe et al. [43]	EFRNet	Bleeding, angiodysplasia, ulceration, and neoplastic lesions	93.40	97.80	-	
Xie et al. [44]	EfficientNet-B5	CD	94.80	96.80	96.30	0.989
		Noninflammatory stenosis	95.10	95.90	95.70	0.989
		Inflammatory stenosis	96.00	96.90	96.70	0.992
Zhu et al. [45]	ResNet50+ YOLO	Protruding lesion, diverticulum, erosion & ulcer and angioectasia	92.00	93.00	86.00	0.947
Urban et al. [46]	CNN	Colon polyp	-	-	96.40	0.991
Wang et al. [47]	CNN	Colon polyp	94.38	95.92	-	0.984
Chen et al. [48]	DNN	Colorectal polyps	96.30	78.10	-	-
Wang et al. [49]	Multiple CNNs	Polyps, inflammation, tumor, and normal	-	-	94.48	-
Sarwinda et al. [50]	ResNet	Colorectal cancer	87.00	83.00	80.00	-
Fan et al. [51]	ResNet50	UC	87.50	96.68	86.54	

Note: The multi lesions in literature include inflammation, ulcer, polyps, lymphangiectasia, bleeding, vascular disease, protruding lesion, lymphatic follicular hyperplasia, diverticulum, parasite, and other. CD, Crohn's disease; GIA, gastrointestinal angioectasia; UC, ulcerative colitis; CNN, convolutional neural networks; DNN, deep neural networks; SSD, Single-Shot Multibox Detector.

Mayo and Ulcerative Colitis Endoscopic Index of Severity visual classification tasks and automatically scoring inflammatory activity across full-length endoscopy videos. Experimental results indicate that the system provides accurate, efficient, and reproducible scores. These studies not only showcase the immense potential of DL in diagnosing colonic diseases but also provide robust support for future clinical applications. An overview of deep learning in intestinal diseases is provided in **Table 3**.

Conclusion

In recent years, deep learning, as an important branch in the field of artificial intelligence, has made significant strides in the medical field, particularly in the diagnosis and treatment of GI diseases. This review highlights the high prevalence of GI diseases and the importance of early diagnosis, followed by a brief introduction to the concept and applications of deep learning. It then explores the background of deep learning in the context of GI diseases, noting that its application provides new approaches for early screening and precise diagnosis. Deep learning-assisted diagnosis has shown remarkable results in improving diagnostic accuracy, reducing clinician workload, and increasing early detection rates for diseases such as esophageal cancer, gastric cancer, small bowel ulcers, and colon cancer.

Despite the significant progress made in applying deep learning for GI disease diagnosis, several challenges remain. Most current studies are retrospective and single-center, limiting their generalizability. Future research should focus on conducting prospective, multi-center studies to enhance the robustness and reliability of the models. Furthermore, deep learning models requires large, high-quality labeled datasets, which are often difficult to obtain in the medical field due to issues such as limited incidence and data sensitivity. Lastly, as deep learning-based intelligent diagnostic technologies continue to evolve, addressing issues related to their legal and compliant use in clinical practice, as well as safeguarding patient privacy and data security, will be crucial.

The future of deep learning in intelligent diagnostic assistance remain promising. As technology advances, deep learning models are expected to demonstrate improved generalization capabilities and higher accuracy. Additionally, research on multimodal learning will allow deep learning models to integrate data from various sources, providing more personalized healthcare services and comprehensive diagnostic support. As deep learning applications become more standardized and secure in clinical practice, technology is poised to play an increasingly important role in healthcare. It is believed that with the continuous advancements of technology, deep learning is expected to significantly impact the medical field and transform the landscape of disease diagnosis and treatment.

Author contributions: Liying Pang performed the data analyses and wrote the manuscript; Xudong Guo, Qin Zhang supervised the study, making significant contributions to the writing and revision of the paper.

References

[1] Freddie B, Jacques F, Isabelle S, et al. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2018;68(6):394-424.

[2] Wang Y, Huang Y, Chase RC, et al. Global Burden of Digestive Diseases: A Systematic Analysis of the Global Burden of Diseases Study, 1990 to 2019. Gastroenterology 2023;165(3):773-783.e715.

[3] Talaei Khoei T, Ould Slimane H, Kaabouch N. Deep learning: systematic review, models, challenges, and research directions. Neural Comput Appl 2023;35(31):23103-23124.

[4] Dong S, Wang P, Abbas K. A survey on deep learning and its applications. Comput Sci Rev 2021;40:100379.

[5] Cui M, Zhang DY. Artificial intelligence and computational pathology. Lab Invest 2021;101(4):412-422.

[6] Pang X, Zhao Z, Weng Y. The Role and Impact of Deep Learning Methods in Computer-Aided Diagnosis Using Gastrointest Endosc. Diagnostics 2021;11(4):694.

[7] Zeng H, Zheng R, Sun K, et al. Cancer survival statistics in China 2019–2021: a multicenter, population-based study. J Natl Cancer Cent 2024;4(3):203-213.

[8] Voron T, Julio C, Pardo E. Cancers œsophagiens: nouveautés et défis des prises en charge chirurgicales. Bulletin du Cancer 2023;110(5):533-539.

[9] Ghatwary N, Zolgharni M, Ye X. Early esophageal adenocarcinoma detection using deep learning methods. Int J Comput Assist Radiol Surg 2019;14(4):611-621.

[10] Ebigbo A, Mendel R, Probst A, et al. Real-time use of artificial intelligence in the evaluation of cancer in Barrett's oesophagus. Gut 2020;69(4):615-616.

[11] de Groof AJ, Struyvenberg MR, van der Putten J, et al. Deep-Learning System Detects Neoplasia in Patients With Barrett's Esophagus With Higher Accuracy Than Endoscopists in a Multistep Training and Validation Study With Benchmarking. Gastroenterology 2020;158(4):915-929.e914.

[12] Hashimoto R, Requa J, Dao T, et al. Artificial intelligence using convolutional neural networks for real-time detection of early esophageal neoplasia in Barrett's esophagus (with video). Gastrointest Endosc 2020;91(6):1264-1271.e1261.

[13] Hussein M, González-Bueno Puyal J, Lines D, et al. A new artificial intelligence system successfully detects and localises early neoplasia in Barrett's esophagus by using convolutional neural networks. United European Gastroenterol J 2022;10(6):528-537.

[14] Sung H, Ferlay J, Siegel RL, et al. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. CA Cancer J Clin 2021;71(3):209-249.

[15] Cai SL, Li B, Tan WM, et al. Using a deep learning system in endoscopy for screening of early esophageal squamous cell carcinoma (with video). Gastrointest Endosc 2019;90(5):745-753.e2.

[16] Everson M, Herrera L, Li W, et al. Artificial intelligence for the real-time classification of intrapapillary capillary loop patterns in the endoscopic diagnosis of early oesophageal squamous cell carcinoma: A proof-of-concept study. United European Gastroenterol J 2019;7(2):297-306.

[17] Guo L, Xiao X, Wu C, et al. Real-time automated diagnosis of precancerous lesions and early esophageal squamous cell carcinoma using a deep learning model (with videos). Gastrointest Endosc 2020;91(1):41-51.

[18] Li B, Cai SL, Tan WM, et al. Comparative study on artificial intelligence systems for detecting early esophageal squamous cell carcinoma between narrow-band and white-light imaging. World J Gastroenterol 2021;27(3):281-293.

[19] Horie Y, Yoshio T, Aoyama K, et al. Diagnostic outcomes of esophageal cancer by artificial intelligence using convolutional neural networks. Gastrointest Endosc 2019;89(1):25-32.

[20] Sui H, Ma R, Liu L, et al. Detection of

Incidental Esophageal Cancers on Chest CT by Deep Learning. Front Oncol 2021;11:700210.
[21] Takeuchi M, Seto T, Hashimoto M, et al. Performance of a deep learning-based identification system for esophageal cancer from CT images. Esophagus 2021;18(3):612-620.
[22] Yasaka K, Hatano S, Mizuki M, et al. Effects of deep learning on radiologists' and radiology residents' performance in identifying esophageal cancer on CT. Br J Radiol 2023;96(1150):20220685.
[23] Niu PH, Zhao LL, Wu HL, et al. Artificial intelligence in gastric cancer: Application and future perspectives. World J Gastroenterol 2020;26(36):5408-5419.
[24] Klang E, Sourosh A, Nadkarni GN, et al. Deep Learning and Gastric Cancer: Systematic Review of AI-Assisted Endoscopy. Diagnostics (Basel) 2023;13(24):3613.
[25] Hirasawa T, Aoyama K, Tanimoto T, et al. Application of artificial intelligence using a convolutional neural network for detecting gastric cancer in endoscopic images. Gastric Cancer 2018;21(4):653-660.
[26] Ishioka M, Hirasawa T, Tada T. Detecting gastric cancer from video images using convolutional neural networks. Dig Endosc 2019;31(2):e34-e35.
[27] Ikenoyama Y, Hirasawa T, Ishioka M, et al. Detecting early gastric cancer: Comparison between the diagnostic ability of convolutional neural networks and endoscopists. Dig Endosc 2021;33(1):141-150.
[28] Namikawa K, Hirasawa T, Nakano K, et al. Artificial intelligence-based diagnostic system classifying gastric cancers and ulcers: comparison between the original and newly developed systems. Endoscopy 2020;52(12):1077-1083.
[29] Noda H, Kaise M, Higuchi K, et al. Convolutional neural network-based system for endocytoscopic diagnosis of early gastric cancer. BMC Gastroenterol 2022;22(1):237.
[30] Jin J, Zhang Q, Dong B, et al. Automatic detection of early gastric cancer in endoscopy based on Mask region-based convolutional neural networks (Mask R-CNN)(with video). Front Oncol 2022;12:927868.
[31] Zhang K, Wang H, Cheng Y, et al. Early gastric cancer detection and lesion segmentation based on deep learning and gastroscopic images. Sci Rep 2024;14(1):7847.
[32] Zhang L, Zhang Y, Wang L, et al. Diagnosis of gastric lesions through a deep convolutional neural network. Dig Endosc 2021;33(5):788-796.
[33] Xia J, Xia T, Pan J, et al. Use of artificial intelligence for detection of gastric lesions by magnetically controlled capsule endoscopy. Gastrointest Endosc 2021;93(1):133-139.e134.
[34] Fu B, Zhang M, He J, et al. StoHisNet: A hybrid multi-classification model with CNN and Transformer for gastric pathology images. Comput Methods Programs Biomed 2022;221:106924.
[35] Bhardwaj P, Kim S, Koul A, et al. Advanced CNN models in gastric cancer diagnosis: enhancing endoscopic image analysis with deep transfer learning. Front Oncol 2024;14:1431912.
[36] Fan S, Xu L, Fan Y, et al. Computer-aided detection of small intestinal ulcer and erosion in wireless capsule endoscopy images. Phys Med Bio 2018;63(16):165001.
[37] Aoki T, Yamada A, Aoyama K, et al. Automatic detection of erosions and ulcerations in wireless capsule endoscopy images based on a deep convolutional neural network. Gastrointest Endosc 2019;89(2):357-363.e352.
[38] Ding Z, Shi H, Zhang H, et al. Gastroenterologist-Level Identification of Small-Bowel Diseases and Normal Variants by Capsule Endoscopy Using a Deep-Learning Model. Gastroenterology 2019;157(4):1044-1054.e1045.
[39] Tsuboi A, Oka S, Aoyama K, et al. Artificial intelligence using a convolutional neural network for automatic detection of small-bowel angioectasia in capsule endoscopy images. Dig Endosc 2020;32(3):382-390.
[40] Klang E, Barash Y, Margalit RY, et al. Deep learning algorithms for automated detection of Crohn's disease ulcers by video capsule endoscopy. Gastrointes Endosc 2020;91(3):606-613. e602.
[41] Klang E, Grinman A, Soffer S, et al. Automated Detection of Crohn's Disease Intestinal Strictures on Capsule Endoscopy Images Using Deep Neural Networks. J Crohns Colitis 2021;15(5):749-756.
[42] Leenhardt R, Vasseur P, Li C, et al. A neural network algorithm for detection of GI angiectasia during small-bowel capsule endoscopy. Gastrointest Endosc 2019;89(1):189-194.
[43] Hosoe N, Horie T, Tojo A, et al. Development of a Deep-Learning Algorithm for Small Bowel-Lesion Detection and a Study of the Improvement in the False-Positive Rate. J Clin Med 2022;11(13):3682.
[44] Xie W, Hu J, Liang P, et al. Deep learning-based lesion detection and severity grading of small-bowel Crohn's disease ulcers on double-balloon endoscopy images. Gastrointest Endosc 2024;99(5):767-777.e765.
[45] Zhu Y, Lyu X, Tao X, et al. A newly developed deep learning-based system for automatic

detection and classification of small bowel lesions during double-balloon enteroscopy examination. BMC Gastroenterol 2024;24(1):10.
[46] Urban G, Tripathi P, Alkayali T, et al. Deep Learning Localizes and Identifies Polyps in Real Time With 96% Accuracy in Screening Colonoscopy. Gastroenterology 2018;155(4):1069-1078.e1068.
[47] Wang P, Xiao X, Glissen Brown JR, et al. Development and validation of a deep-learning algorithm for the detection of polyps during colonoscopy. Nat Biomed Eng 2018;2(10):741-748.
[48] Chen PJ, Lin MC, Lai MJ, et al. Accurate Classification of Diminutive Colorectal Polyps Using Computer-Aided Analysis. Gastroenterology 2018;154(3):568-575.
[49] Wang Y, Feng Z, Song L, et al. Multiclassification of Endoscopic Colonoscopy Images Based on Deep Transfer Learning. Comput Math Methods Med 2021;2021:2485934.
[50] Sarwinda D, Paradisa RH, Bustamam A, et al. Deep Learning in Image Classification using Residual Network (ResNet) Variants for Detection of Colorectal Cancer. Proc Comput Sci 2021;179:423-431.
[51] Fan Y, Mu R, Xu H, et al. Novel deep learning–based computer-aided diagnosis system for predicting inflammatory activity in ulcerative colitis. Gastrointest Endosc 2022;97(2):335-346.

Progress in Medical Devices
ISSN: 2957-5478

Zentime

Review Article https://doi.org/10.61189/606109gzqkrg

Research progress on simulation of radiofrequency ablation for liver cancer treatment

Shaobo Wang[1], Yuan Yao[2], Haipo Cui[1]

[1]Shanghai Institute for Minimally Invasive Therapy, University of Shanghai for Science and Technology, Shanghai 200093, China. [2]Shanghai Songyu Medical Device Co., Ltd., Shanghai 200050, China.

Corresponding author: Haipo Cui.

Acknowledgement: This work was supported by State Key Laboratory of Systems Medicine for Cancer Foundation (KF2407-93).

Declaration of conflict of interest: None.

Received April 26, 2025; Accepted June 9, 2025; Published June 30, 2025

Highlights

- Analyzing the influence of biological tissue parameters on the effectiveness of radiofrequency ablation.
- Comparing the advantages and limitations of four biological heat transfer models.
- Comparing three evaluation methods for assessing ablation-induced thermal damage.

Abstract

Radiofrequency ablation (RFA) is a widely used minimally invasive treatment for non-surgical hepatocellular carcinoma. This review synthesizes the technical principles, core components, and key modeling aspects of RFA. RFA induces tumor necrosis via Joule heating from ionic vibration. Electrode needle design critically impacts ablation efficacy and safety, with multipolar needles offering larger zones yet posing power and tissue risks. Crucially, biological tissue parameters exhibit dynamic spatial and thermal variations, necessitating nonlinear modeling for accurate temperature prediction. While the Pennes bioheat model remains mainstream for its simplicity, more advanced models (e.g., porous medium) enhance physiological realism. Thermal damage assessment commonly employs the Arrhenius model and isothermal thresholds, aided by real-time monitoring for intraoperative precision. Future research should prioritize the development of smart electrodes, creation of personalized tissue parameter databases, and exploration of multi-energy techniques to shift RFA from an “empirically oriented” approach to an “accurate prediction” paradigm, ultimately improving hepatocellular carcinoma patient survival and quality of life.

Keywords: Radiofrequency ablation, liver cancer, electron spin, biological heat transfer equation

Introduction

Cancer is a major disease threatening human life and health, characterized by high malignancy and rapid growth. Among cancer types, primary liver cancer is one of the most common malignant tumors, primarily classified into hepatocellular carcinoma and intrahepatic cholangiocarcinoma, with hepatocellular carcinoma accounting for 70%-80% of the cases [1-4]. The main treatment methods for hepatic tumors include surgical resection, radiotherapy, chemotherapy, and thermal ablation, with resection, radiotherapy and chemotherapy still being the preferred choices in clinical practice [5]. Although surgery can directly resect the tumors, it presents high risks in cases involving tumors in challenging locations, large tumor size, or patients in poor physical condition. In such cases, complete tumor resection may not be achievable. Radiotherapy, while effective for some tumors, can cause damage to surrounding normal tissues and may fail to completely eradicate the tumor [6]. With the continuous development of minimally invasive therapies, local thermal ablation techniques have become an effective treatment for liver cancer [7, 8].

Address correspondence to: Haipo Cui, Shanghai Institute for Minimally Invasive Therapy, University of Shanghai for Science and Technology, Shanghai 200093 Jungong Road, China. E-mail: h_b_cui@163.com.

Local ablation techniques mainly include microwave ablation, laser ablation, and radiofrequency ablation (RFA) [9-11]. RFA, first used in the treatment of liver tumors in the 1990s, is a thermal coagulation technique. It offers the advantages of minimal invasion, high safety, effectiveness, simple operation, and high reproducibility, making it the first-line non-surgical treatment for hepatocellular carcinoma [12-14].

However, efficacy of RFA is highly dependent on the precise prediction of thermal distribution and ablation zones [13]. The complex biophysical interactions during RFA, including dynamic variations in tissue parameters (e.g., electrical/thermal conductivity, blood perfusion), heat-sink effects near blood vessels, and nonlinear thermal damage thresholds, pose significant challenges for treatment planning. Computational simulation has thus become indispensable for optimizing electrode design and ablation protocols, predicting temperature distribution and coagulation necrosis boundaries, and personalizing treatments through virtual patient-specific modeling. This review provides a comprehensive overview of advancements in RFA simulation methodologies, integrating experimental insights with computational frameworks to advance precision oncology.

Radiofrequency refers to electromagnetic waves with frequencies between audio and infrared ranges, spanning from approximately 104 to 3×1012 Hz. RFA involves the use of unmodulated sinusoidal alternating current (AC) within this frequency range to destroy biological tissues, with a typical frequency of about 500 KHz. This frequency falls in the mid-wave range, high enough (>20 KHz) to induce molecular frictional heating without triggering neuromuscular responses or electrolysis, and low enough (<20 MHz) to limit the energy transfer to a more controllable tissue mass without excessive radiation. It has been shown that alternating currents exceeding 100 mA can cause fatal electric shocks and ventricular fibrillation at the lower frequencies, such as 50 Hz (the frequency of household electricity) [13]. Unlike higher frequency ionizing electromagnetic waves (e.g., typical X-rays at 1018 Hz), radiofrequency energy is non-ionizing and is considered safe for use when applied properly. In the context of RFA, “ablation” can be understood as a virtual surgical ablation, i.e., the inactivation of tissue without removing it from the body, thus producing results similar to those of surgical resection [14, 15].

A typical RFA system consists of a radiofrequency generator, electrode needles, and a grounding plate, which form a complete current circuit with the human liver during the procedure. During RFA, the physician first uses medical imaging to precisely locate the liver tumor. Then, various types of electrode needles (unipolar or multipolar) are inserted into the tumor area [16]. The radiofrequency generator delivers high-frequency current (100-500 KHz) to the cancerous tissue, causing high-frequency vibration and friction of conductive ions and polarized molecules in the tissues around the electrode needles, thus generating Joule heat [17]. Over time, this heat energy is transmitted to the surrounding tissues, causing a coagulation reaction zone around the tumor. As a result, the tumor’s blood supply is cut off, leading to coagulative necrosis of the tumor tissue. The necrotic tissue then peels off and is either absorbed by the body or discharged, achieving tumor destruction [18].

The potential of RFA for liver cancer treatment was first proposed by Rossi et al., with a related study published in 1993 [19, 20]. The effectiveness of RFA in treating liver tumors lies in its ability to exploit the poor heat resistance of tumor cells. Energy deposition into the tumor triggers thermal damage, which produces a tumor-killing effect. RFA involves the flow of alternating current through tissue, causing ionic agitation and impedance heating of the tissue. To establish this current, RFA systems require a closed-loop circuit comprising an electrical energy generator, a needle electrode, the patient (acting as a resistor), and a large-area electrodes (or “grounding pads”). It has been shown that when tissue temperature exceeds 40°C, tumor cells stop dividing; at temperatures above 50°C, tumor cell proteins denature; at temperatures above 60°C, tumor cells coagulate and necrose; and at temperatures reaching 100°C, tissues are charred [21].

Ablation is a very complex procedure, and the success of current tumor RFA treatments depends on several factors, including the structure and materials of the radiofrequency electrode needle, radiofrequency input power, blood perfusion rate, and the conductivity of the liver tissue [22]. **Figure 1** shows the process of RFA for a liver lesion [23].

With the continuous advancement of RFA technology for liver cancer treatment, the electrode needles, as key components, have also undergone significant improvements in clinical practice [24]. Currently, commonly used RFA electrode needles can be categorized based on function and structure. Functionally, they include multi-needle expandable electrode

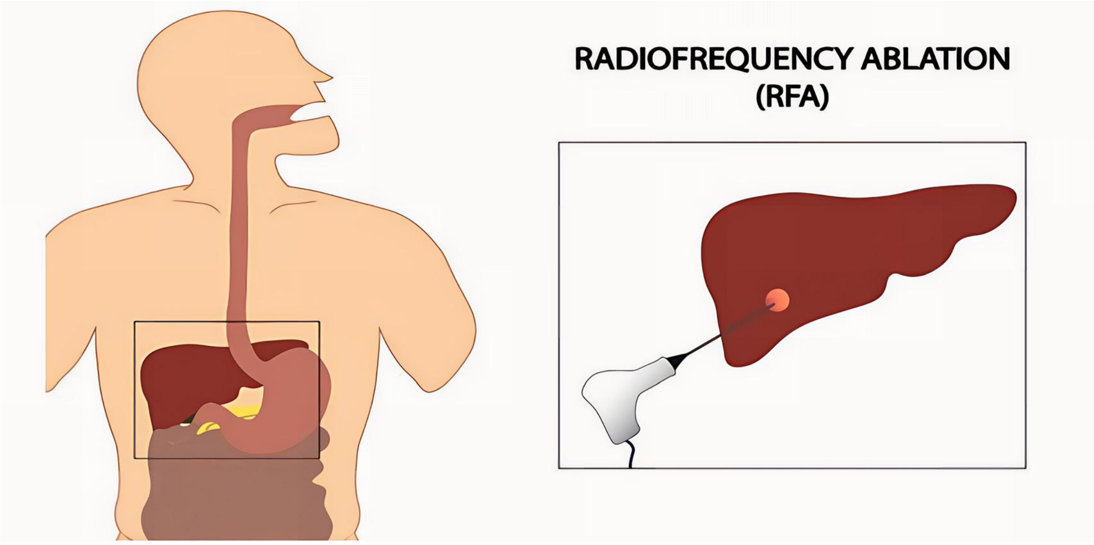


Figure 1. Schematic diagram of radiofrequency ablation. This figure is cited from [23].

needles, internal cooling electrode needles, and perfusion electrode needles. Structurally, they can be classified into unipolar, bipolar, and multipolar needles, depending on the number of needles and subneedles. Among these, multipolar electrode needles can deploy multiple subneedles (typically three or more) at the tip of the needle tube, which are fully expanded in the form of an anchor or umbrella with a uniform spatial structure [25]. A study simulating the performance of spherical and cylindrical electrode tips in cardiac RFA, using a three-dimensional computational model, revealed that the shape of the electrode tip significantly affects the extent of ablation injury, safety, and the risk of complications, providing crucial insights for optimizing catheter design and ablation protocols [26]. Radiofrequency catheter ablation for pulmonary vein isolation is usually achieved by linear point-to-point ablation. However, creating continuous lesions can be time-consuming and requires advanced 3D mapping systems. To overcome these limitations, a multi-electrode ablation system was designed.

Currently, the most commonly used multipolar electrode needles in clinical practice are the multi-needle tip electrode needles developed by MedSphere Technologies [27]. Internal cooling electrode needles can be categorized into liquid-cooled and air-cooled according to the cooling medium. These needles typically feature unipolar or bipolar structures, which reduces tissue temperature around the needle tip by circulating coolant or cooling gas through a built-in hollow cavity. A commonly used internal cooling electrode needle is the Valleylab monopolar needle from Tyco, which is liquid-cooled, with an ablation time of 12 min and an ablation range of up to 3.6 cm in the long axis [28, 29].

In actual RFA procedures, selecting the appropriate electrode needle is crucial for achieving optimal ablation effect, especially considering the size and shape of the tumors. Unipolar and bipolar needles have good therapeutic effects in ablating tumors with a diameter of less than 3 cm. However, due to the small diameter of their needle tubes, these needles often require multiple punctures. Additionally, the temperature distribution around the needle tip is not controllable, potentially causing burning and charring of surrounding tissue, which can cause secondary injuries and increases the risk of complications [30]. The umbrella structure of the multipolar needles allows for a larger ablation range, making them suitable for tumors with a diameter of 3-5 cm. However, their limited output power results in a longer ablation time. Furthermore, fully deploying the multipolar needle may damage adjacent normal tissues, and the tissues in the electrode needle's vicinity are easily burned and charred, limiting their clinical application [31]. **Figure 2** illustrates the schematic diagram of RFA treatment using a multipolar electrode needle [23].

Recent studies have demonstrated that stepwise incremental high-power multi-electrode RFA technology, combined with real-time image guidance, provides a highly effective and safe treatment option for small hepatocellular carcinoma, significantly reducing the rate of local recurrence [32]. However, further validate through multicenter randomized controlled trials is necessary to confirm its long-term efficacy and explore its applicability in a broader patient population.

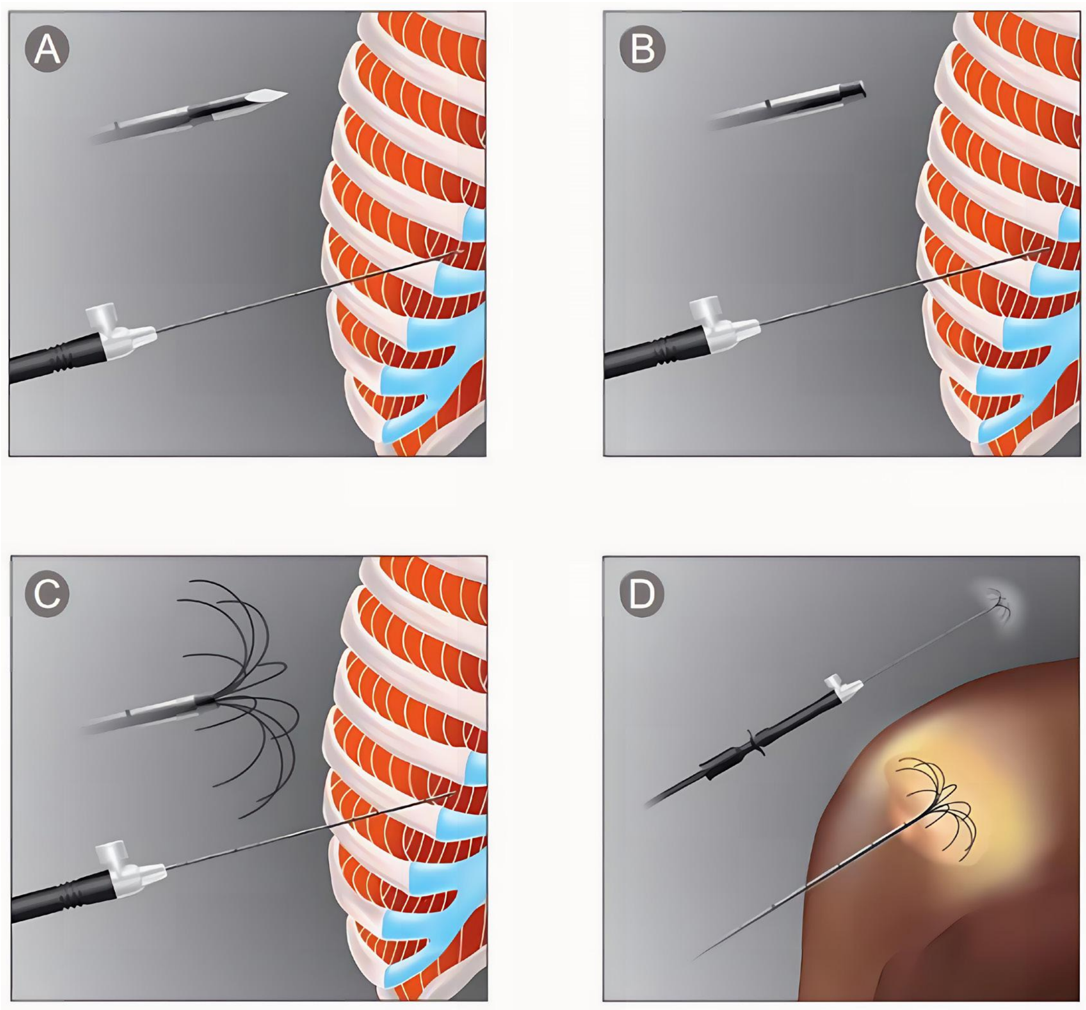


Figure 2. Schematic diagrams of radiofrequency ablation treatment with multipolar electrode needle. (A) a scalpel creates an access route, and the electrode needle is guided toward the liver area in the thoracic - abdominal cavity; (B) the electrode needle is inserted into the targeted liver region under proper guidance; (C) the multipolar electrodes of the needle deploy within the liver, preparing for energy delivery; (D) radiofrequency ablation is carried out, with the electrode acting on the liver lesion to complete the treatment. This figure is cited from [23].

Table 1. Material properties of four electrodes

Electrode Materials	Thermal Conductivity (W/m·K)	Electrical Conductivity (S/m)
Ni-Ti	18	4×10^6
Pt-Ir	71	4×10^6
Au	317	45×10^6
Cu	401	60×10^6

In terms of electrode material selection, Avishek et al. compared the physical properties of four electrode materials (Ni-Ti, Pt-Ir, Au, and Cu) based on numerical modeling, evaluating thermal conductivity and electrical conductivity in the thermal ablation region of a cylindrical liver tissue model (**Table 1**) [33].Despite the theoretical advantages of Au and Cu, the study found that the practical differences in ablation effect between various electrode materials are minimal and have limited clinical significance. When selecting electrodes, it is not necessary to overly focus on the thermophysical properties of the material; factors such as cost and biocompatibility should be prioritized instead.

Biological tissue parameters

Biological tissue parameters refer to the thermophysical properties of liver tissue and its surrounding structures in RFA. These include tissue density, specific heat capacity, thermal conductivity, electrical conductivity, dielectric constant, and blood perfusion rate, among others. These parameters play a crucial role in affecting the ablation temperature field and the extent of thermal damage. Accurate control of these parameters is beneficial for forecasting the ablation temperature distribution, determining the size of the ablation zone, and ensuring the smooth execution of the procedure. Thermal ablation of liver tissue is a complex process, and these parameters are not static; they continuously change with temperature [34]. Moreover, tissue parameters can vary even within different regions of the same tissue [35]. Therefore, it is extremely challenging to measure these parameters simultaneously under non-invasive conditions [36].

Chang et al. used NaCl solution to simulate the liver tissue to measure the temperature-dependent conductivity of the liver [37]. The simulation results showed that liver conductivity exhibited a nonlinear relationship with temperature, largely affecting the final temperature distribution within the tissue. Kröger et al. investigated for the nonlinear variations in tissue parameters with liver tissue temperature and dehydration status, and their simulation results closely matched the actual ablation [34]. Dos et al. used the probabilistic finite element method for the first time to quantify the impact of tissue parameter variations on the ablation damage region of the liver, considering individual and spatial differences [35]. Trujillo et al. observed that the electrical conductivity and permittivity of liver tissue varied linearly at temperatures below 100°C, providing corresponding functional expressions for these relationships [38]. Schutt et al. noted that blood perfusion rate directly affects the heat transfer process during RFA [39]. Their study on the nonlinear changes in blood perfusion rate showed significant differences in ablation outcomes depending on the calculation method used. Johansson et al. incorporated the temperature-dependent changes in thermal conductivity of both normal liver and tumor tissues, representing the conductivity parameters with piecewise expressions [40]. To improve the ablation rate of tumors while minimizing thermal damage to surrounding blood vessels, Fang et al. successfully optimized RFA treatment near vascular structures using a faceted method and computer modeling [41].

These studies underscore the importance of understanding and accurately measuring biological tissue parameters for optimizing RFA treatment protocols. This knowledge is essential for improving treatment efficacy and safety, which requires not only extensive experimental analysis but also robust mathematical modeling.

Biological heat transfer equation

In the study of RFA temperature fields, heat transfer in biological tissues is usually modeled using the biological heat transfer equation. Unlike conventional heat transfer, biological heat transfer is more complex due to the influence of tissue thermophoretic properties, blood perfusion rate, and metabolic activity [42]. This complexity makes the heat transfer process challenging to simulate accurately with a mathematical model, often requiring simplifications and assumptions.

As research into biological heat transfer progresses, various models have been continuously optimized and refined. Currently, four main biological heat transfer models are commonly used: the Pennes biological heat transfer model, the porous medium model, the Weinbaum-JiJi heat transfer model, and the Chen-Holmes heat transfer model [43-46].

• The Pennes biological heat transfer model is based on Fourier's law of heat conduction and simplifies human forearm tissues into a single cylindrical model. It incorporates blood perfusion rate and metabolic heat production, making it applicable for studying heat transfer in tissues.

• The porous medium model focuses on the heat transfer process between solid and liquid heat phases.

• The Weinbaum-JiJi model primarily focuses on venous heat transfer within muscle tissue, employing a three-layer skin-muscle composite model for more accurate simulations.

• The Chen-Holmes heat transfer model considers the relationship between biological heat transfer and the size and location of blood vessels within the tissues.

Among these models, the Pennes biological heat transfer model takes into account the effects of blood perfusion rate, local arterial blood temperature, and biological metabolic activities. It offers reliable computational results for thermal field simulation closely matching experimental findings. This model is characterized by its simplicity, convenience, and applicability for describing heat transfer process in soft tissues exposed to high temperatures, and it has been widely used in model simulations [47].

During RFA, Wang et al. derived an analytical solution for the blood and tissue temperature distributions, as well as an overall heat exchange correlation in cylindrical coordinates, based on porous medium theory [48]. This was done under local thermal non-equilibrium (LTNE) conditions to better represent the thermal dynamics during RFA. Sheu et al. conducted a 3D thermoelectric analysis of a system consisting of the liver, hepatic artery, and a 4-mm-diameter tumor, focusing on the impact of blood perfusion on heat transfer [49]. Vidya et al. found that closed-form expressions for heat transfer coefficients could accelerate RFA model calculations and identified vessel radius conditions necessary for directional effects on thermal lesions [50]. Tucci et al. developed a biothermal mathematical model based on variable porosity, providing more accurate predictions of the coagulation zone, coagulation diameter, and ablation volume, which can improve medical protocols and device design [51].

Overall, the heat transfer process in biological tissues is complex and difficult to simulate accurately with mathematical models. As numerical models continue to evolve, biological heat transfer equations for RFA will be further refined, aiming to minimize tissue damage and enhance accuracy in treatment protocols.

Evaluation of ablative thermal damage

The evaluation of ablative thermal damage is an important indicator of the effectiveness of RFA. In the process of RFA, tumor cell proteins undergo denaturation, leading to cell necrosis and the formation of a certain range of thermal damage area. Since this process is irreversible, the size (diameter, area, or volume) of the tumor necrotic region serves as the primary metric for thermal damage evaluation. Currently, the commonly used methods for tissue thermal damage assessment include Isothermal Contour method, Arrhenius model, and Thermal Isoeffective Does [52].

The isothermal contour method is the most commonly used quantitative technique in liver tumor RFA simulations. This method involves setting a specific temperature threshold within the ablation temperature field. When the temperature of the ablated tissue exceeds this threshold, it is assumed that the tissue has been thermally damaged by the electrode needle. The isothermal contour corresponding to this threshold serves as the boundary of the ablation thermal damage area. Since protein denaturation and coagulative necrosis of tumor cells typically occur within the temperature range of 50°C-60°C, thresholds of 50°C, 55°C and 60°C are commonly used as evaluation benchmarks for treatment [53]. The equation for the isotherm threshold method is:

$$V = \iiint_{\Omega} \mathrm{dv} \ (\Omega \geq \mathrm{T}°\mathrm{C}) \tag{1}$$

Where V represents volume of the ablated tissue, Ω represents the tissue region affected by thermal damage, and T is the temperature threshold.

Thermal Isoeffective Dose is a standardized metric used to quantify tissue damage in thermal therapies (e.g., radiofrequency ablation, microwave ablation). Its core concept is to compare the effectiveness of different treatment regimens by mathematically converting varying temperature-time exposures into an equivalent dose value. In clinical practice, the commonly used thermal dose unit is CEM43 (Cumulative Equivalent Minutes at 43°C), which indicates the equivalence of the actual temperature-time curve to the exposure time at 43°C [54]:

$$\mathrm{CEM43} = \sum_t R^{(43-T(t))} \cdot \Delta t \tag{2}$$

Where, t represents time, R= 0.5 (T>43°C) or R=0.25 (T<43°C), T means the thermodynamic temperature (K) at time t.

The Arrhenius model is a mathematical model derived from chemical reaction kinetics, commonly used to quantify tissue damage during thermotherapy or thermal ablation (e.g., radiofrequency ablation, microwave ablation). The core concept is that tissue damage is exponentially related to both temperature and time, with the probability of cell necrosis being predicted by integrating the effects of temperature and time [50]. The damage index (Ω) in the Arrhenius model is calculated as follows:

$$\Omega(\mathrm{t}) = \int_0^{\mathrm{t}} A \cdot e^{-\frac{Ea}{R \cdot T(\tau)}} \cdot \mathrm{d}\tau \tag{3}$$

Where A represents the frequency, Ea is the activation energy, R is the universal gas constant, and T (τ) is the temperature at time τ.

With the development of computer technology, the future direction of ablation thermal injury evaluation will focus on multimodal real-time assessment (combining impedance monitoring, temperature imaging, and artificial intelligence algorithms), dynamic tissue parameter correction (integrating individual differences such as blood perfusion and conductivity), and precise ablation boundary-defining techniques (e.g., elastography, high-resolution imaging), to enhance the personalized efficacy assessment and improve the precision of intraoperative modulation.

RFA temperature field simulation

The computational modeling of liver thermal ablation is primarily based on the Pennes equation, integrated with numerical methods such as finite element method and finite surface method, along with tissue damage models [55]. With the continuous development of computer technology and numerical computation, reconstructing a three-dimensional model of bioheat transfer using finite element modeling software has become a standard approach. This allows for finite element simulation and analysis to simulate the distribution of the temperature field during RFA. The utilization of finite element method greatly simplifies the complexity of controlling the temperature distribution in RFA and enhances the ability to predict the ablation region.

In the finite element simulation of RFA, several studies have contributed valuable insights. Ooi et al. established a three-dimensional impedance finite element model of RFA, focused on the dynamics of ablation temperature field, thermal damage area, and input voltage under different electro-thermal boundary conditions [56]. Chen et al. investigated the effects of variations in vessel size, horizontal distance between vessels, and the positioning of ablation electrode needles on the ablation temperature field through a three-dimensional simulation model of RFA that incorporated blood vessels [57]. Goldberg et al. coupled the radiofrequency electric field with the heat conduction process in their simulation of liver tissue during RFA [28]. Their study successfully predicted the temperature distribution within the liver during the ablation process, yielding ideal results that matched clinical outcomes. Lee et al. focused on a "claw-shaped" radiofrequency ablation electrode [58]. Their finite element simulation demonstrated that this electrode design effectively expanded the ablation range. Since then, finite element simulation has been widely used in the prediction of temperature field in RFA procedures.

Discussion

RFA has emerged as the cornerstone in the treatment of hepatocellular carcinoma, offering a minimally invasive, safe, and effective alternative to traditional surgical methods. This review synthesizes the technical principles, critical components (including electrode needle typology and materials), dynamic behavior of biological tissue parameters, bioheat transfer modeling approaches, and thermal damage assessment methodologies pertinent to RFA. Our analysis highlights the significant advancements made in this field, while also acknowledging the persistent challenges that remain.

Selection of electrode needles is a critical determinant of RFA therapeutic efficacy. Electrode needles are broadly classified as unipolar, bipolar, or multipolar configurations. The choice of electrode is influenced by tumor characteristics, including morphology and size. Unipolar and bipolar systems demonstrate high efficacy for tumors <3 cm in diameter, whereas multipolar electrodes are frequently employed for larger lesions (>3 cm) to achieve an expanded ablation volume in a single session. While computational simulations suggest that materials with high thermal conductivity (e.g., Cu, 401 W/m·K) or high electrical conductivity (e.g., Au, 45 $*10^6$ S/m) could theoretically enlarge the ablation zone, empirical validation has shown that material selection has only a negligible impact on the ablation volume, with differences of less than 0.1% observed in clinical settings. Consequently, clinical electrode material selection should prioritize biocompatibility, cost-effectiveness, and device engineering considerations over the pursuit of marginal thermophysical property optimization.

Thermal ablation of hepatic tissue is an intricate biophysical process influenced by several dynamic parameters. Notably, tissue thermal conductivity, electrical conductivity, and blood perfusion rate exhibit significant spatiotemporal variations, which are dependent on factors such as temperature and anatomical location. These variations directly influence the distribution of the resulting ablation temperature field. Although contemporary research employs nonlinear modeling frameworks (e.g., probabilistic finite element methods) to partially mitigate parameter uncertainty, the absence of non-invasive, real-time measurement techniques remains a fundamental bottleneck constraining the accurate prediction of ablation margins. Ongoing refinement of numerical systems and the bioheat transfer equations governing RFA is imperative to enhance predictive accuracy and minimize collateral tissue damage.

Bioheat transfer modeling for RFA predominantly utilizes four principal frameworks: the Pennes bioheat transfer model, the porous medium model, the Weinbaum-Jiji model, and the Chen-Holmes model. Among these, the Pennes model, valued for its computational efficiency and relative simplicity, remains the most widely adopted in RFA simulations for hepatic malignancies. The integration of numerical methods, such as finite element method and finite surface method, with tissue damage models allows for the creation of three-dimensional biothermal simulations using finite element software. Subsequent finite element simulation of the RFA temperature field distribution constitutes the prevailing methodological paradigm. Notably, advanced models, such as the porous medium and variable porosity approaches, enhance the physiological accuracy of these simulations. For thermal damage assessment, the isothermal threshold method and the Arrhenius integral model represent established standards; however, their intraoperative precision is significantly improved by integrating multimodal real-time monitoring techniques, such as impedance feedback and artificial intelligence-driven dynamic correction algorithms.

Conclusions

Optimizing RFA technology necessitates a multidisciplinary approach focused on three critical domains: the advancement of smart electrode systems incorporating shape-adaptive designs, the integration of individualized tissue parameter databases to better reflect patient-specific tissue properties, and the strategic exploration of multi-energy ablation techniques, such as RFA- microwave ablation synergy to broaden therapeutic indications. Crucially, achieving a paradigm shift from empirically guided interventions toward predictive precision requires rigorous large-scale clinical validation coupled with iterative refinement of computational models. Successfully implementing this “accurate prediction” framework holds significant potential to substantially improve long-term survival outcomes and quality of life for patients with hepatocellular carcinoma.

Author contributions: Shaobo Wang: Writing – original draft, Investigation, Data curation. Yuan Yao: Writing – review & editing, Methodology. Haipo Cui: Writing – review & editing, Supervision, Project administration.

References

[1] Mazzanti R, Gramantieri L, Bolondi L. Hepatocellular carcinoma: epidemiology and clinical aspects. Mol Aspects Med 2008;29(1-2):130-143.

[2] Sia D, Villanueva A, Friedman SL, et al. Liver Cancer Cell of Origin, Molecular Class, and Effects on Patient Prognosis. Gastroenterology 2017;152(4):745-761.

[3] Parikh ND, Pillai A. Recent Advances in Hepatocellular Carcinoma Treatment. Clin Gastroenterol Hepatol 2021;19(10):2020-2024.

[4] Benson AB 3rd, D’Angelica MI, Abbott DE, et al. NCCN Guidelines Insights: Hepatobiliary Cancers, Version 1.2017. J Natl Compr Canc Netw 2017;15(5):563-573.

[5] Ozcan M, Altay O, Lam S, et al. Improvement in the Current Therapies for Hepatocellular Carcinoma Using a Systems Medicine Approach. Adv Biosyst 2020;4(6):e2000030.

[6] Anwanwan D, Singh SK, Singh S, et al. Challenges in liver cancer and possible treatment approaches. Biochim Biophys Acta Rev Cancer 2020;1873(1):188314.

[7] Liu CY, Chen KF, Chen PJ. Treatment of Liver Cancer. Cold Spring Harb Perspect Med 2015;5(9):a021535.

[8] Solbiati L, Ierace T, Tonolini M, et al. Radiofrequency thermal ablation of hepatic metastases. Eur J Ultrasound 2001;13(2):149-158.

[9] Groeschl RT, Pilgrim CH, Hanna EM, et al. Microwave ablation for hepatic malignancies: a multiinstitutional analysis. Ann Surg 2014;259(6):1195-1200.

[10] Caspani B, Ierardi AM, Motta F, et al. Small nodular hepatocellular carcinoma treated by laser thermal ablation in high risk locations: preliminary results. Eur Radiol 2010;20(9):2286-2292.

[11] Kim YS, Lim HK, Rhim H, et al. Ten-year outcomes of percutaneous radiofrequency ablation as first-line therapy of early hepatocellular carcinoma: analysis of prognostic factors. J Hepatol 2013;58(1):89-97.

[12] Lencioni R, Crocetti L. Radiofrequency ablation of liver cancer. Tech Vasc Interv Radiol 2007;10(1):38-46.

[13] Hoffer EK, Drinane MC, Bhatnagar V, et al. Radiofrequency ablation of hepatocellular carcinoma guided by real-time physics-based ablation simulation: a prospective study. Int J Hyperthermia 2024;41(1):2331704.

[14] Spiliotis AE, Gäbelein G, Holländer S, et al. Microwave ablation compared with radiofrequency ablation for the treatment of liver cancer: a systematic review and meta-analysis. Radiol Oncol 2021;55(3):247-258.

[15] Habibi M, Berger RD, Calkins H. Radiofrequency ablation: technological trends, challenges, and opportunities. Europace 2021;23(4):511-519.

[16] Chu KF, Dupuy DE. Thermal ablation of tumours: biological mechanisms and advances in therapy. Nat Rev Cancer 2014;14(3):199-208.

[17] Cheong JKK, Yap S, Ooi ET, et al. A computational model to investigate the influence of electrode lengths on the single probe bipolar radiofrequency ablation of the liver. Comput Methods Programs Biomed 2019;176:17-32.

[18] McWilliams JP, Yamamoto S, Raman SS, et al. Percutaneous ablation of hepatocellular carcinoma: current status. J Vasc Interv Radiol

2010;21(8 Suppl):S204-S213.

[19] Rossi S, Fornari F, Pathies C, et al. Thermal lesions induced by 480 KHz localized current field in guinea pig and pig liver. Tumori 1990;76(1):54-57.

[20] Geoghegan R, Ter Haar G, Nightingale K, et al. Methods of monitoring thermal ablation of soft tissue tumors - A comprehensive review. Med Phys 2022;49(2):769-791.

[21] Deng Q, He M, Fu C, et al. Radiofrequency ablation in the treatment of hepatocellular carcinoma. Int J Hyperthermia 2022;39(1):1052-1063.

[22] Wang Z, Aarya I, Gueorguieva M, et al. Image-based 3D modeling and validation of radiofrequency interstitial tumor ablation using a tissue-mimicking breast phantom. Int J Comput Assist Radiol Surg 2012;7(6):941-948.

[23] Chen ZL. Design of minimally invasive radiofrequency ablation needle for tumor. University of shanghai for science and technology 2022.

[24] Zou YW, Ren ZG, Sun Y, et al. The latest research progress on minimally invasive treatments for hepatocellular carcinoma. Hepatobiliary Pancreat Dis Int 2023;22(1):54-63.

[25] Lee DH, Lee JM. Recent Advances in the Image-Guided Tumor Ablation of Liver Malignancies: Radiofrequency Ablation with Multiple Electrodes, Real-Time Multimodality Fusion Imaging, and New Energy Sources. Korean J Radiol 2018;19(4):545-559.

[26] Petras A, Moreno Weidmann Z, Echeverría Ferrero M, et al. Impact of electrode tip shape on catheter performance in cardiac radiofrequency ablation. Heart Rhythm O2 2022;3(6Part A):699-705.

[27] Yang Y, Chen Y, Ye F, et al. Late recurrence of hepatocellular carcinoma after radiofrequency ablation: a multicenter study of risk factors, patterns, and survival. Eur Radiol 2021;31(5):3053-3064.

[28] Goldberg SN, Gazelle GS, Halpern EF, et al. Radiofrequency tissue ablation: importance of local temperature along the electrode tip exposure in determining lesion shape and size. Acad Radiol 1996;3(3):212-218.

[29] Fang Y, Yan R, Jiao C, et al. Measurement methods for positioning accuracy of multileaf collimators in radiation therapy: A mini review. Prog Med Devices 2025;3(1):21-25.

[30] Lin SM, Lin CC, Chen WT, et al. Radiofrequency ablation for hepatocellular carcinoma: a prospective comparison of four radiofrequency devices. J Vasc Interv Radiol 2007;18(9):1118-1125.

[31] Jia Z, Zhang H, Li N. Evaluation of clinical outcomes of radiofrequency ablation and surgical resection for hepatocellular carcinoma conforming to the Milan criteria: A systematic review and meta-analysis of recent randomized controlled trials. J Gastroenterol Hepatol 2021;36(7):1769-1777.

[32] Hwang S, Kim JH, Yu SJ, et al. Incremental high power radiofrequency ablation with multi-electrodes for small hepatocellular carcinoma: a prospective study. BMC Gastroenterol 2024;24(1):280.

[33] Avishek S, Samantaray S. A study on the selection of electrode materials for hepatic radiofrequency ablation: A numerical approach. Mater Today Proc 2023;74:974-979.

[34] Kröger T, Altrogge I, Preusser T, et al. Numerical simulation of radio frequency ablation with state dependent material parameters in three space dimensions. Med Image Comput Comput Assist Interv 2006;9(Pt 2):380-388.

[35] Dos Santos I, Haemmerich D, Schutt D, et al. Probabilistic finite element analysis of radiofrequency liver ablation using the unscented transform. Phys Med Biol 2009;54(3):627-640.

[36] Radmilović-Radjenović M, Radjenović D, Radjenović B. Finite element analysis of the effect of microwave ablation on the liver, lung, kidney, and bone malignant tissues. Europhys Lett 2021;136(2):28001.

[37] Chang I. Finite element analysis of hepatic radiofrequency ablation probes using temperature-dependent electrical conductivity. Biomed Eng Online 2003;2:12.

[38] Trujillo M, Berjano E. Review of the mathematical functions used to model the temperature dependence of electrical and thermal conductivities of biological tissue in radiofrequency ablation. Int J Hyperthermia 2013;29(6):590-597.

[39] Schutt DJ, Haemmerich D. Effects of variation in perfusion rates and of perfusion models in computational models of radio frequency tumor ablation. Med Phys 2008;35(8):3462-3470.

[40] Johansson JD, Eriksson O, Wren J, et al. Radio-frequency lesioning in brain tissue with coagulation-dependent thermal conductivity: modelling, simulation and analysis of parameter influence and interaction. Med Biol Eng Comput 2006;44(9):757-766.

[41] Fang Z, Wei H, Zhang H, et al. Radiofrequency ablation for liver tumors abutting complex blood vessel structures: treatment protocol optimization using response surface method and computer modeling. Int J Hyperthermia 2022;39(1):733-742.

[42] Hahn DW, Özişik MN. Heat Conduction, 3rd Edition. Hoboken (USA). John Wiley and Sons, 2012.

[43] Pennes HH. Analysis of skin, muscle and brachial arterial blood temperatures in the

resting normal human forearm. Am J Med Sci 1948;215(3):354.
[44] Song S, Shen X, Ge Y, et al. Study of Heat Setting Machine Based on Porous Media Model. J Converg Inf Technol 2012;7(22):652-659.
[45] Jiji LM, Weinbaum S, Lemons DE. Theory and experiment for the effect of vascular microstructure on surface tissue heat transfer--Part II: Model formulation and solution. J Biomech Eng 1984;106(4):331-341.
[46] Chen MM, Holmes KR. Microvascular contributions in tissue heat transfer. Ann N Y Acad Sci 1980;335:137-150.
[47] Tucci C, Trujillo M, Berjano E, et al. Pennes' bioheat equation vs. porous media approach in computer modeling of radiofrequency tumor ablation. Sci Rep 2021;11(1):5272.
[48] Wang K, Tavakkoli F, Wang S, et al. Analysis and analytical characterization of bioheat transfer during radiofrequency ablation. J Biomech 2015;48(6):930-940.
[49] Sheu TW, Chou CW, Tsai SF, et al. Three-dimensional analysis for radio-frequency ablation of liver tumor with blood perfusion effect. Comput Methods Biomech Biomed Engin 2005;8(4):229-240.
[50] Vidya M, Divya MS, Priyadarshini N, et al. Computational modelling and analysis of thermal characteristics of DBS electrode in application to Parkinson's disease. 2014 International Conference on Advances in Electrical Engineering (ICAEE) 2014;1-4.
[51] Tucci C, Trujillo M, Berjano E, et al. Mathematical modeling of microwave liver ablation with a variable-porosity medium approach. Comput Methods Programs Biomed 2022;214:106569.
[52] Zhang B, Moser MA, Zhang EM, et al. A review of radiofrequency ablation: Large target tissue necrosis and mathematical modelling. Phys Med 2016;32(8):961-971.
[53] Arenas J, Perez JJ, Trujillo M, et al. Computer modeling and ex vivo experiments with a (saline-linked) irrigated electrode for RF-assisted heating. Biomed Eng Online 2014;13:164.
[54] Gas P, Kurgan E. Evaluation of Thermal Damage of Hepatic Tissue During Thermotherapy based on the Arrhenius Model. 2018 Progress in Applied Electrical Engineering (PAEE) 2018;1-4.
[55] van Erp GCM, Hendriks P, Broersen A, et al. Computational Modeling of Thermal Ablation Zones in the Liver: A Systematic Review. Cancers (Basel) 2023;15(23):5684.
[56] Ooi EH, Lee KW, Yap S, et al. The effects of electrical and thermal boundary condition on the simulation of radiofrequency ablation of liver cancer for tumours located near to the liver boundary. Comput Biol Med 2019;106:12-23.
[57] Chen R, Lu F, Wu F, et al. An analytical solution for temperature distributions in hepatic radiofrequency ablation incorporating the heat-sink effect of large vessels. Phys Med Biol 2018;63(23):235026.
[58] Lee ES, Lee JM, Kim WS, et al. Multiple-electrode radiofrequency ablations using Octopus® electrodes in an in vivo porcine liver model. Br J Radiol 2012;85(1017):e609-615.

Progress in Medical Devices
ISSN: 2957-5478

Zentime

Research Article https://doi.org/10.61189/389587wfmupp

A novel instrument guidance device for enhanced surgical navigation

Xuelong Wang[1], Changjiang Zhou[2], Zhenyi Huang[2], Hao Zhao[2], Shiju Yan[1]

[1]School of Health Science and Engineering, University of Shanghai for Science and Technology, Shanghai 200093, China. [2]Department of Ultrasound, Jinan City People's Hospital, People's Hospital Affiliated to Shandong First Medical University, Jinan 250000, Shandong Province, China.

Corresponding author: Changjiang Zhou.

Declaration of conflict of interest: None.

Received September 28, 2024; Accepted December 16, 2024; Published June 30, 2025

Highlights

- A compact instrument guidance device for minimally invasive surgery was developed, capable of real-time monitoring of angle and key status.
- The device provides navigation information and estimates the lesion depth, aiding in preoperative planning.

Abstract

Objectives: The key to an ultrasound-guided system is to accurately determine the spatial position of the surgical instrument. The size and shape of the optical positioning target in the optical positioning device can restrict the doctor's hand movements. Additionally, the use of a binocular camera occupies valuable operating room space. To reduce both the space occupation and the impact on surgical procedures, this paper presents a device designed to provide enhanced navigation information to assist the surgeon. **Methods:** The proposed device is equipped with an angle sensor and a button. The angle sensor measures the relative angle between the puncture device and the ultrasound probe in real time. The button assists the surgeon in measuring the depth of the surgical target before performing the puncture. Furthermore, we propose a filtering algorithm to reduce noise in the angle sensor's output signal. The device's performance was evaluated using an ultrasound instrument and an experimental setup for data collection and testing. **Results:** The filtering algorithm effectively reduced noise in the angular data. After filtering, the Signal Smoothness decreased from 0.036989 to 0.0010376, the Coefficient of Variation decreased from 0.00046682 to 0.0004323, and the Count of Local Extrema decreased from 251 to 135. Additionally, we collected and compared the coordinates of the guide line endpoints at the tip of the puncture needle and found an average point-to-point error of 3.53 pixel. **Conclusions:** The developed guidance device and filtering algorithm provide valuable navigation data, supporting the surgeon's operations.

Keywords: Ultrasound surgical navigation, filtering algorithm, minimally invasive surgery, angle sensor

Introduction

Minimally invasive techniques have become a mainstream approach in modern surgery due to their advantages of less trauma and faster recovery. These techniques are widely used in various clinical fields, including thoracic surgery, eye surgery, orthopedic surgery, abdominal surgery, neurosurgery, and obstetric and gynecological surgery [1-4]. Ultrasound guidance plays a crucial role in minimally invasive surgery by providing real-time navigation information, reducing physician fatigue, and improving surgical success rates. Studies have shown that ultrasound-guided techniques result in higher success rates in peripheral venous catheter insertion [5-7]. Pavone et al. showed that ultrasound-guided robotic surgery can shorten the duration of minimally invasive procedures [8]. Liang et al. found that ultrasound-guided acupotomy for trigger finger significantly increased the surgical success rate while reducing recurrence and complication rates compared to the traditional approach [9].

The core function of an ultrasound guidance system is to accurately determine the spatial position of surgical instruments and provide

Address correspondence to: Changjiang Zhou, Department of Ultrasound, Jinan City People's Hospital, No. 001 Xuehu Street, Jinan 250000, Shandong Province, China. Tel: +86-18956153985. E-mail: yanshiju@usst.edu.cn.

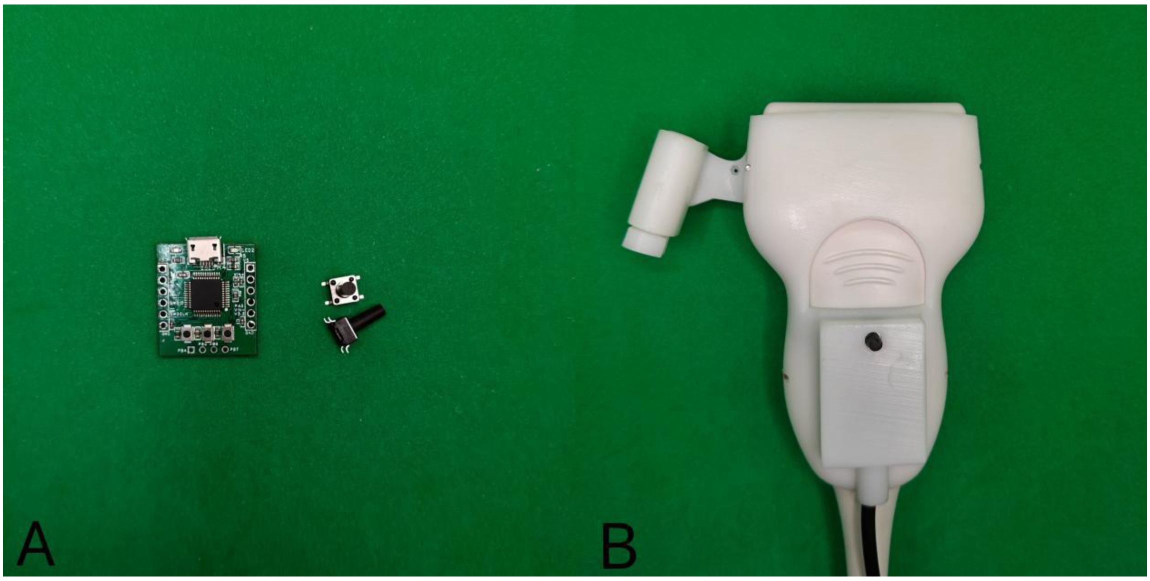


Figure 1. Circuit Board, Buttons and Instrument Guide Device. (A) Independently designed circuit board and button; (B) 3D printed and installed instrument guidance device.

navigation information to the surgeon. One commonly used method is optical localization-assisted technology, which involves placing optical tracking targets on both the ultrasound probe and surgical instruments, with binocular cameras positioned in the operating room [10-13]. The size and shape of the optical tracking sensor directly influence the surgeon's hand movements, and the presence of binocular camera occupies valuable space in the operating room. These factors can impact the efficiency and ease of surgery.

To reduce space occupation and minimize interference with the surgeon's operations, various small-scale surgical guidance techniques have been developed. Golafsoun et al. developed an MR-UIG system for central line procedures, which uses a six-degree-of-freedom transducer to track the spatial position of the ultrasound probe [14]. Minchev et al. designed a micro-robotic guidance device and tested it in 150 stereotactic procedures [15]. In this paper, we present a mechatronics-based instrument guidance device that provides navigation information to the surgeon. The device includes an angle sensor for real-time measurement of the relative angle between the puncture device and the ultrasound probe, and a keypad for assisting the surgeon in virtually measuring the depth of the surgical target before performing the puncture. We have designed the instrument to be compact and easy to hold, minimizing the impact on the surgeon's hand movements and reducing space occupation. Additionally, we propose a filtering algorithm to reduce noise in the sensor data.

Design and realization of instrument guidance device

Structural design of instrument guidance device

In this study, we designed a novel instrument guidance device for surgical navigation. The device is compact, easy to hold, and capable of obtaining precise angles for surgical instruments. It also assists the surgeon in virtually measuring the depth of the surgical target prior to puncture.

The device consists of an upper and lower casing, which are fixed together using screws and nuts. The lower casing has a pocket for the angle sensor, with a perforated design to allow the puncture guide to be mounted on the rotation axis of the angle sensor. The puncture guide transmits angle change to the sensor, enabling precise measurement of the relative angle between the puncture device and the ultrasound probe. This angle data is then used by the ultrasound navigation system to provide real-time visual guidance for the surgeon.

The upper casing includes grooves to house the sensor drive circuit board and a depth measurement button, as shown in **Figure 1A**. The button opening is located below the circuit board casing, allowing the surgeon to press it after determining the puncture angle to measure the depth of the surgical target. The virtual depth information obtained from this measurement aids in preoperative planning.

Additionally, we designed a puncture guide

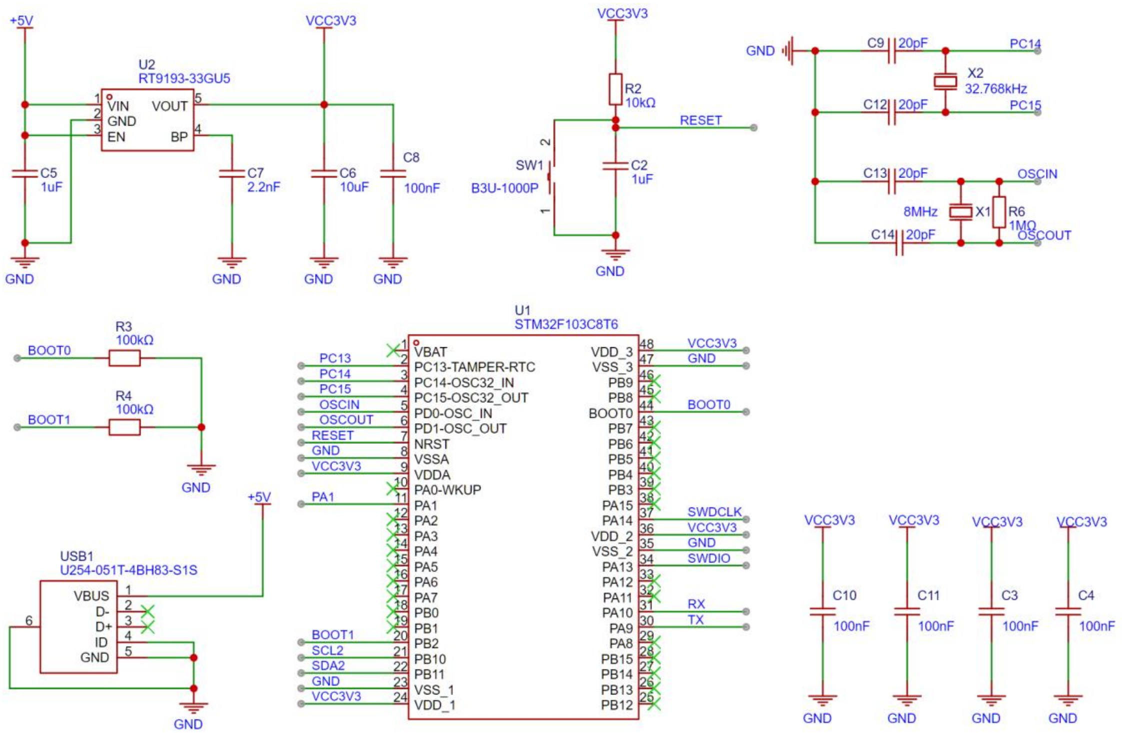


Figure 2. Schematic Diagram.

sleeve that accommodates puncture needles of varying sizes. The sleeve consists of an outer and an inner cylinder. A sensor fixing column at the lower part of the outer cylinder transmits the relative angle between the puncture guide sleeve and the ultrasound probe to the angle sensor. The inner cylinder has three different inner groove sizes to accommodate three different puncture needles. Finally, the design is realized by 3D printing and shown in **Figure 1B**.

Hardware design of the sensor drive module

The main function of the angle sensor drive module is to collect angle data and key status signals. The main controller of this module is STM32F103C8T6, which is sufficiently powerful to run the data acquisition and filtering algorithm designed for this project. The main control chip circuit is shown in **Figure 2**.

The STM32F103C8T6 microcontroller includes two built-in 12-bit analog-to-digital converters (ADC) that can be used for external analog signal acquisition. However, to enhance detection accuracy and provide more accurate navigation information, we have integrated an external 16-bit high-precision ADC chip, the ADS1110. This ADC chip features a built-in gain amplifier and supports four common sampling rates: 15SPS, 30SPS, 60SPS, and 240SPS.

The ADS1110 uses differential signal input pins (VIN+ and VIN-). It supports two acquisition modes: bipolar differential signal acquisition and single-ended acquisition. In this project, the single-ended acquisition mode is employed, with VIN- grounded and the input signal connected to VIN+. The single-ended signal range is from 0 V to 2.048 V.

For ADC chips, the conversion of input analog voltage to a digital value is governed by the following formula for an N-bit ADC:

$$Value = (\frac{V_{out}}{V_{ref}}) \times (2^N - 1) \qquad (1 - 1)$$

Where: Value is the digital output from the ADC, V_{out} is the analog output voltage of the potentiometer, V_{ref} is the reference voltage for ADC (2.048 V in this case), and N is the ADC resolution (16 bits in this project).

This project uses the RDC501051A angle sensor to obtain the relative angle between the the surgical puncture instrument and the ultrasound probe. The RDC501051A is a resistive position sensor known for its compact design and high accuracy. The linear relationship between the displacement of this sensor and its output voltage value is suitable for this project. The sensor has a resistance of 10 kΩ and a linearity accuracy of ± 2%. The position of the wiper of this potentiometer is linearly related to the output analog voltage, described by the following formula:

$$V_{out} = V_{CC} \times (\frac{P}{P_{max}}) \qquad (1 - 2)$$

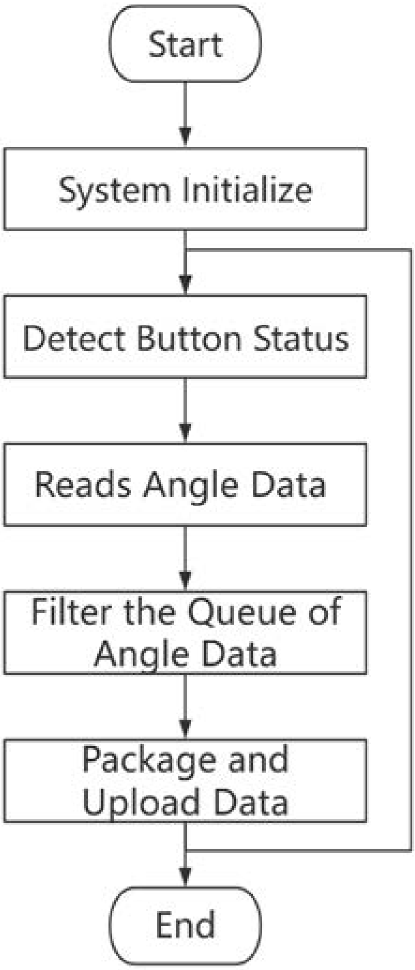


Figure 3. Software Design Process Flowchart.

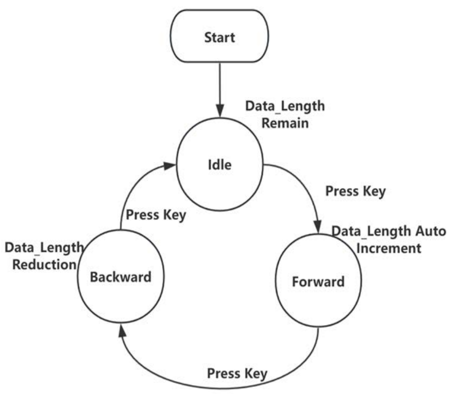


Figure 4. State Transition Diagram.

Where, V_{cc} is the specified value (2.048 V), P is the current position of the potentiometer wiper, and $Pmax$ is the maximum position of the potentiometer. When the wiper is at the starting position (P=0), the output voltage is 0 V. When the wiper is at the end position (P=$Pmax$), the output voltage is V_{cc}.

Software design of sensor drive modules

The angle sensor driver module is responsible for collecting and transmitting angle data and key status signals. The software program flowchart of this module is shown below in **Figure 3**.

The software follows a cyclic structure. After initializing each module, the chip continuously detects button status and angle data, and the filtered angle data and button status are then packaged and sent to the host computer. A key part to the software algorithm is filtering the angle data to eliminate outliers and minimize the impact of noise.

Noise is a common issue in data acquisition systems caused by various factors, such as electromagnetic interference, mechanical vibrations, motion artifacts, and imprecise circuit design [16-18]. This noise can affect the linearity of sensor and distort statistical properties, complicating accurate data analysis [19]. Signal-to-noise separation is a fundamental challenge in biological and data acquisition systems [20]. This project uses two approaches to filter and reduce noise: hardware circuit improvements and software filtering algorithms.

Hardware approach: A filter capacitor is connected between the sampling input pin and ground of the ADC chip. This capacitor can effectively filter high-frequency noise, improving the stability and accuracy of the ADC's signal conversion.

Software approach: Based on the observation of continuous static data waveforms, we identified the noise as impulse noise and Gaussian noise. The software filtering algorithm is divided into two parts: Part one, the angle detection driver module removes the maximum and minimum values from ten consecutive data points, then calculates the mean value; Part two, The host computer applies the standard deviation method to remove outliers from 100 consecutive data points, followed by mean filtering with a window size of 10. The standard deviation method used to filter impulse noise is calculated using the following formula:

$$s = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})}{n-1}} \qquad (1 - 3)$$

Where, s is the sample standard deviation, n is the number of data points in the sample, [illegible] is each individual data point, and $\bar{x}$ is the mean of the sample. The denominator uses n-l instead of n, which is known as degrees of freedom correction to provide an unbiased estimate of the population standard deviation.

In the board-level filtering process, we compute the mean $\bar{x}$ and standard deviation. If a data point exceeds the mean plus or minus three times, the data is determined to be outlier. The remaining data are then subjected to mean filtering to produce the final filtered result.

Depth measurement algorithm design

In response to feedback from the doctor, we designed a depth measurement function to aid in pre-surgical planning. This function allows the doctor to estimate the surgical target's penetration depth after determining the puncture angle, facilitating better surgical preparation.

A user button (model SW-6 mm) is added to the sensor driver module, see **Figure 1A**, enabling the doctor to adjust the depth measurement. The algorithm cycles between idle, forward, and backward states, with each state transition triggering a corresponding action, as shown in **Figure 4**. Pressing the button changes the user state, allowing for the adjustment of the depth measurement line in both forward and backward directions.

In the algorithm, we define a depth value, L_d (length of data), which is the modal length of the depth measurement line. This value increases when the button is pressed to enter "forward" mode and decreases for the "backward" mode. Importantly, this value is not the actual physical world penetration depth. To establish the relationship between L_d and the true puncture depth, we conducted data acquisition for 110 points. At each point, we measured the actual depth and performed polynomial fitting to obtain the following equation:

$$y = P_1 \times L_d + P_2 \quad \begin{cases} P_1 = 16.32 \\ P_2 = -1.074 \end{cases} \qquad (1-4)$$

Where x is the depth value L_d and y is the physical real puncture depth.

Based on the above equation, we derived the relationship between the physical puncture depth and the pixel coordinates in the ultrasound image:

$$y = P_1 \times \sqrt{(P_x + P_y)^2} + P_2 \qquad (1-5)$$

Where P_x and P_y are the pixel coordinates on the ultrasound image. The above formulas and algorithms were implemented on the host computer to enable the depth measurement function.

Results

We used a variety of data metrics to measure the performance of the devices and algorithms.

Signal Smoothness measures the degree of fluctuation and smoothness in the signal. It was assessed using the first and second derivatives:

$$S_1 = \frac{1}{N-1}\sum_{i=1}^{N-1} |y_{i+1} - y_i| \qquad (2-1)$$

$$S_2 = \frac{1}{N-1}\sum_{i=1}^{N-2} |y_{i+2} - 2y_{i+1} + y_i| \qquad (2-2)$$

Where y_i is the signal value at point i, and n is the total number of points, and y'_i and y''_i are its first-order and second-order derivatives.

The Coefficient of Variation (CV) measures the dispersion of the data relative to its mean. A lower CV indicates less data volatility:

$$CV = \frac{\sigma}{\mu} \times 100\% \qquad (2-3)$$

Where σ is the standard deviation and μ is the mean of the data.

Count of Local Extrema counts the number of local maxima and minima in the filtered data. A reduction in local extrema indicates fewer sharp changes:

For each point, compare it to its neighboring points;

If a point is greater than its neighboring, it is a local maximum;

If a point is smaller than its neighbors, it is a local minimum;

The total number of local maxima and minima is then calculated.

Root Mean Square Error (RMSE) measures the magnitude of the error between the predicted values and the actual observed values:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \qquad (2-4)$$

Where y_i is the true value and $\hat{y}_i$ is the predicted value for the i-th data point.

Mean Absolute Error (MAE) measures average absolute error between the predicted and actual values:

$$MAE = \frac{1}{n}\sum_{i=1}^{n} |y_i - \hat{y}_i| \qquad (2-5)$$

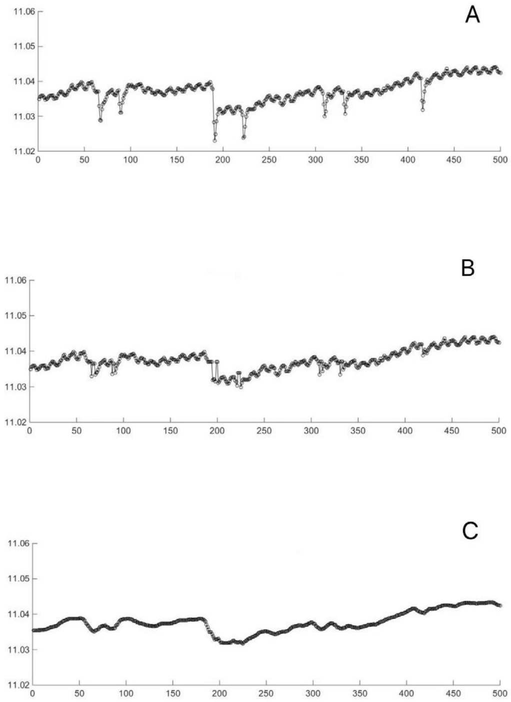


Figure 5. Filter algorithm result data. (A) Raw data; (B) Intermediate results; (C) Filtered results.

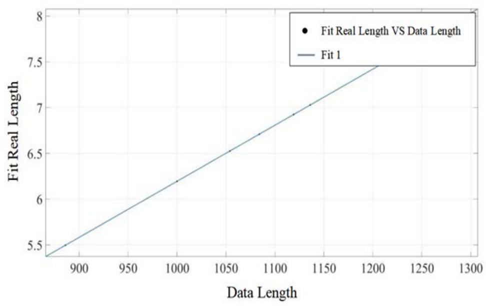


Figure 6. Fitting Function and Fitted Data Points.

Where y_i is the true value and $\hat{y}_i$ is the predicted value for the i-th data point.

The R-squared value indicates the proportion of variance in the dependent variable explained by the independent variables in the regression model. A value closer to 1 indicates better explanatory power:

$$R^2 = 1 - \frac{SSE}{SST} \qquad (2 - 6)$$

Where, SSE (Sum of Squares Errors) stands for the sum of squared residuals, which is the differences between the observed and predicted values. SST (Sum of Squares Total) stands for sum of squares total, which is the sum of squared differences between each observation and the overall mean.

Experimental results of filtering algorithm

The angle acquisition module was kept stationary, and 1,000 data points were collected to reflect the fluctuations caused by the instrument's inherent noise, with no external interference. Afterwards, the designed filtering algorithm was then applied to the data to obtain the result data.

The pre-filtered and post-filtered data are shown in **Figure 5**. **Figure 5A** shows the original data before filtering; **Figure 5B** shows the intermediate results after filtering using the standard deviation method with outlier data removed; **Figure 5C** shows the final filtered results after applying both the standard deviation method and mean filtering, with outliers and random noise removed.

As shown in the figures, the standard deviation method effectively filters out the transient impulse noise in the original data, while mean filtering removes random noise, resulting in smoother data. To evaluate the effectiveness of the filtering algorithm, we applied the performance index tests mentioned earlier to the data before and after filtering. The results of the performance metrics are summarized in **Table 1**.

The Signal Smoothness significantly improved, decreasing from 0.036989 to 0.0010376, the CV decreased from 0.00046682 to 0.0004323, and the Count of Local Extrema decreased from 251 to 135. These improvements suggest that the filtering algorithm enhances the accuracy and stability of angle data acquisition. The reduction in extreme values and peaks further indicates that the filtered data are more stable.

Depth measurement algorithm test results

To establish the relationship between the depth measurement value L_d and the actual penetration depth, we utilized MATLAB's polynomial fitting tool to generate a fitting formula. The formula, along with the fitted data points, is shown as **Figure 6**.

The results demonstrate a unitary function correspondence between L_d and true depth. The performance metrics of the fitted function are presented in the **Table 2**.

The coefficient of determination of the fitting formula is 1, and the RMSE and SSE values are extremely small, which indicates that the fitting formula has strong explanatory power and can

Table 1. Performance metrics of raw data, intermediate results and filtered results

	Signal Smoothness	CV	Count of Local Extrema
Raw data	0.03698900	0.00046682	251
Intermediate results	0.02928600	0.00044010	291
Filtered results	0.00103760	0.00043230	135

Note: CV, Coefficient of Variation.

Table 2. Performance parameters for fitting functions

SSE	R-squared	RMSE
5.9875e-28	1	2.1218e-15

Note: SSE, Sum of Squared Errors; RMSE, Root Mean Square Error.

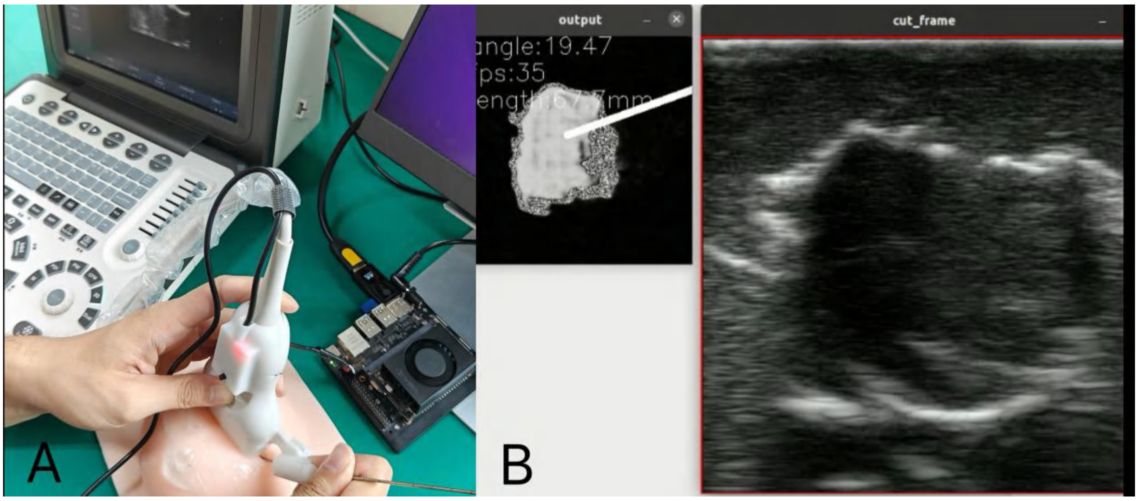


Figure 7. Experimental Bench and Experimental Effect. (A) Experimental Bench; (B) Upper Computer Operation Effect.

accurately describe the relationship between L_d and the real depth in the physical world.

We then developed a test program that used an image processing algorithm to measure the pixel coordinates of the puncture needle tip in the ultrasound image. The algorithm then compared these pixel coordinates with those of the guide-line endpoint, calculating the error to assess the performance.

In the tests, we collected coordinate data from 150 pairs of points. The average error between points was 3.53 pixels. Using the fitting formula, the corresponding average error for the depth measurement algorithm was found to be 0.216 mm. Finally, we constructed an experimental bench and performed simulation testing using the puncture needle with scale, ferrule, breast body model, ultrasound instrument, and probe fixation bracket designed in this paper. The setup and test results are shown in the **Figure 7**. The upper computer running results show that the angle at the center of the lesion is 19.47 degrees, with a calculated virtual depth of 67.7 mm.

Discussion

In this paper, an instrument guidance device capable of providing navigation information is presented. During the design process, several iterations were performed to optimize both the hand-holding sensation and device's size, resulting in a compact, user-friendly version.

We developed a filtering algorithm that processes data from the ADS1110 chip to provide accurate angle information to the physician. The filtered data meets the required accuracy and smoothness, with reduced noise. Further optimizations can be made at both the hardware and software levels. On the hardware side, adding an RC filter circuit could further reduce noise. On the software side, increasing the sampling rate to 240 sps could improve the real-time response speed and enhance the data volume, ensuring a faster and more responsive system.

To assist the surgeon in estimating the puncture depth before the procedure, we introduced a virtual guide-line feature. This feature has shown strong performance, but it can also be optimized further. Hardware improvements, such as incorporating a magnetic sensor for real-time depth measurement, could increase accuracy. On the software side, implementing a tip detection algorithm would allow for more precise real-time depth calculations.

Additionally, we developed an angular depth measurement algorithm tailored to the doctor's needs. This algorithm aids in preoperative planning by estimating the puncture depth using intraoperative imaging. The final test results demonstrate that the virtual bathymetry algorithm we designed performs well and meets the expected accuracy and reliability.

Conclusions

In this paper, we present an instrument guidance device that provides real-time navigation information to assist physicians during procedures. The device's mechanical structure was carefully engineered and 3D printed. An angle sensor, capable of measuring angle between the puncture device and the ultrasound probe, is integrated into the design. A sensor driver module drives the angle sensor and capture real-time keystroke information. We also developed a filtering algorithm to reduce noisy in the angle sensor data and a depth measurement algorithm to estimate the virtual depth before puncture.

The filtering algorithm effectively addresses the noise characteristics of the sensor data, ensuring accurate angle navigation for the physician. The depth measurement algorithm is precise in both positioning and measurement, helping doctors determine the accurate depth of the surgical target prior to surgery.

The instrument guidance device is compact, lightweight, and easy for the doctor to hold, without interfering with hand movements during the procedure. Data collected using an ultrasound instrument and experimental bench validated the performance of the algorithms. The results confirm that the device provides accurate and reliable navigation data to assist physicians in their procedures.

Author Contributions: Shiju Yan, as the leader of the project, formulated the detailed project plan, coordinated the resources of all parties, and comprehensively coordinated the planning and promotion of the project. Xuelong Wang is responsible for the design of the device structure, circuit board design, algorithm design and thesis writing. Changjiang Zhou is the co-corresponding author of the paper. He is responsible for researching the background of the project, planning and conceptualizing the paper, and providing suggestions and thoughts from clinical experience and doctors' perspectives. Zhenyi Huang was responsible for the assembly and testing of the device structure, as well as the operation of the device to collect a large amount of data. Hao Zhao assisted Zhenyi Huang in data collection and performance testing of the filtering algorithm and the device designed in this paper.

References

[1] Balafoutas D, Vlahos N. The role of minimally invasive surgery in gynaecological cancer: an overview of current trends. Facts Views Vis Obgyn 2024;16(1):23-33.

[2] Ditonno F, Franco A, Manfredi C, et al. Trends and Costs of Minimally Invasive Surgery for Kidney Cancer in the US: A Population-based Study. Urology 2024;189:41-48.

[3] Li R, Liu H, Zhang K, et al. Global tendency and research trends of minimally invasive surgery for glaucoma from 1992 to 2023: A visual bibliometric analysis. Heliyon 2024;10(16):e36591.

[4] Sang D, Guo J, Meng H, et al. Global Trends and Hotspots of Minimally Invasive Surgery in Lumbar Spinal Stenosis: A Bibliometric Analysis. J Pain Res 2024;17:117-132.

[5] Doniger SJ, Ishimine P, Fox JC, et al. Randomized controlled trial of ultrasound-guided peripheral intravenous catheter placement versus traditional techniques in difficult-access pediatric patients. Pediatr Emerg Care 2009;25(3):154-159.

[6] Skulec R, Callerova J, Vojtisek P, et al. Two different techniques of ultrasound-guided peripheral venous catheter placement versus the traditional approach in the pre-hospital emergency setting: a randomized study. Intern Emerg Med 2020;15(2):303-310.

[7] Otani T, Morikawa Y, Hayakawa I, et al. Ultrasound-guided peripheral intravenous access placement for children in the emergency department. Eur J Pediatr 2018;177(10):1443-1449.

[8] Pavone M, Seeliger B, Teodorico E, et al. Ultrasound-guided robotic surgical procedures: a systematic review. Surg Endosc 2024;38(5):2359-2370.

[9] Liang YS, Chen LY, Cui YY, et al. Ultrasound-guided acupotomy for trigger finger: a systematic review and meta-analysis. J Orthop Surg Res 2023;18(1):678.

[10] Xiao G, Bonmati E, Thompson S, et al. Electromagnetic tracking in image-guided laparoscopic surgery: Comparison with optical tracking and feasibility study of a combined

laparoscope and laparoscopic ultrasound system. Med Phys 2018;45(11):5094-5104.
[11] Yamada A, Tokuda J, Naka S, et al. Magnetic resonance and ultrasound image-guided navigation system using a needle manipulator. Med Phys 2020;47(3):850-858.
[12] Gao W, Jiang B, Kacher DF, et al. Real-time probe tracking using EM-optical sensor for MRI-guided cryoablation. Int J Med Robot 2018;14(1):10.1002/rcs.1871.
[13] Manni F, Elmi-Terander A, Burström G, et al. Towards Optical Imaging for Spine Tracking without Markers in Navigated Spine Surgery. Sensors (Basel) 2020;20(13):3641.
[14] Ameri G, Baxter JSH, Bainbridge D, et al. Mixed reality ultrasound guidance system: a case study in system development and a cautionary tale. Int J Comput Assist Radiol Surg 2018;13(4):495-505.
[15] Minchev G, Wurzer A, Ptacek W, et al. Development of a miniaturized robotic guidance device for stereotactic neurosurgery. J Neurosurg 2021;137(2):479-488.
[16] Liu J, Wang H, Kogos LC, et al. Optical spatial filtering with plasmonic directional image sensors. Opt Express 2022;30(16):29074-29087.
[17] Nathan V, Jafari R. Particle Filtering and Sensor Fusion for Robust Heart Rate Monitoring Using Wearable Sensors. IEEE J Biomed Health Inform 2018;22(6):1834-1846.
[18] Wang M, Li Z, Zhang Q, et al. Removal of Motion Artifacts in Photoplethysmograph Sensors during Intensive Exercise for Accurate Heart Rate Calculation Based on Frequency Estimation and Notch Filtering. Sensors (Basel) 2019;19(15):3312.
[19] Zhang Y, Wang R, Li S, et al. Temperature Sensor Denoising Algorithm Based on Curve Fitting and Compound Kalman Filtering. Sensors (Basel) 2020;20(7):1959.
[20] Taub M, Yovel Y. Segregating signal from noise through movement in echolocating bats. Sci Rep 2020;10(1):382.

Progress in Medical Devices
ISSN: 2957-5478

Zentime

Research Article
https://doi.org/10.61189/409228vdbtgj

Design and experimental study of novel plasma ablation electrodes

Siqi Zhao[1], Lin Mao[1], Chengli Song[1], Zhuotianhao Wang[2]

[1]Shanghai Institute for Minimally Invasive Therapy, School of Health Science and Engineering, University of Shanghai for Science and Technology, Shanghai 200093, China. [2]Department of Biology, Shenzhen MSU-BIT University, Shenzhen 518000, Guangdong Province, China.

Corresponding author: Lin Mao.

Declaration of conflict of interest: None.

Received October 23, 2024; Accepted January 21, 2025; Published June 30, 2025

Highlights

- Three different electrode structures, including ring-needle electrode, needle electrode, and cylinder electrode, were designed to accommodate diverse surgical requirements.
- A novel low-temperature plasma ablation electrode features interchangeable tips, enabling surgeons to select appropriate tips based on specific operative situations.
- With increased electric field strength, ring-needle electrodes facilitates liquid medium breakdown, thereby opti mizing plasma generation and ablation efficacy.

Abstract

Objective: This study utilizes finite element modeling to investigate the coupled electric-thermal field distributions for three novel electrode designs, assessing their ability to induce electric breakdown in saline for plasma generation. Furthermore, the ablation effects of various electrode shapes were validated through ex vivo tissue ablation experiments, ensuring both the safety and feasibility of the electrodes. **Methods**: Three electrode structures were designed: the ring-needle electrode, the needle electrode, and the cylinder electrode. A COMSOL Multiphysics finite element model simulated their behavior in saline. This modeling approach enabled a detailed analysis of the spatial variations in both the electric and temperature fields. Furthermore, the electrodes were tested at four voltages (180 V, 220 V, 260 V, and 300 V) and a frequency of 100 kHz on porcine liver tissue to evaluate ablation performance. **Results**: Simulations showed temperatures of 25-70 °C at 0.3 mm above the three electrodes, with electric field strength exceeding 1 *10^6 V/m, which are sufficient to trigger electric breakdown and plasma formation. Ex vivo experiments confirmed ablation efficacy, with the ring-needle electrode exhibited the best performance at a voltage of 300 V, achieving a 2.3 mm ablation depth and 0.72 mm^2 thermal damage area. **Conclusion**: Finite element simulation and ex vivo experiments demonstrate the feasibility of the proposed electrodes, highlighting their potential as an innovative solution for plasma-based ablation technologies.

Keywords: Low-temperature plasma, ablation electrode, structural design, Multiphysics simulation

Introduction

Tonsils, or palatine tonsils, are lymphoid tissues located in the upper oropharyngeal and nasopharyngeal regions, serving a crucial immunological function in infection defense. In clinical practice, patients may experience varying degrees of pharyngeal pain, which may significantly impair both physical and psychological well-being, ultimately affecting their overall quality of life. Therefore, prompt and effective intervention is crucial to mitigate these adverse effects [1]. For patients with chronic pharyngeal inflammation secondary to recurrent acute tonsillitis, peritonsillar abscesses, or tonsillar hypertrophy-induced obstructive sleep apnea, symptom severity dictates the necessity for immediate treatment. Cases demonstrating severe symptoms require urgent medical attention. Appropriate therapeutic intervention must be implemented promptly to minimize adverse health impacts and maintain quality of life. Un-

Address correspondence to: Lin Mao, Shanghai Institute for Minimally Invasive Therapy, School of Health Science and Engineering, University of Shanghai for Science and Technology, NO.516 Jungong Road, Yangpu District, Shanghai 200093, China. Tel: +86-21-55572159. E-mail: linmao@usst.edu.cn.

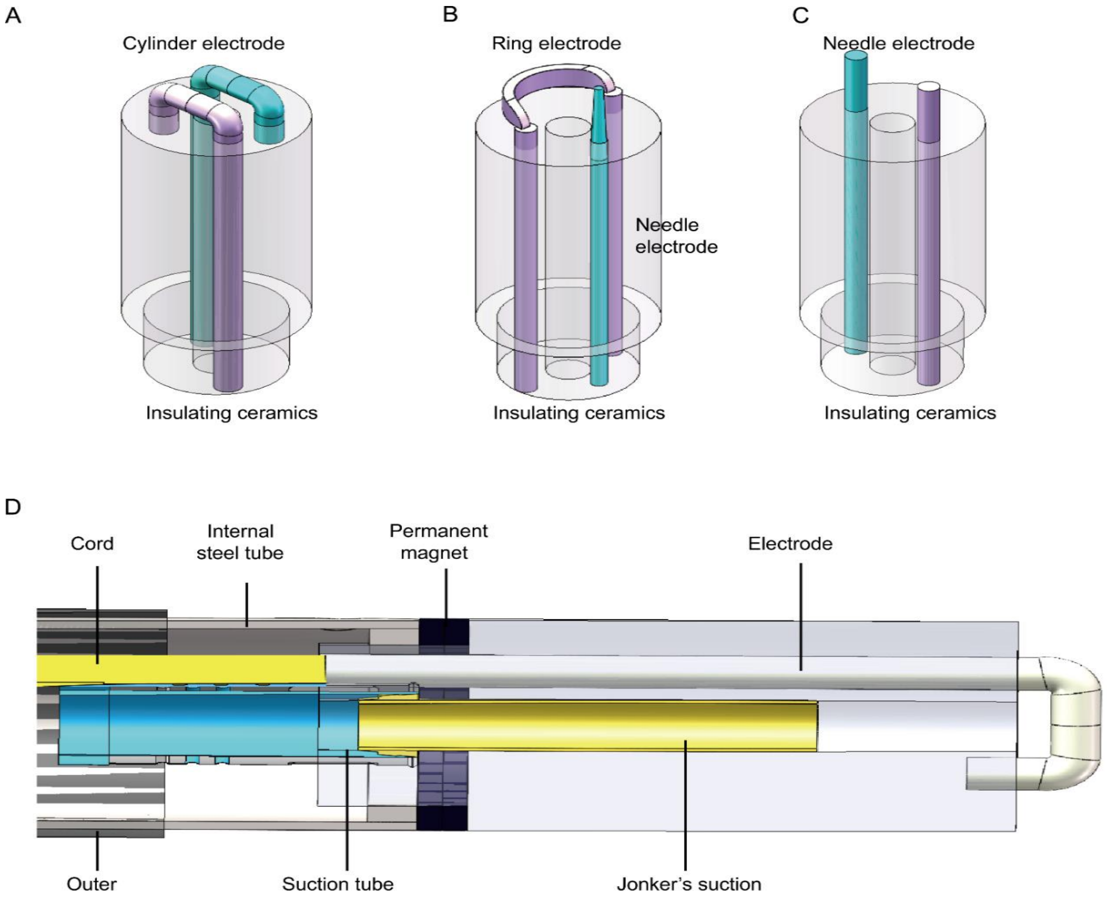


Figure 1. Three-dimensional model of a low-temperature plasma ablation tip. (A) Cylinder electrode; (B) Ring-needle electrode; (C) Needle electrode; (D) Detailed structural view of the ablation head.

der clinical supervision, tonsillectomy may represent a more effective treatment option over conservative management [2].

In otolaryngological practice, tonsillectomy techniques are classified into two categories: cold dissection (conventional surgical excision) and thermal methods employing high-frequency electrosurgery [3, 4]. Currently, electric surgical tools widely used in medical practice, such as high-frequency instruments, function by generating resistive heat during current passage through electrodes to achieve tissue sectioning. This significant thermal mechanism frequently induces substantial collateral tissue damage, potentially compromising wound healing and promoting excessive scar formation [5-8]. Conversely, traditional scalpel-based procedures are often associated with significant postoperative pain due to their mechanical cutting action. Furthermore, cold knife surgery lacks inherent hemostatic property, potentially increasing both intraoperative and postoperative hemorrhage, thereby requiring supplementary hemostatic measures [9-11].

Recent advances have introduced low-temperature plasma ablation technology as a minimally invasive alternative that overcomes limitations of conventional techniques regarding surgical precision, hemorrhage control, and thermal injury. This technology facilitates precise dissection and efficient coagulation through small incisions, offering distinct advantages including reduced postoperative discomfort, faster recovery, improved hemostasis, and decreased thermal damage compared to traditional methods [12-16]. These characteristics demonstrate the significant clinical potential and promising applications of low-temperature plasma technology in surgical practice.

China has made significant strides in the field of low-temperature plasma ablation scalpel technology, with research primarily concentrating on liquid-phase plasma generation [17, 18]. Key developments include multi-channel drainage circulation systems and optimized suction ports. Current plasma scalpel designs integrate anti-clogging mechanisms and ensure uniform saline coverage of electrodes. Notably, significant advancements have been made in liquid injection systems, which facilitate stable plasma layer formation while providing effective cooling and cleaning of ablation sites. However, previous studies have primarily focused on innovative suction and irrigation systems to minimize electrode-induced thermal damage. Less attention has been paid to issues such as non-replaceable cutter heads and the monomorphic emitter electrode designs. These design constraints prevent surgical adaptation to varying operative conditions, potentially compromising ablation efficacy - an issue requiring urgent investigation. Electrode design must carefully balance thermal damage minimization with surgical effectiveness, considering both surgical outcomes and patient safety.

Precise control of plasma scalpel thermal effects is essential to prevent collateral tissue damage. Empirical evidence indicates that optimized electrode structural design, along with the fine-tuning of plasma parameters, can significantly contribute to the precise temperature control and the mitigation of unintended thermal injury [19, 20]. Finite Element Modeling and Finite Element Analysis serve as potent tools within the biomedical domain, they facilitate an in-depth investigation of electrothermal field variations and foster the development of novel technologies [21, 22]. By employing COMSOL Multiphysics, we conducted an in-depth analysis of the coupled thermal - electrical interactions of various electrode configurations in physiological saline, yielding essential theoretical data for assessing system stability and safety.

Materials and methods

3D models

A three-dimensional model of the low-temperature plasma ablation head was developed, as depicted in **Figure 1**. This model is primarily comprised of a titanium electrode head, an insulating ceramic component, and essential ancillary devices. The present study introduces three distinct electrode configurations: the cylinder electrode (**Figure 1A**), the ring-needle electrode (**Figure 1B**), and the needle electrode (**Figure 1C**), each fabricated from titanium. These electrodes feature one side electrically grounded and the opposing side serving as the anode. The design incorporates insulated permanent magnets at the proximal ends of both the ceramic head and inner steel tube. These magnets are arranged with opposing polarities, ensuring secure attachment to the scalpel handle through magnetic coupling, as illustrated in **Figure 1D**.

Physical model control equations

Finite element analysis was performed using the COMSOL Multiphysics (COMSOL, UK) to investigate the coupling between electromagnetic and thermal fields. The simulation incorporated the ‘Electric Currents’ and ‘Heat Transfer in Solids’ modules for multiphysics coupling. The electric field distribution was governed by fundamental electromagnetic principles:

Current continuity described by Ohm’s law ($J=\sigma E$), where J represents current density (A/m^2), σ denotes conductivity (S/m), and E is electric field intensity (V/m) [23];

Electrostatic field distribution following the Laplace equation: $\nabla \cdot(\sigma \nabla V)=0$ [24];

Electric field-potential relationship: . $E=-\nabla V$

The thermal field was modeled using the heat transfer equation [25]:$\rho C\ \partial T/\partial t-\nabla\cdot(k\nabla T)=J\cdot E$. In this context, T represents the temperature (°C), ρ denotes the mass density (kg/m^3), C indicates the specific heat capacity (J/(kg*K)), t refers to time (s), and J·E accounts for Joule heating.

Material properties

The electrothermal characteristics of plasma generation in saline were analyzed using COMSOL 6.0. Material parameters were selected according to simulation requirements, with key properties listed in **Table 1**. The temperature-dependent saline properties were derived from empirical equations, where T represents the real-time temperature (°C) [26, 27].

Simulation models and boundary conditions

To optimize computational efficiency, the original saline model was simplified to a cylindrical structure. This streamlined model reduces computational complexity, accelerates calculation speed, while preserving model accuracy, enabling faster simulations with lower computational costs. A detailed schematic of the simplified model is presented in **Figure 2A.**

Boundary conditions were defined for both the contact surface between the geometric model and the surrounding environment, and the internal contact surface within the model. It is necessary to add a potential boundary condition and a ground boundary condition for the current (ec). Specifically, in the simulation of the ring-needle electrode, one side was assigned a positive potential, while the opposite side was grounded. The global parameters, as indicated in **Table 1**, were set to 260 V. Addi-

Table 1. Material properties

Position	Materials	Displayed formula	Value	Properties
Ceramic Glass	rho_Gl	2595 [kg/m^3]	2595 kg/m^3	Densities, Ceramic Glass
	k_Gl	1.09 [W/(m*K)]	1.09 W/(mK)	Thermal conductivity, Ceramic Glass
	Cp_Gl	750 [J/(kg*K)]	750 J/(kgK)	Heat capacity, Ceramic Glass
	sigma_Gl	1e-5 [S/m]	1E-5 S/m	Electrical conductivity, Ceramic Glass
Electrode	rho_Ti	4500 [kg/m^3]	4500 kg/m^3	Densities, Titanium
	k_Ti	21.90 [W/(m*K)]	21.90 W/(mK)	Thermal conductivity, Titanium
	Cp_Ti	523 [J/(kg*K)]	523 J/(kgK)	Heat capacity, Titanium
	sigma_Ti	2.5e5 [S/m]	2.5E5 S/m	Electrical conductivity, Titanium
	V0	260 [V]	260 V	Initial voltage
	T0	19.85 [deg C]	293 K	Initial temperature

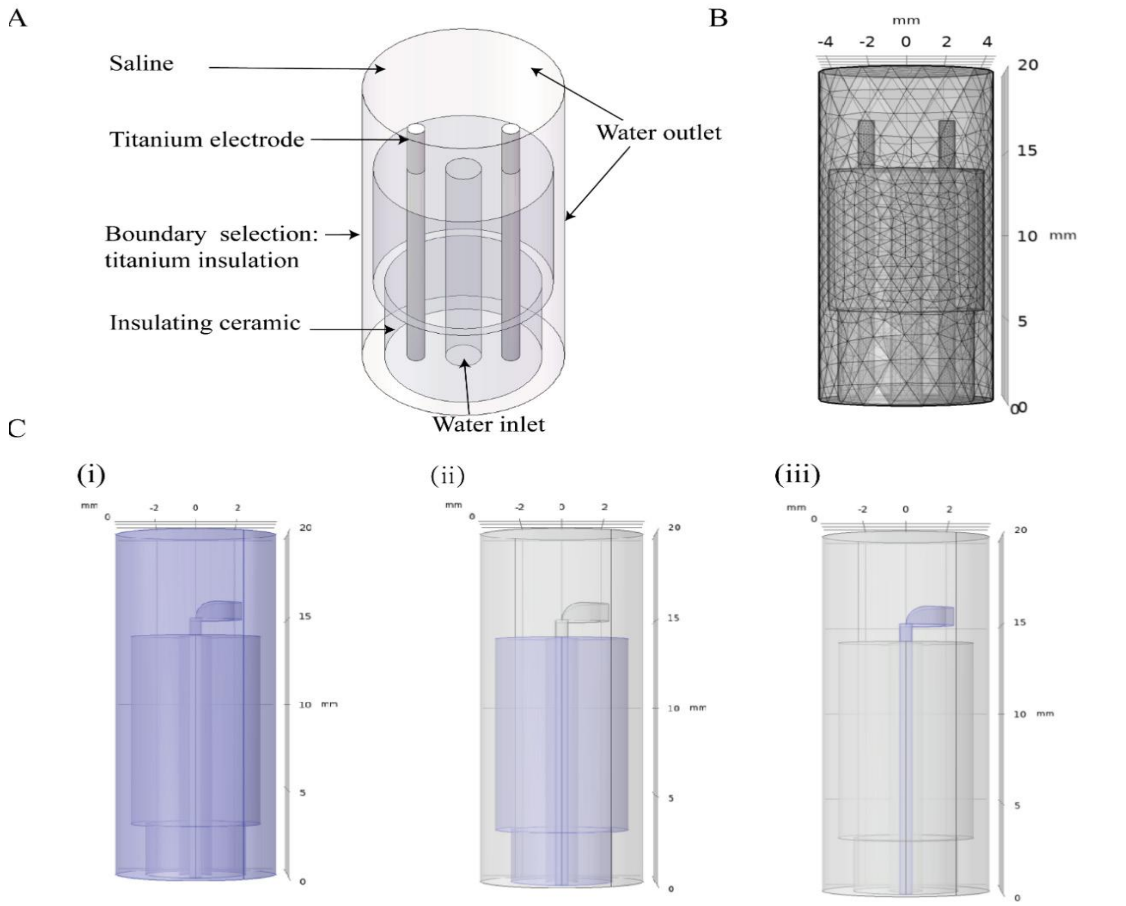


Figure 2. Simulation model and its specific parameter settings. (A) Simplified model of the electrode tip; (B) 3D meshing diagram of the model; (C) Electrode boundary geometric entity layers: electrode (i); insulating ceramic (ii); physiological saline (iii).

tionally, the surface in contact with the physiological saline and the external environment was designated as electrically insulating. In terms of solid heat transfer, the initial temperature was set to 293 K, as shown in **Table 1**, with the boundary conditions illustrated in **Figure 2C**.

To balance accuracy and computational efficiency, the grid size was optimized, and the model was discretized into tetrahedral meshes. A needle electrode was selected as a representative model for the simulation. The geometry included four domains, 26 boundaries, 44 edges, and 28 vertices. The mesh consisted of 102,482 domain elements, 10,412 boundary elements, and 752 edge elements. **Figure 2B** depicts a 3D diagram of the generated mesh. The coupled electric-thermal fields were solved using transient time stepping, employing PARDISO direct solver for the simulation [28].

Electrode ablation experiment

Three different electrode tip configurations were fabricated using 3D printing technology and subsequently assembled. Discharge and

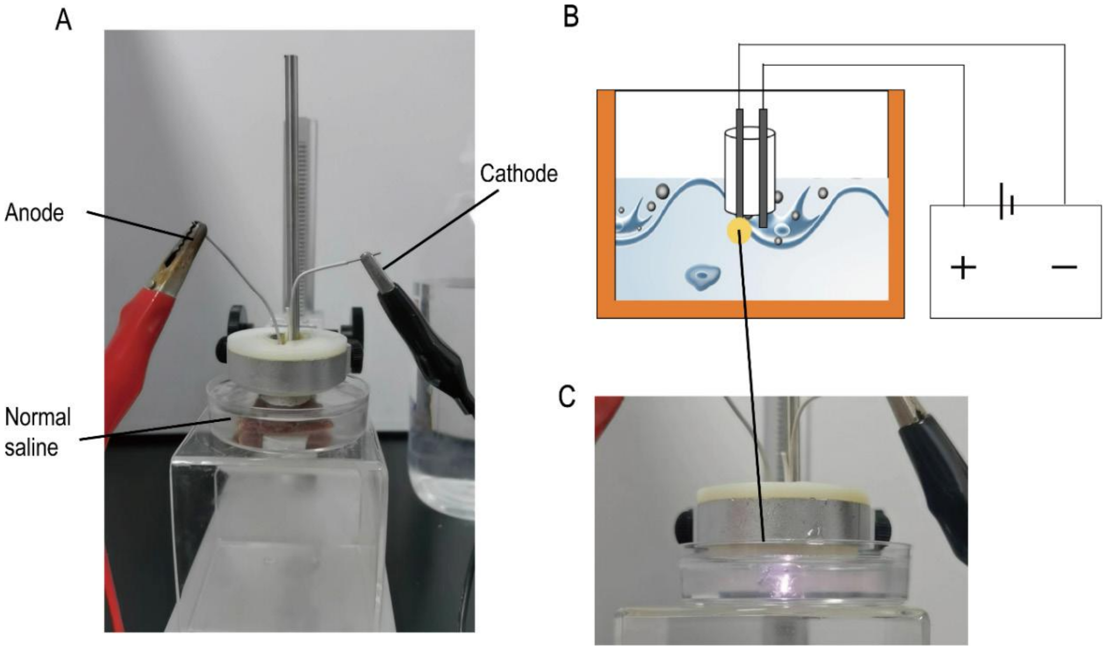


Figure 3. Discharge and ablation experiments. (A) Physical diagram; (B) Schematic diagram; (C) Detailed diagram.

ablation experiments were performed in physiological saline at various voltage levels, with the experimental setup shown in **Figure 3**.

At standard temperature and pressure, discharge measurements were conducted on plasma equipment connected to the same electrode at different voltage levels: 180 V, 220 V, 260 V, and 300 V, and the observed phenomena were documented. Simultaneously, the discharge characteristics of plasma equipment with differently structured electrodes were measured at the same voltage levels. A total of 12 photographs were taken to illustrate the experimental phenomena, with particular emphasis on the generation of orange plasma at varying voltages. Due to variations in electrode shapes, the voltage required for stable plasma generation differed among plasma devices. Voltages of 180 V, 220 V, 260 V, and 300 V were applied, and ablation of porcine liver tissue was performed at these voltage levels. To compare the ablation effects of the three electrodes, the ablation process was carried out at four distinct voltages. In total, four sets of experiments were conducted for electrodes with different shapes, resulting in 12 experimental sets.

Results

Temperature field simulation of the 3D electrode model

The temperature field simulation results for three electrodes are shown in **Figures 4A-C**. The temperature of the physiological saline, in regions distal to electrodes, exhibited slight fluctuations over time, maintaining a stable value of 19.85 °C. Notably, the temperature increased around the electrode as the distance decreases, reaching a maximum of 110.298 °C for the needle electrode (**Figure 4A**). The peak temperature in the ring-needle electrode simulation was approximately 106°C (**Figure 4B**), while the peak temperature of the cylinder electrode was 94 °C (**Figure 4C**). Temperature variations were examined at a distance of 0.3 mm above the electrodes. For the needle electrode (**Figure 4D**), the temperature of the physiological saline gradually increased over time. At time t = 1E-5, the temperature 0.3 mm above the needle electrode ranged from 20 °C to 110 °C, with higher temperatures observed near the electrode and temperatures remaining below 40 °C further away. For the ring-needle electrode (**Figure 4E**), a noticeable temperature gradient was observed in a localized area centered around the electrode. These results suggest a concentration of thermal energy near the electrode tips, which may influence tissue interaction and ablation effectiveness. At time t = 1E-5, the temperature of the physiological saline near the electrode peaked at approximately 100°C, while temperatures in the surrounding regions dropped slightly below 35 °C. As depicted in Figure 4F, the temperature near the cylinder electrode gradually escalated before reaching a plateau, with a range of approximately 32 °C to 47 °C.

Electric field simulation of the 3D electrode model

The one-dimensional electric field distributions for three electrode configurations are shown in **Figures 5A-C**. This simulation study investigates

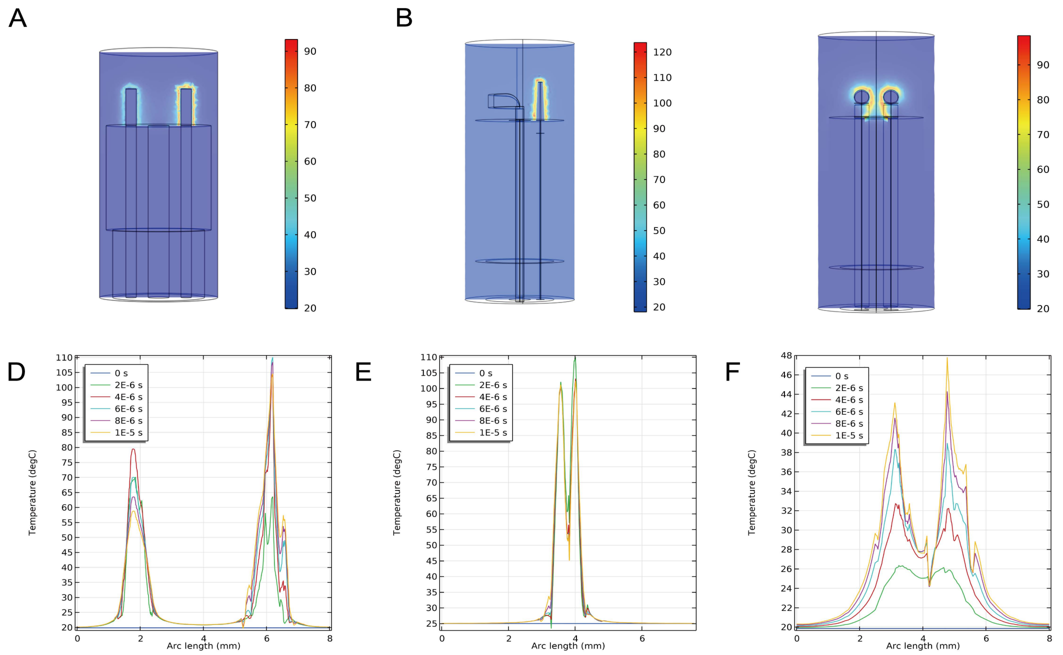


Figure 4. Temperature field simulation of different electrodes and the temperature variation curve over time at 0.3 mm above different electrodes. (A) Needle electrode; (B) Ring-needle electrode; (C) Cylinder electrode; (D) Needle electrode; (E) Ring-needle electrode; (F) Cylinder electrode.

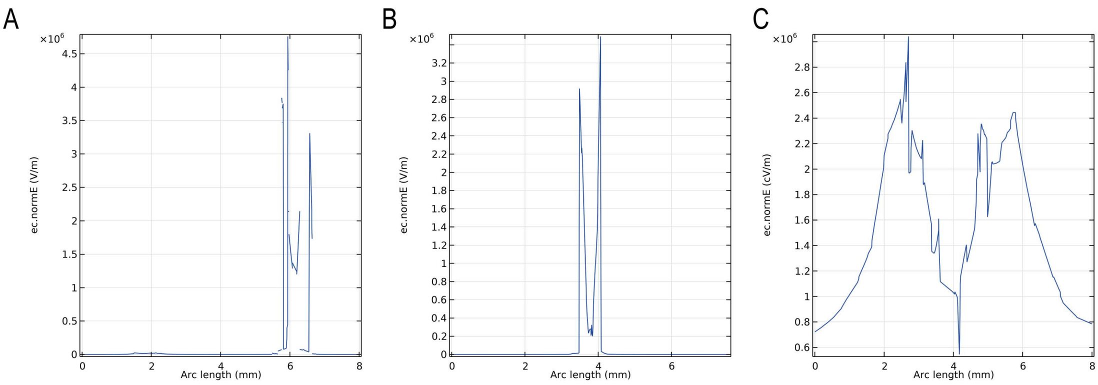


Figure 5. One-dimensional diagram of the electric field distribution for different electrode discharging in physiological saline. (A) Needle electrode; (B) Ring-needle electrode; (C) Cylinder electrode.

the correlation between electric field intensity and arc length, with a particular focus on field strength variations along the discharge path.

As shown in **Figure 5A**, the field strength in the central region of the needle electrode reached 1.57*10^6 V/m, meeting the minimum plasma formation threshold. **Figure 5B** illustrates the one-dimensional electric field model of the ring-needle electrode. The field strength in the central ring region was 1.2*10^6 V/m, while the field strength in most discharge areas of the electrode exceeded 1.8*10^6 V/m, and the maximum field strength reached 3.4*10^6 V/m. For the cylinder electrode, as shown in **Figure 5C**, the field strength near the electrode was 1.6*10^6 V/m. Notably, the field strength exceeded the 1*10^6 V/m threshold in majority regions. A comparative analysis of the one-dimensional electric field distribution images for the three electrode configurations under identical voltage and temperature conditions reveals that the local electric field intensity of the ring-needle electrode was markedly higher than that of the needle and cylinder electrodes.

Electrode ablation experiment

The discharge properties of the electrode were systematically examined across a range of voltages. **Figure 6A** illustrates the electrodischarge phenomena observed for the needle electrode at four distinct voltage levels (180 V, 220 V, 260 V, and 300 V). In **Figure 6A(i)**, it is evident that no orange substance was generated. **Figures 6A(ii-iii)** demonstrate that at volt-

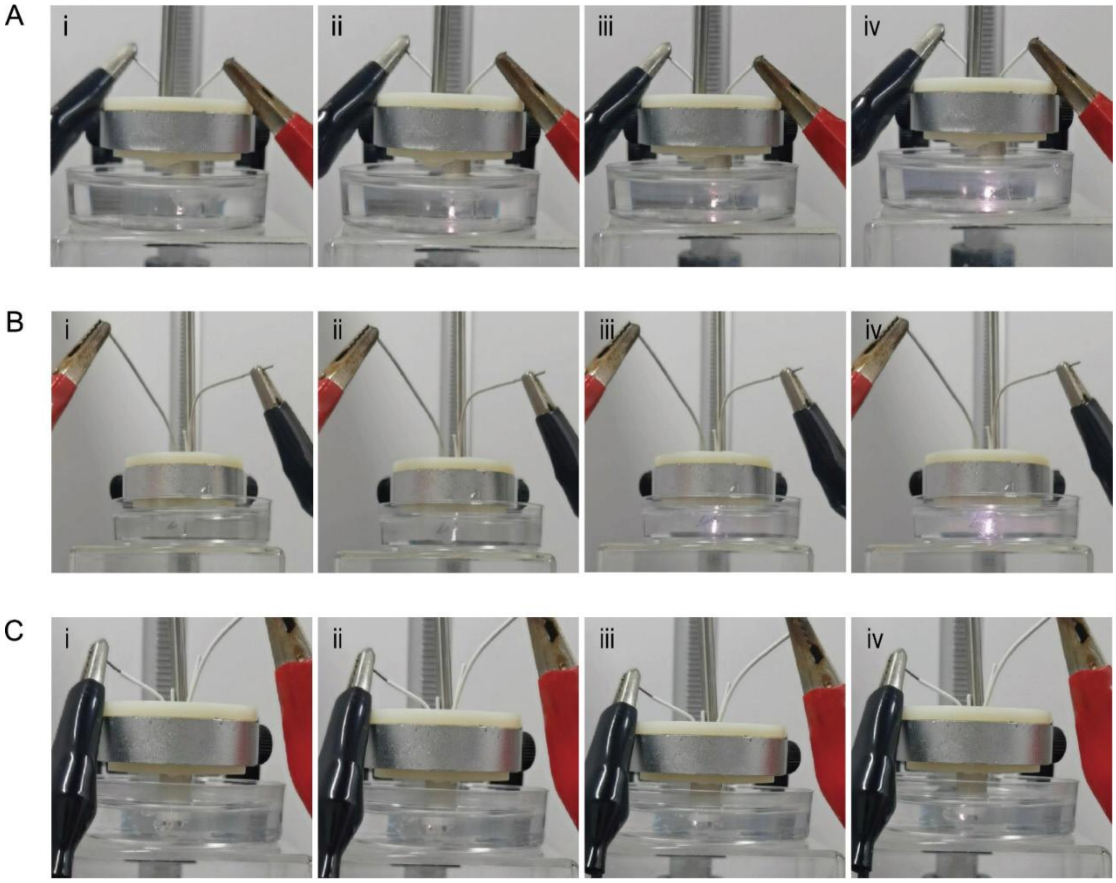


Figure 6. Discharge behavior of three types of electrodes in saline solution at different voltages. (A) Needle electrode; (B) Ring-needle electrode; (C) Cylinder electrode; i-iv: 180 V, 220 V, 260 V, and 300 V.

ages of 220 V and 260 V, orange substances, identified as plasma, were generated, although the color remains relatively dark. **Figure 6A(iv)** reveals that at 300 V, the electrode produced a stable bright orange substance. As the voltage increasing, the effects of the needle electrode became increasingly pronounced (**Figure 6B**). Experimental results indicate that at 300 V, the orange plasma light emitted by the needle electrode reached its maximum brightness. **Figure 6C** shows that while the intensity of the orange discharge from the cylinder electrode augmented with increasing voltage, it remained lower compared to the needle electrode, even at the highest voltage tested. Among the four voltage levels evaluated, the cylindrical electrode exhibited the most favorable discharge characteristics at 300 V. A comparative longitudinal analysis of the discharge phenomena for the three distinct electrode types, as depicted in **Figures 6A(iv), 6B(iv), and 6C(iv)**, indicates that the ring-needle electrode generated an orange-colored discharge material with the highest luminosity. This is followed by the needle electrode, which exhibits a comparatively lower luminosity, while the cylinder electrode demonstrates the least luminosity among the three.

Building on this foundation, a detailed analysis was conducted on the electrode ablation effects under varying voltage conditions. **Figure 7A** shows the thermal damage area and effective ablation depth from a 2-second needle electrode ablation of porcine liver tissue. As the voltage increased from 180 V to 260 V, the effective ablation depth increased rapidly from 0.4 mm to 1.7 mm. However, from 260 V to 300 V, the depth of the effective ablation area experienced a deceleration, increasing only from 1.7 mm to 2.1 mm. The effective ablation area was significantly enlarged at this voltage level, while the thermal damage to surrounding normal tissue was relatively minimized. **Figure 7B** delineates that at a voltage of 300 V, the ring-needle electrode yielded the smallest thermal damage area, measured at 0.72 mm^2, and the deepest effective ablation, reaching 1.13 mm in depth. Furthermore, **Figure 7C** illustrates that with increasing voltage, there was a corresponding reduction in the thermal damage area associated with cylinder electrode ablation, suggesting an optimization of the ablation process at higher voltages. At 180 V, the thermal damage area reached its maximum. Although the effective ablation depth increased with voltage, the increase was minimal, with a change of only 0.7 mm.

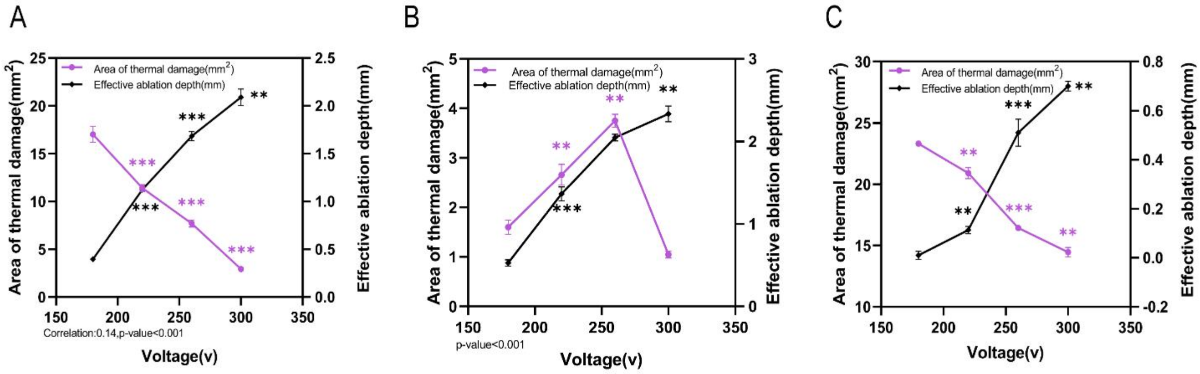


Figure 7. Curves showing the change in thermal injury area and effective ablation depth in porcine liver tissue over time for different electrodes. (A) Needle electrode; (B) Ring-needle electrode; (C) Cylinder electrode.

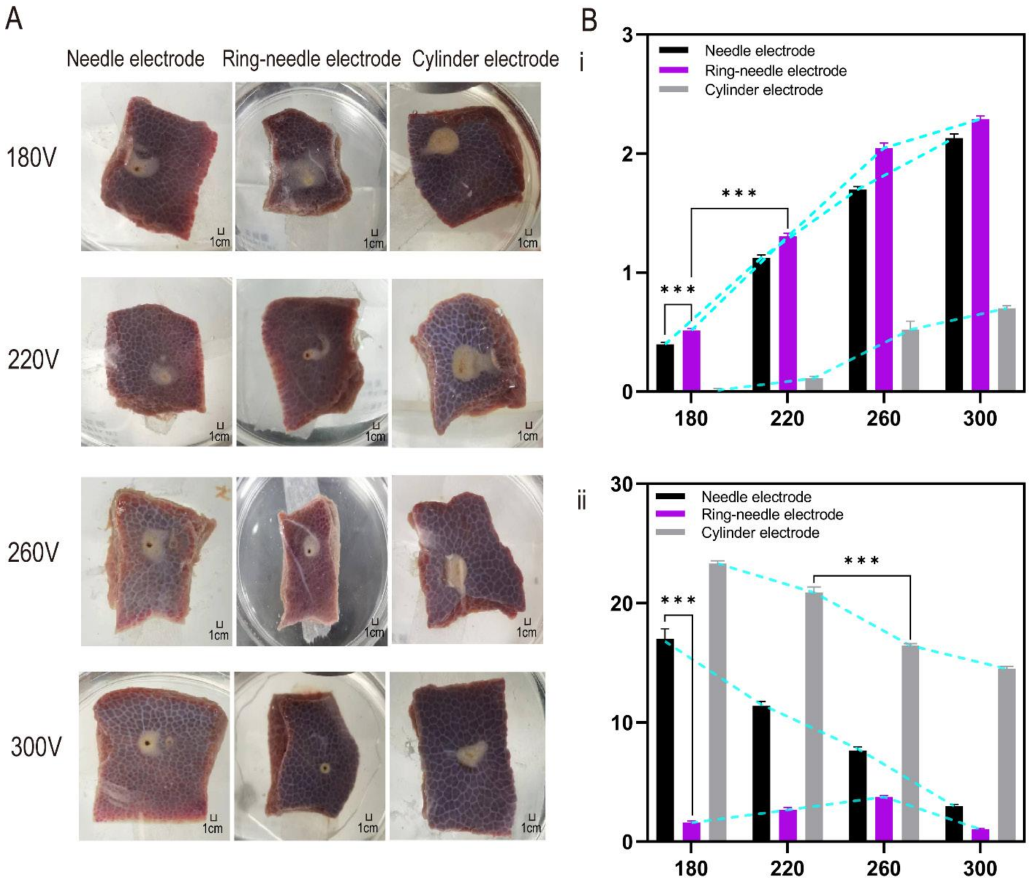


Figure 8. Voltage-dependent ablation effects of three types of electrodes. (A) Pictures of ablation effects on porcine liver tissue from three types of electrodes; (B) Curves of effective ablation depth (i) and thermal damage area (ii) for three types of electrodes.

This study systematically evaluates the ablation performance of three electrode configurations under standardized voltage conditions. Comparative analysis of ablation effects (**Figure 8A**) reveals a consistent efficacy hierarchy: ring-needle electrodes demonstrate superior performance, followed by needle electrodes, with cylinder electrodes exhibiting the lowest efficacy. Specifically, at the maximum test voltage of 300 V, the ring-needle electrode exhibited the smallest area of protein denaturation and discoloration, significantly lower than that observed with the needle and cylinder electrodes. The voltage-dependent ablation characteristics are presented in **Figure 8B**. Both ablation depth and thermal damage area exhibit non-linear relationships with applied voltage. As voltage increasing, the effective ablation depth in porcine liver tissue increased for all three electrodes, ranked from largest to smallest as follows: ring-needle type electrode, needle type electrode, and cylinder electrode (**Figure 8B(i)**). Furthermore, **Figure 8B(ii)** indicates that, with increasing voltage, the thermal damage area for both the needle and cylinder electrodes exhibited a gradual reduction. In contrast, ring-needle electrode demonstrated an initial increase in thermal damage with subsequent decrease. A comprehensive analysis of **Figures 8A-B** reveals that the ring-needle electrode results in reduced thermal damage and a larger ablation area at the same voltage.

Discussion

Low-temperature plasma ablation represents a significant advancement in tissue ablation technology, integrating cutting, hemostasis, and suction capabilities while maintaining a low-temperature condition. This methodology offers multiple clinical benefits, including reduced postoperative inflammation, minimized patient discomfort, improved surgical field visualization, and enhanced procedural efficiency. These advantages have established low-temperature plasma ablation in tonsillectomy surgery as an increasingly important tool in tonsillectomy procedures, with growing recognition of its clinical value [14]. While existing research has primarily focused on optimizing plasma layer formation and implementing rapid cooling/cleaning mechanisms through multi-channel drainage systems, critical limitations in current device design have remained unaddressed. The fixed configuration of conventional plasma scalpel heads, featuring non-replaceable electrodes with uniform emission geometries, presents significant constraints in adapting to diverse surgical requirements. In this study, three distinct electrodes configurations were designed to address the aforementioned issue.

The simulation results demonstrate distinct thermal gradients surrounding the three electrode types, with proximal temperatures reaching 70-100 °C while rapidly decreasing with distance from the electrode surface. This finding confirms the localized effect of low-temperature plasma ablation, where the elevated temperatures are confined to the immediate tissue interface while maintaining sub-lethal temperatures in surrounding regions. This temperature range not only facilitates effective ablation but also minimizes thermal injury to adjacent tissues. Through extensive research and precise calculations, significant results have been achieved in temperature distribution, and accurate simulations of the electric field distribution have been demonstrated. According to the dielectric breakdown theory of gas discharge ignition, when the electric field intensity within the vapor layer exceeds a certain critical threshold, the gaseous medium in the vapor layer undergoes electrical breakdown, leading to electron avalanche ionization and the formation of plasma [29]. The electric field intensity E within the vapor layer can be rigorously expressed as: $E = \frac{U_V}{l_V}$,where l_V is the thickness of the vapor layer, and U_V is the voltage applied across the electrodes. Plasma discharge in saline liquid media involves complex vapor-layer dynamics and localized electric field distributions. The breakdown threshold is determined based on the properties of the gas mixture. In this study, the vapor layer primarily consists of a mixture of hydrogen and water vapor, and the breakdown field strength of the vapor layer is set to an empirical value of 1*10^6 V/m. This empirical value is commonly used to describe the breakdown behavior of air or similar gases (e.g., water vapor, hydrogen, etc.), with the specific magnitude determined based on factors including the ionization energy of the gas molecules, collision cross-section, and dielectric properties of the gas mixture [30]. In the experiment, the applied voltages (180 V, 220 V, 260 V, and 300 V) were combined with the corresponding vapor layer thicknesses to calculate the resulting electric field strengths. These values were then compared to the empirical breakdown threshold ($1*10^6$ V/m) to validate whether the breakdown conditions were satisfied.

Image analysis revealed that the three electrode structures designed in this study were capable of generating bubble layer with thickness ranging from 100 to 300 μm in saline medium, as determined by finite element analysis. Within these bubble layers, the electric field strength exceeded 1*10^6 V/m, sufficient to induce electrical breakdown and ensure plasma formation, thereby establishing favorable conditions for subsequent experiments.

The discharge phenomena of cylinder electrode in the saline solution under four different voltages were not significant. This indicates that at this stage, only minimal plasma was generated, primarily causing thermal damage through heating. Notably, in sparse discharge regions, such localized heating can readily lead to tissue burns [14].

A comprehensive study was conducted on the ablation effects of electrodes with different shapes under the same voltage conditions. The results indicate that the effective ablation depth decreases as follows: ring-needle, needle, and cylinder electrodes. Bio-thermal damage is determined by the combined effect of temperature and exposure time. When the cumulative thermal dose exceeds a critical threshold, the affected tissue is deemed to have undergone thermal necrosis [31, 32]. For instance, at 300 V, the ring-needle electrode demonstrated the smallest area of protein denaturation and discoloration, indicating the least thermal damage to surrounding healthy tissue. Concurrently, as voltage increased, the thermal damage area decreased for both needle and cylinder electrodes. This trend may be attributed to enhanced electric field intensity at higher voltages, which triggers electrical breakdown and subsequent plasma generation. When employing the ring-needle electrode for ablation of porcine liver tissue, the thermal damage area initially increased and then decreased. This observation suggests that as voltage rises, both the quantity of generated plasma and its energy content also increase. Therefore, higher voltage correlates with enhanced ablation capability and improved resection efficacy. However, it should be noted that while increasing voltage can expand the effective ablation zone, exceeding a certain threshold may result in more severe thermal damage to surrounding normal tissue. Hence, the application of higher voltages in clinical practice requires cautious consideration.

This study comprehensively investigated the mechanical design, finite element simulation, and ablation experiments of three types of electrodes. However, long-term studies remain necessary to confirm their safety and efficacy, particularly regarding whether the electrodes can reliably generate plasma discharges in saline environments. If the findings are extended to low-temperature plasma ablation procedures for tonsillectomy, we can anticipate that both the needle and ring needle electrodes would demonstrate satisfactory temperature and electric field distributions to meet clinical requirements. A significant challenge lies in the performance of cylinder electrodes. Specifically, when ablating porcine liver tissue, the cylinder electrode failed to achieve effective plasma-mediated ablation, resulting in substantial thermal damage. This failure can be attributed to insufficient electric field strength, not reaching the threshold required for electrical breakdown, which inhibits plasma generation. Potential causes for this phenomenon include insufficient extension length of the cylinder electrode and/or incomplete contact between the bipolar cutting surface and the tissue. Additionally, different types of material processing require distinct plasma characteristics, which in turn affect the plasma discharge process and its characteristics [33]. To address this issue, potential solutions include optimizing electrode length, adjusting electrode geometry, and modifying the conductive properties of electrode material to promote more stable and intense plasma generation, thereby increasing plasma yield. Furthermore, the effects of different voltage levels on the discoloration area and effective ablation depth can help determine optimal voltage parameters.

Conclusion

Through finite element simulation analysis and ablation experiments involving three distinct electrode shapes, the feasibility and advantages of each electrode were validated. The findings indicate that all three electrodes can effectively achieve tissue cutting and ablation during surgical procedures while significantly minimizing thermal damage to adjacent normal tissues. Furthermore, these electrodes generate vapor layers where the electric field intensity is sufficient to induce electrical breakdown, thereby facilitating plasma formation. Experimental results demonstrate that as the voltage increases, the electrodes produce brighter and more stable plasma. Notably, at 300 V, all three electrodes generated the most stable and luminous plasma. Additionally, as the voltage increases, thermal damage to surrounding normal tissue decreases, while the effective ablation area expands. Among the tested electrodes, the ring-needle electrode demonstrated the highest ablation efficiency.

In summary, these three electrodes can effectively achieve tissue cutting and hemostasis while ensuring the safety of therapeutic outcomes. This research provides a novel approach for the design of electrodes for low-temperature plasma ablation, thereby promoting their medical applications and clinical development.

Author Contributions: Chengli Song, Lin Mao and Siqi Zhao conceived this project. Chengli Song and Siqi Zhao carried out structural design and finite element simulation. Lin Mao and Chengli Song designed an electrode ablation experiment. Siqi Zhao and Zhuotianhao Wang performed experiments and collected the relevant data and analyzed. All the authors contributed to the writing of manuscript, with Lin Mao making important contributions throughout the research and writing process.

References

[1] Harris E. Tonsillectomy Effective for People With Frequent Sore Throats. JAMA 2023;329(23):2011-2011.

[2] Wilson JA, O'Hara J, Fouweather T, et al. Conservative management versus tonsillectomy in adults with recurrent acute

tonsillitis in the UK (NATTINA): a multicentre, open-label, randomised controlled trial. Lancet 2023;401(10393):2051-2059.
[3] Kandemir S, Pamuk AE, Özel G, et al. Comparison of Three Tonsillectomy Techniques: Cold Dissection, Monopolar Electrocautery, and Coblation. Int Arch Otorhinolaryngol 2023;27(04):e694-e698.
[4] Subaşı B, Oğhan F, Taşlı H, et al. A comparison of the harmonic scalpel, coblation, bipolar, and cold knife tonsillectomy methods in adult patients. J Surg Med 2021;5(12):1198-1201.
[5] Zhong Y, Wei Y, Min N, et al. Comparative healing of swine skin following incisions with different surgical devices. Ann Transl Med 2021;9(20):1514.
[6] Dandu N, Nelson BB, Easley JT, et al. Quantifying the magnitude of local tendon injury from electrosurgical transection. J Shoulder Elbow Surg 2022;31(4):832-838.
[7] Lacitignola L, Desantis S, Izzo G, et al. Comparative Morphological Effects of Cold-Blade, Electrosurgical, and Plasma Scalpels on Dog Skin. Vet Sci 2020;7(1):8.
[8] Chang EI, Carlson GA, Vose JG, et al. Comparative Healing of Rat Fascia Following Incision with Three Surgical Instruments. J Surg Res 2011;167(1):e47-e54.
[9] Clark CM, Schubart JR, Carr MM. Trends in the management of secondary post-tonsillectomy hemorrhage in children. Int J Pediatr Otorhinolaryngol 2018;108:196-201.
[10] Haq AU, Bansal C, Pandey AK, et al. Analysis of Different Techniques of Tonsillectomy: An Insight. Indian J Otolaryngol Head Neck Surg 2022;74(3):5717-5730.
[11] Lieberg N, Aunapuu M, Arend A. Coblation tonsillectomy versus cold steel dissection tonsillectomy: a morphological study. J Laryngol Otol 2019;133(9):770-774.
[12] Karthik C, Sarngadharan SC, Thomas V. Low-Temperature Plasma Techniques in Biomedical Applications and Therapeutics: An Overview. Int J Mol Sci 2024;25(1):524.
[13] Samdani SK, Kishore J, Yogi V, et al. A Comparative Study of Cold Dissection Tonsillectomy and Harmonic Scalpel Tonsillectomy Under Microscope- Our Experience. Indian J Otolaryngol Head Neck Surg 2022;74(Suppl 3):5395-5403.
[14] Elmore L, Minissale NJ, Israel L, et al. Evaluating the Healing Potential of J-Plasma Scalpel-Created Surgical Incisions in Porcine and Rat Models. Biomedicines 2024;12(2):277.
[15] Belloso A, Chidambaram A, Morar P, et al. Coblation tonsillectomy versus dissection tonsillectomy: Postoperative hemorrhage. Laryngoscope 2003;113(11):2010-2013.
[16] Parab SR, Khan MM. Endoscope holder-assisted endoscopic coblation tonsillectomy. Eur Arch Oto-rhino-l 2020;277(11):3223-3226.
[17] Chen L, Xie L, Wu T, et al. An experimental study on low-temperature plasma tissue ablation and its thermal effect. J Phys D Appl Phys 2024;57(36):365202.
[18] Xiao A, Liu D, He D, et al. Plasma Scalpels: Devices, Diagnostics, and Applications. Biomedicines 2022;10(11):2967.
[19] Guo J, Wang Y, Liu H, et al. Thermal–Electrical Optimization of Lithium-Ion Battery Conductor Structures Under Extreme High Amperage Current. Appl Sci 2025;15(10):5338.
[20] Trieschmann J, Vialetto L, Gergs T. Review: Machine learning for advancing low-temperature plasma modeling and simulation. arXiv 2023;arXiv:2307.00131.
[21] Wiechmann EP, Morales AS, Aqueveque P, et al. Three-Dimensional FEM Thermal and Electrical Analysis of Copper Electrowining Intercell Bars. IEEE Trans Ind Appl 2017;53(1):638-644.
[22] Cicciù M. Bioengineering Methods of Analysis and Medical Devices: A Current Trends and State of the Art. Materials (Basel) 2020;13(3):797.
[23] Panchal B, Bhadauria A, Varghese S. 3D FEM Simulation and Analysis of Fractal Electrode-Based FBAR Resonator for Tetrachloroethene (PCE) Gas Detection. Fractal Fract 2022;6(9):491.
[24] Stawarz JE, Matteini L, Parashar TN, et al. Comparative Analysis of the Various Generalized Ohm's Law Terms in Magnetosheath Turbulence as Observed by Magnetospheric Multiscale. J Geophys Res Space Physics 2021;126(1):2020JA028447.
[25] Tungjitkusolmun S, Staelin ST, Haemmerich D, et al. Three-dimensional finite-element analyses for radio-frequency hepatic tumor ablation. IEEE Trans Biomed Eng 2002;49(1):3-9.
[26] Qasem NAA, Generous MM, Qureshi BA, et al. A Comprehensive Review of Saline Water Correlations and Data: Part II—Thermophysical Properties. Arab J Sci Eng 2021;46(3):1941-1979.
[27] Generous MM, Qasem NAA, Qureshi BA, et al. A Comprehensive Review of Saline Water Correlations and Data-Part I: Thermodynamic Properties. Arab J Sci Eng 2020;45(11):8817-8876.
[28] Fic A, Białecki RA, Kassab AJ. Solving Transient Nonlinear Heat Conduction Problems by Proper Orthogonal Decomposition and the Finite-Element Method. Numer Heat Tr B-Fund 2005;48(2):103-124.
[29] Babich L, Loĭko TV. Generalized Paschen's Law for Overvoltage Conditions. IEEE Trans

Plasma Sci 2016;44(12):3243-3248.
[30] Škoro N, Marić D, Malović G, et al. Electrical Breakdown in Water Vapor. Phys Rev E Stat Nonlin Soft Matter Phys 2011;84(5 Pt 2):055401.
[31] Shih T-C, Huang H-W, Wei W-C, et al. Parametric analysis of effective tissue thermal conductivity, thermal wave characteristic, and pulsatile blood flow on temperature distribution during thermal therapy. Int J Heat Mass Transfer 2014;52:113-120.
[32] Shih T-C, Horng T-L, Huang H-W, et al. Numerical analysis of coupled effects of pulsatile blood flow and thermal relaxation time during thermal therapy. Int J Heat Mass Transfer 2012;55(13):3763-3773.
[33] Schutt DJ, Haemmerich D. Sequential Activation of a Segmented Ground Pad Reduces Skin Heating During Radiofrequency Tumor Ablation: Optimization via Computational Models. IEEE Trans Biomed Eng 2008;55(7):1881-1889.

Progress in Medical Devices
ISSN: 2957-5478

Zentime

Research Article https://doi.org/10.61189/483711ymetia

Design and experimental verification of a portable multi-degree-of-freedom electric needle holder

Lin Xin, Wenjie Yu, Yangzhi Liu, Liuxiao Cheng, Zhongxin Hu, Chengli Song

School of Health Sciences and Engineering, University of Shanghai for Science and Technology, Shanghai 200093, China.

Corresponding author: Chengli Song.

Declaration of conflict of interest: None.

Received December 3, 2024; Accepted March 4, 2025; Published June 30, 2025

Highlights

- Innovative Design: A wire-driven multi-degree-of-freedom electric needle holder enhances laparoscopic suturing flexibility and stability.
- Advanced Motion Control: Two rotational joints enable yaw and pitch movements for precise needle placement in confined spaces.
- Superior Clamping Performance: With a maximum clamping force of 29.6 N an extraction force of 12.9 N, the needle holder prevents slippage.
- 28% Faster Suturing: Reduced operation time due to minimized repositioning in complex procedures.
- Clinical Potential: Validated through diverse tests, demonstrating accuracy and efficiency in minimally invasive surgery.

Abstract

Objective: To design and validate a portable multi-degree-of-freedom electric needle holder to improve operational flexibility and precision in laparoscopic surgery, while reducing the operator's workload. The design optimizes the transmission system to enable precise control of surgical instruments, addressing the challenges of complex surgical environments. **Methods:** A multi-degree-of-freedom transmission structure, driven by steel wires, was designed. It integrates motors and angle sensors to achieve multi-directional rotation and force control for the grasping forceps. The experimental section includes tests for maximum gripping force, maximum extraction force, and a porcine suturing experiment. The performance of the electric needle holder was compared with traditional needle holders in simulated laparoscopic surgery to verify its effectiveness. **Results:** Experimental results show that the maximum gripping force of the electric needle holder at various angles reached 29.6 N, significantly higher than that of traditional needle holders. The maximum extraction force for needles of different sizes is approximately 12.9 N, effectively preventing needle slippage. Additionally, the porcine suturing experiment demonstrated that the electric needle holder significantly reduced operation time and improved suturing precision in complex surgical settings. **Conclusion:** The portable multi-degree-of-freedom electric needle holder enhances operational flexibility and safety in laparoscopic surgery by increasing the rotational degrees of freedom and precise motor control. Its superior gripping force and operational efficiency suggest strong clinical potential, offering significant improvements over traditional surgical instruments.

Keywords: Needle holder, multi-degree-of-freedom, motor-driven, portable, structural design

Introduction

Laparoscopic surgery is a minimally invasive technique that involves inserting instruments through small incisions and using real-time imaging from an endoscope to guide precise operations. It is considered one of the most significant advancements in surgery in the 20th century. Traditional laparoscopic procedures typically require 3 to 5 small incisions, ranging from 5 to 20 millimeters in size. In contrast, Single-Incision Laparoscopic Surgery offers a less invasive alternative, with several potential advantages, including reduced postoperative pain, shorter recovery time, and improved cosmetic outcomes [1].

Address correspondence to: Chengli Song, School of Health Sciences and Engineering, University of Shanghai for Science and Technology, No.516 Jungong Road, Shanghai 200093, China. Tel: +86-15000513058; E-mail: csong@usst.edu.cn.

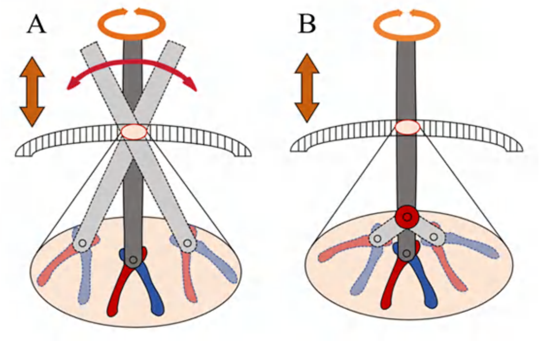


Figure 1. Comparison of Degrees of Freedom in Different Needle Holders. (A) Traditional Needle Holder; (B) Multi-Degree-of-Freedom Needle Holder.

However, Single-Incision Laparoscopic Surgery still faces technical challenges, such as instrument collisions, lack of appropriate triangulation, and difficulty in tissue retraction [2, 3].

The needle holder is a key instrument in Single-Incision Laparoscopic Surgery. As an indispensable tool during suturing, it is used to clamp and control the suture needle, enabling complex suturing operations within the confined space of the surgical field. Given the limited operating space in minimally invasive surgery, the instrument's degrees of freedom (DOF) and stability are critical for procedural success [4]. An ideal needle holder should provide high-precision control within the narrow surgical view while maintaining sufficient stability to ensure the quality of suturing. However, traditional rigid needle holders have limitations due to their restricted DOF, leading to longer surgical times and steeper learning curves [5, 6]. **Figure 1** compares the operating modes of traditional needle holders and multi-degree-of-freedom needle holders. Therefore, the development of multi-degree-of-freedom needle holders has become a primary approach to improving both surgical efficiency and the learning curve [7].

Wire-driven mechanisms, with their low inertia, lightweight design, and high precision, are well-suited for long-distance transmission in confined spaces [8-10]. Consequently, most existing multi-degree-of-freedom instruments employ wire drive systems, which can be broadly classified into two categories: rigid structures that form rotational joints through interconnected links, and flexible integrated structures that achieve bending through material deformation. Examples of the first category include the RealHand instrument developed by Novare Surgical System, the Hand-X instrument developed by Human Extension, and handheld instruments used in robotic systems such as the da Vinci surgical robot [11-13]. The second category includes instruments like the Autonomy Lapro-Angle developed by Cambridge Endo [14]. While wire-driven systems offer advantages such as being lightweight and cost-effective, the holding force applied by the operator can lead to wire deformation, requiring greater force to drive the tool's end-effector, which can result in operator fatigue [15, 16].

To address these challenges, this study proposes a novel multi-degree-of-freedom electric needle holder. The proposed needle holder features a compact wire-driven joint structure and utilizes an electric motor to provide the clamping force, enabling precise multi-degree-of-freedom control and enhancing operational comfort. This paper presents the design principles, key technologies, and functional validation results of the proposed electric needle holder.

Materials and methods

Design principles of the needle holder

Traditional needle holders typically consist of a rigid long shaft, an end clamp, and a handle, providing two DOF: axial insertion and rotation. Due to the limited DOF and confined working space, movements such as yaw and pitch rely heavily on the operator's large-arm motions, which are achieved through leverage [17]. Additionally, the rotation direction of the needle holder is opposite to that of the operator's arm movements, meaning that reaching the desired position with the instrument often requires considerable time and practice. This operating method not only increases the risk of interference between instruments but can also cause secondary damage to the patient's incision.

In certain surgical procedures, instruments must maneuver around organs, requiring the end-effector to possess a wide range of motion and large-angle directional adaptability. Traditional needle holders, limited to two DOF—axial insertion and rotation—force operators to rely on arm movements for adjustments, increasing the risk of collisions and procedural complexity. To address these challenges, the proposed needle holder integrates a front-end mechanical joint, introducing two additional DOF (yaw and pitch). This expands the total DOF to four (axial insertion, rotation, yaw, and pitch), allowing for precise needle angle adjustments in confined spaces. Compared to conventional instruments (2 DOF vs. 4 DOF), this design enhances the gripper's motion range, reduces collision risks, and improves suturing efficiency and safety by minimizing reliance on operator arm leverage.

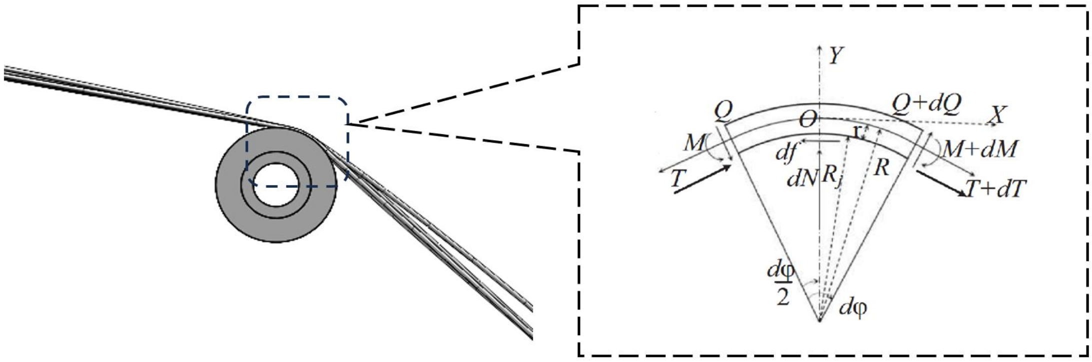


Figure 2. Force Analysis of the Infinitesimal Element of the Steel Wire.

Due to tension loss caused by the steel wire passing through the guide pulleys in a wire-driven pulley system, this loss directly affects the force and displacement transmission characteristics [18]. By conducting a force analysis on the infinitesimal elements of the steel wire, the tension transfer characteristics in a single-pulley system can serve as the foundation to derive the tension transmission characteristics throughout the entire drive system.

Figure 2 shows the force analysis diagram at the contact point between the steel wire and pulley. In the force transmission process, the steel wire possesses elasticity and bending stiffness. Therefore, when modeling, it is necessary to account for the elastic deformation, shear force, and bending moment of the steel wire. By establishing the force equilibrium equations along the X and Y axes and the force and moment equilibrium equations around point O, the following equation can be derived:

$$\begin{cases} \sum F_x = (T+dT)\cos\frac{d\varphi}{2} - T\cos\frac{d\varphi}{2} + (Q+dQ)\sin\frac{d\varphi}{2} + Q\sin\frac{d\varphi}{2} - df = 0 \\ \sum F_Y = (Q+dQ)\cos\frac{d\varphi}{2} - Q\cos\frac{d\varphi}{2} - (T+dT)\sin\frac{d\varphi}{2} - T\sin\frac{d\varphi}{2} + dN = 0 \\ \sum M_O = (M+dM) - M - (Q+dQ)\frac{R\,d\varphi}{2} - Q\frac{R\,d\varphi}{2} + r\,df = 0 \end{cases} \quad (1)$$

Where T is the tension in the steel wire; Q is the shear force in the steel wire; M is the bending moment in the steel wire; N is the normal force at the contact surface of the steel wire; f is the Coulomb frictional force at the contact surface of the steel wire; r is the radius of the steel wire; R is the sum of the radius of the guide pulley and the radius of the steel wire; and φ is the wrap angle at the contact surface of the steel wire. Since the diameter of the steel wire is typically small, its bending stiffness and radius are negligible, so it can be treated as a fine flexible rope. Therefore, $Q=M=0$ and $r=0$. When considering its nonlinear friction effects, the frictional force and normal pressure follow Amontons' law [19]:

$$f = \alpha N^n, n \le 1 \quad (2)$$

Where α and n are constants that depend on the materials of the two contacting objects.

From the above derivation, the relationship between the input tension and output tension of the steel wire between the $i-1$ pulley and the i pulley can be expressed as:

$$T_{i-1,i} = \alpha\left[\left(\frac{T_{i,i+1}}{\alpha}\right)^{\frac{1-n}{n}} + \alpha(1-n)\varphi\right]^{\frac{n}{1-n}} \quad (3)$$

From the equation (3), it can be seen that when the materials of the steel wire and pulley remain unchanged, both α and n are constants. The primary factor influencing the frictional loss of the input tension is the wrap angle of the rope on the pulley. Therefore, when designing the transmission system, the contact angle between the steel wire and the pulley should be minimized to reduce frictional losses.

Drive system design of the needle holder

In laparoscopic surgery, the movements of the needle holder primarily involve two aspects: first, the rotational movement of the needle holder within the abdominal cavity for suturing operations; and second, the application of a strong clamping force to the suture needle to prevent relative displacement between the clamp and needle during suturing. Due to the limited operating space in the laparoscopic environment, a short-range, precise driving mechanism is essential. Among various drive and connection mechanisms, the wire-driven joint mechanism occupies less space and allows for precise control. Therefore, a wire-pulley system with two joints is adopted to drive the needle holder.

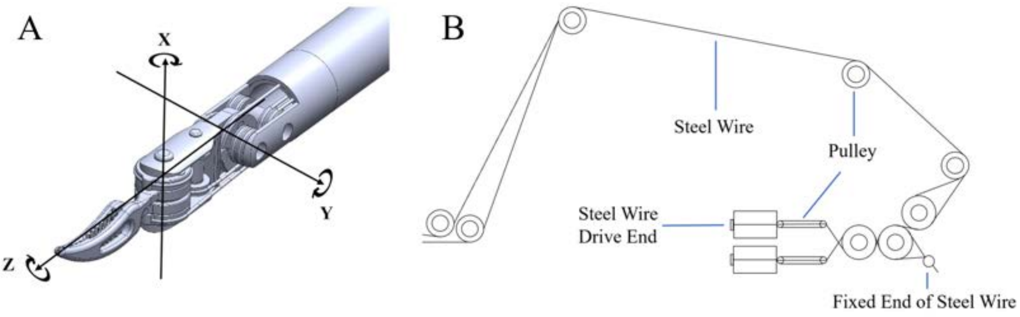


Figure 3. Needle Holder Structure. (A) Schematic Diagram of the Gripping Joint; (B) Steel Wire Drive System.

The schematic of the joint structure is shown in **Figure 3A**. Joint 1 rotates around the X-axis, while Joint 2 rotates around the Y-axis, simultaneously controlling the opening and closing of the gripper. Rotation along the Z-axis is achieved by rotating the entire instrument to prevent internal interference and wire entanglement. This mechanism provides two additional rotational DOF compared to traditional needle holders, granting the operator greater flexibility in performing laparoscopic procedures.

To align the rotational angles of the handle with the gripper and increase the handle's movement range, a long-distance driving mechanism using a specific steel wire transmission method was designed. The gripper's rotation is controlled by two joints. The first joint, controlling rotation in the XZ plane, uses two steel wires. One end of each wire is fixed, while the other end is connected to a driving pulley. When the handle rotates upward, one steel wire displaces under tension, controlling the gripper's rotation along the Y-axis. When the handle rotates downward, the other steel wire controls the reverse rotation. The second joint simultaneously controls the gripper's rotation in the YZ plane and its opening/closing motion, using two closed-loop steel wires. Each wire is fixed to both the driving and driven pulleys. The driving pulleys are fixed to the left and right grippers, enabling independent control. When the gripper rotates around the X-axis, the fixed axis of the driving pulley rotates, and the relative positions of the left and right grippers remain unchanged, rotating together.

For opening and closing the gripper, both grippers rotate, driving the rotation of the driving pulley, which in turn drives the steel wires and driven pulleys to rotate, causing the grippers to open and close. This design synchronizes the gripper's rotational angles with the operator's wrist movements, making the gripper more flexible and easier to control. **Figure 3B** presents a schematic diagram of the steel wire-pulley system structure. As the wrap angle of the pulley increases, the loss in output tension also increases. Therefore, minimizing the wrap angle in the steel wire drive system design is crucial while ensuring that the handle has sufficient movement space to guarantee the gripper's operational range. Based on these design requirements, the steel wire drive system shown in the figure uses multiple pulley sets to guide the input tension while minimizing the wrap angle to avoid exceeding 90°, thus reducing frictional losses.

Motion simulation of the needle holder

To prevent force transfer leading to steel wire deformation and insufficient clamping force, a kinematic simulation of the wire-pulley system was conducted using Automatic Dynamic Analysis of Mechanical Systems. In the simulation, components without relative displacement were treated as a single unit. Frictional and elastic contact were set between the gripper and the suture needle. A constant force ranging from 10 to 50 N was applied to the steel wire's end, and the resultant forces along the X and Y axes, as well as the resultant force, were obtained.

The simulation results showed a linear relationship between the clamping force and the operating force, with slopes of 0.016 and 0.188 for F_x and F_y respectively, and a resultant force slope of 0.19. It is evident that the system suffers significant force transmission losses. As the operator applies a larger force, the clamping force provided by the gripper decreases. This issue can lead to a significant load on the operator's hands during prolonged suturing. To address this, an electric motor was added to provide the clamping force, reducing operator hand fatigue.

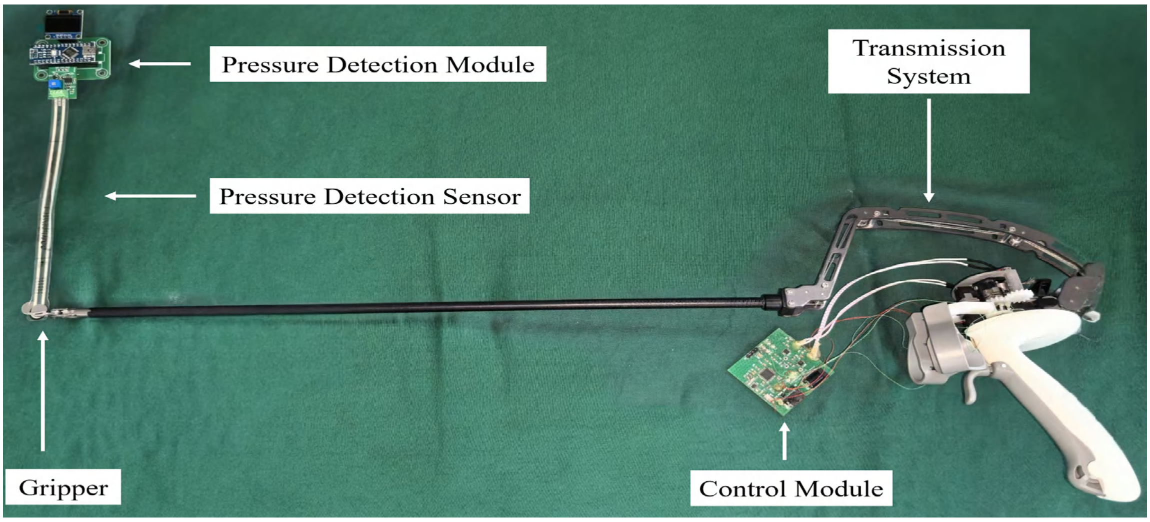


Figure 4. Maximum Clamping Force Experiment.

Needle holder circuit and control program design

A PIHER PT10 linear potentiometer was selected as the angle sensor, mounted on the handle with its central hexagonal hole matching the fixed shaft. As the handle rotates, the output resistance of the angle sensor changes, controlling the motor's forward/reverse rotation and speed, which in turn governs the movement of the gripper.

The motor drive and control are implemented through an H-bridge DC motor driver chip (DRV8832). The STM32F103CBT6 microcontroller (MCU) processes the analog signals from the angle sensor and generates PWM signals to control the motor's operation.

For the power supply design, considering that the system may share a power source with high-frequency devices such as electrosurgical tools, 220 V AC is converted to 5 V DC via a power adapter. A K7803-500 power module then converts the 5 V to 3.3 V to supply the MCU and motor driver chip.

To synchronize the rotational angle and speed of the gripper with the handle, the focus during programming was on controlling the Pulse Width Modulation (PWM) signal output from the MCU. As the resistance of the angle sensor increases, the MCU processes this analog signal, calculates the handle's rotation angle based on the slope of the resistance change, and outputs the corresponding PWM duty cycle to control the motor's forward rotation, causing the gripper to close. The motor speed is controlled by adjusting the PWM duty cycle. When the resistance of the angle sensor decreases, the MCU reverses the motor to open the gripper. Based on these functions, the circuit board design and motor control code were completed.

Experimental methodology for the needle holder

To comprehensively evaluate the performance of the multi-degree-of-freedom electric needle holder, three sets of experiments were designed: the maximum clamping force experiment, the maximum extraction force experiment, and the pork suturing experiment. Each experiment was repeated multiple times, and the data were averaged to reduce errors.

Maximum clamping force experiment

The maximum clamping force experiment aims to quantitatively assess the clamping ability of the multi-degree-of-freedom electric needle holder at different angles, providing scientific and quantitative data for its application in actual surgeries. A Flexiforce thin-film pressure sensor kit with a precision of 0.1 N was used to measure the maximum pressure exerted by the gripper on the suture needle.

Figure 4 displays the prototype of the multi-DoF electric needle holder and its maximum clamping force test setup. During the experiment, the deflection and pitch angles of the needle holder were adjusted, and measurements were taken every 10° within the range of -60° to 60°. This range covers the various operational angles the needle holder may encounter in actual surgeries, allowing for a comprehensive evaluation of its clamping performance at different postures. At each angle, the handle of the needle holder was rotated, causing the gripper to slowly close.

When the gripper completely clamps the sensor

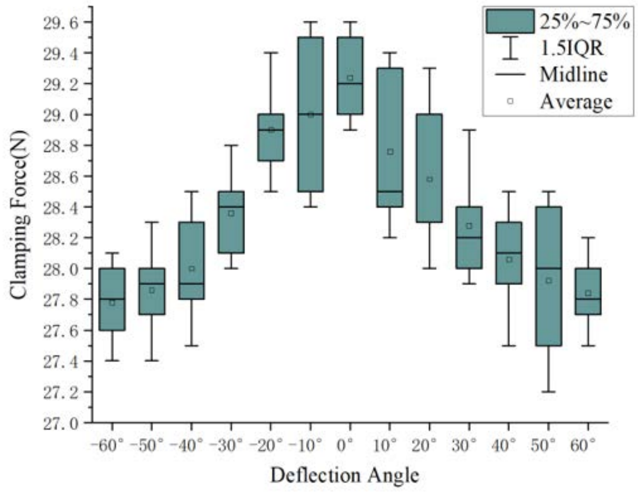


Figure 5. Clamping force at different deflection angles.

and the motor automatically stops due to the maximum clamping force being reached, the maximum clamping force displayed by the sensor was recorded.

Maximum extraction force experiment

The maximum extraction force experiment evaluates the maximum force the multi-degree-of-freedom electric needle holder can exert on the suture needle without causing it to slip out, when clamping different types of suture needles. This test indirectly measures the maximum resistance the suture needle can withstand and can be cross-validated with the maximum clamping force data, providing key insights into the needle holder's performance during actual surgeries. The experiment was conducted using an Instron 5900 tensile testing machine. The maximum force required to pull the suture needle out of the gripper, while securely clamped, was measured by gradually increasing the tensile force.

During the experiment, different types of suture needles were clamped in the needle holder's gripper to ensure the needle remained fixed after clamping. The needle holder was mounted on the tensile testing machine to maintain stability during testing. The machine was started, and the tensile force was gradually increased until the suture needle was completely pulled out, recording the maximum extraction force at this point.

Pork suturing experiment

The pork suturing experiment is designed to evaluate the usability and learning efficiency of the multi-degree-of-freedom electric needle holder during actual operations. The experiment involved two participants, both of whom had no prior experience using a needle holder. They were each tasked with performing suturing operations using both a traditional needle holder and the multi-degree-of-freedom electric needle holder. To ensure the comparability of the results, both participants received the same training before the experiment. Each participant practiced with the assigned needle holder for 10 hours to become familiar with the basic operation and feel of the tool.

After the practice session, the participants were asked to perform a real-world task using a laparoscopic surgery simulator: suturing a 5 cm length on a piece of pork and completing 10 standard surgical knots. This task simulated the suturing requirements encountered in surgery and served as an important metric for evaluating the needle holder's performance and ease of use. During the experiment, the time taken by each participant to complete the task and their subjective feedback were recorded. By comparing the performance of the two types of needle holders under the same conditions, the study assessed whether the multi-degree-of-freedom electric needle holder offered advantages in reducing operation time and improving learning efficiency.

Results

This study systematically evaluated the clamping force, extraction force, and suturing effectiveness of the multi-degree-of-freedom electric needle holder. The experimental results demonstrate that the device performs well in achieving stable clamping and improving suturing efficiency, meeting the expected design goals. The detailed analysis is as follows:

Maximum clamping force analysis

The maximum clamping force of the multi-degree-of-freedom electric needle holder varies at different angles. The experimental results show that when the gripper is at 0°, the maximum clamping force is 29.6 N. At a deflection angle of 60°, the clamping force decreases to a minimum of 27.2 N. **Figure 5** illustrates the maximum clamping force results under different gripping angles. As the deflection angle increases, the clamping force slightly decreases, likely due to changes in the lever arm of the gripper at larger deflection angles. Since the variation is small, it can be considered negligible in actu-

Table 1. Maximum Extraction Force (N) for Different Types of Suture Needles

Group	Diameter (mm) × Length (mm)				
	0.6×14	0.7×17	0.8×20	0.9×24	1.0×28
1	12.86	12.90	12.91	12.90	12.88
2	12.91	12.89	12.88	12.91	12.89
3	12.88	12.92	12.89	12.90	12.92

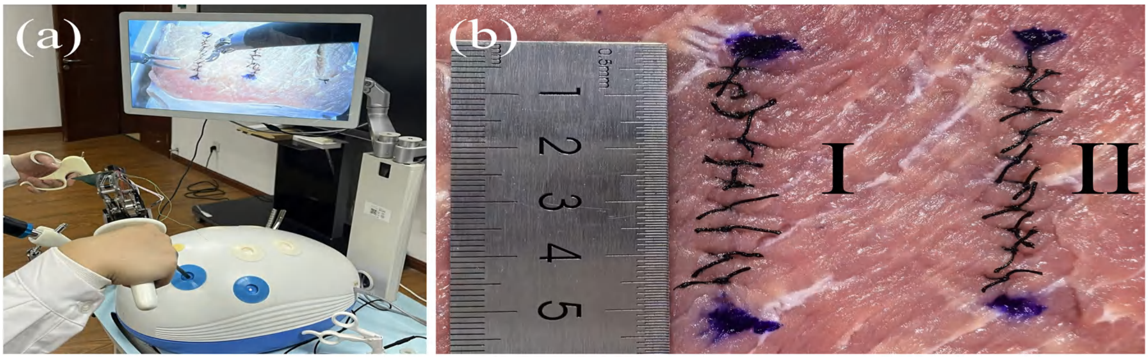


Figure 6. Pork Suturing Experiment. (a) Laparoscopic Surgery Simulation Training Scene; (b) Suturing Outcomes on Pork Tissue. I represents marks from the traditional needle holder; II represents marks from the multi-degree-of-freedom electric needle holder.

al operation.

For a conventional needle holder, the average clamping force required to securely hold the suture needle is 8.9 N [20]. In typical surgical operations, the clamping force of the gripper in the da Vinci system ranges from 7.77 to 19.94 N [21]. Therefore, the maximum clamping force of 29.6 N provided by the electric needle holder is much higher than the force needed to hold the suture needle. This indicates that the clamping force design of the electric needle holder fully meets the requirements for clamping at different angles during routine surgery, ensuring the needle's stability during suturing and reducing the risk of needle displacement or detachment.

Maximum extraction force analysis

The experimental data show that the maximum extraction force for various types of suture needles clamped by the multi-degree-of-freedom electric needle holder remains around 12.9 N. No significant difference in extraction force is observed across different types of suture needles under the same conditions. **Table 1** summarizes the maximum extraction forces when the multi-DoF electric needle holder grips suturing needles of different specifications. The extraction force is smaller than the clamping force, likely due to insufficient friction between the gripper and the suture needle. The force required for a suture needle to penetrate various tissues is typically below 2 N, which is much smaller than the extraction force of the needle holder [22]. This suggests that the extraction force of 12.9 N is more than sufficient to meet the demands of suturing during surgery, preventing the suture needle from being pulled out during the procedure. Furthermore, the uniformity of the extraction force demonstrates the needle holder's versatility in clamping different sizes of suture needles, showcasing its excellent adaptability. This is crucial for its widespread use in actual surgeries, as suture needle sizes vary across different types of procedures. The multi-size compatibility of the electric needle holder enhances its versatility and practicality.

Suturing efficiency analysis

In the pork suturing experiment, the average suturing time for the traditional needle holder group was 18 minutes, whereas the average time for the multi-degree-of-freedom electric needle holder group was reduced to 12 minutes. This significant time difference resulted in a roughly 28% improvement in suturing efficiency. **Figure 6** provides a comparative analysis of pork tissue suturing experiments. The primary reason for this improvement is the flexibility and multi-degree-of-freedom design of the electric needle holder, which allows the operator to adjust the angle of the suture needle without needing to reposition the needle holder. This significantly reduces the adjustment time during the procedure.

From the operators' subjective evaluations, the multi-degree-of-freedom electric needle holder was found to be easier to learn and more convenient to operate, effectively reducing hand fatigue. Particularly in complex suturing tasks, the time advantage of the electric needle holder was especially apparent. In situations that required frequent adjustments to the suture needle's angle, the electric needle holder reduced operational steps and increased suturing speed. Additionally, the smooth operation characteristics of the electric needle holder minimized hand fatigue, contributing to improved suturing quality and accuracy. This advantage is especially noticeable in long and complex surgeries, providing surgeons with a more comfortable operational experience and potentially increasing the surgery's success rate.

Discussion

The excellent performance of the multi-degree-of-freedom electric needle holder in terms of clamping force, withdrawal force, and operational efficiency highlights its significant advantages in meeting the demands of surgery. With a maximum clamping force of 29.6 N and a withdrawal force of 12.9 N, both values are significantly higher than those required in actual surgical procedures, ensuring stable operation under various complex conditions. In particular, the multi-degree-of-freedom design provides the electric needle holder with flexibility and adaptability that traditional needle holders cannot match, helping to shorten surgery time and improve suturing quality. Moreover, the electric needle holder demonstrated consistent withdrawal force when clamping different specifications of suturing needles, indicating high stability. This stability is crucial in clinical settings where various needle types are used, showcasing the broad applicability of the device. In practice, the instrument's versatility and adaptability will reduce the frequency of instrument changes, further enhancing overall surgical efficiency.

The experimental results show that the multi-degree-of-freedom electric needle holder can significantly improve suturing efficiency in complex surgical scenarios while maintaining operational stability and safety. Its performance in the pork suturing experiment confirms that the design goals—namely, enhancing suturing efficiency while ensuring the stability of the suture needle—have been achieved. Overall, the electric needle holder performs well, demonstrating clinical applicability, and has the potential to become an efficient tool for minimally invasive surgeries and complex suturing operations in the future.

For future clinical applications, further optimization of the grasper structure to increase friction with the suturing needle could enhance its adaptability and safety in various surgeries. Additionally, further research on the compatibility of the needle holder with other surgical instruments, such as scissors, could enable flexible multi-degree-of-freedom operation across a range of tools, improving surgical efficiency. Moreover, the integration of force sensors at the front of the instrument could allow for more precise control of the clamping force, enhancing the overall performance and safety of the device.

Conclusion

This paper presents the design of a multi-degree-of-freedom electric needle holder that can improve the efficiency of surgical training and operations while reducing surgeon fatigue. The needle holder uses a steel-wire pulley system that adapts to multi-angle operations in surgery, effectively reducing the frequency of angle adjustments and providing surgeons with flexible and stable support.

Experimental results validate the needle holder's performance: in clamping force tests, the device achieved a maximum clamping force of 29.6 N at 0°, with a minimum of 27.2 N, demonstrating the reliability and stability of the design. In withdrawal force tests, different specifications of suturing needles consistently achieved 12.9 N, showing that the device provides consistent holding ability for various needles and ensures operational safety. The pork suturing experiment demonstrated that the electric needle holder significantly reduced suturing time, improving efficiency by approximately 28% compared to the traditional needle holder. These results confirm that the design objectives have been met, demonstrating the device's wide clinical applicability and its potential to significantly improve surgical efficiency and accuracy in complex suturing scenarios, with considerable clinical value.

Author contributions: Lin Xin: Conceptualization, Methodology, Investigation, Formal analysis, Validation, Writing-Original Draft. Wenjie Yu: Control system design, Code implementation, Writing-Review&Editing. Yangzhi Liu: Experimental execution, Data curation. Liuxiao Cheng: Writing-Review&Editing, Visualization. Zhongxin Hu: Writing-Review&Editing, Resources. Chengli Song: Supervision, Project administration, Writing-Review&Editing (Corresponding Author).

References

[1] Ahmed I, Paraskeva P. A clinical review of single-incision laparoscopic surgery. Surgeon 2011;9(6):341-351.

[2] Rieder E, Martinec DV, Cassera MA, et al. A triangulating operating platform enhances bimanual performance and reduces surgical workload in single-incision laparoscopy. J Am Coll Surg 2011;212(3):378-384.

[3] Tang B, Hou S, Cuschieri SA. Ergonomics of and technologies for single-port lapaxroscopic surgery. Minim Invasive Ther Allied Technol 2012;21(1):46-54.

[4] Sato A, Hongo K, Sasaki T, et al. A novel needle holder for suturing in a deep and narrow surgical field: a technical note. J Neurosurg 2022;139(2):512-516.

[5] Dhumane PW, Diana M, Leroy J, et al. Minimally invasive single-site surgery for the digestive system: A technological review. J Minim Access Surg 2011;7(1):40-51.

[6] Kang SH, Won Y, Lee K, et al. Single-Incision Proximal Gastrectomy With Double-Flap Esophagogastrostomy Using Novel Laparoscopic Instruments. Surg Innov 2021;28(1):151-154.

[7] Jelinek F, Arkenbout EA, Henselmans PWJ, et al. Classification of Joints Used in Steerable Instruments for Minimally Invasive Surgery-A Review of the State of the Art. J Med Devices 2015;9(1):010801.

[8] Hu Y, Li W, Zhang L, et al. Designing, Prototyping, and Testing a Flexible Suturing Robot for Transanal Endoscopic Microsurgery. IEEE Rob Autom Lett 2019;4(2):1669-1675.

[9] Kim YJ, Cheng SB, Kim S, et al. A Stiffness-Adjustable Hyperredundant Manipulator Using a Variable Neutral-Line Mechanism for Minimally Invasive Surgery. IEEE Trans Rob 2014;30(2):382-395.

[10] Zheng Y, Wu BB, Chen YY, et al. Design and validation of cable-driven hyper-redundant manipulator with a closed-loop puller-follower controller. Mechatronics 2021;78:102605.

[11] Yoshiki N. Single-incision laparoscopic myomectomy: A review of the literature and available evidence. Gynecol Minim Invasive Ther 2016;5(2):54-63.

[12] 26th International Congress of the European Association for Endoscopic Surgery (EAES), London, United Kingdom, 30 May-1 June 2018: Poster Presentations. Surg Endosc 2018:483-614.

[13] Freschi C, Ferrari V, Melfi F, et al. Technical review of the da Vinci surgical telemanipulator. Int J Med Robot 2013;9(4):396-406.

[14] Kommu SS, Rané A. Devices for laparoendoscopic single-site surgery in urology. Expert Rev Med Devices 2009;6(1):95-103.

[15] Minor-Martinez A, Ordorica-Flores RM, Mota-Carmona JR, et al. Evaluating the performance of a new ergonomic laparoscopic needle holder for intracorporeal suturing. PLoS One 2025;20(1):e0313568.

[16] Faraj MK, Abed RI, Hamandi SJ, et al. Innovative surgical precision: The electronic pen needle holder based on neurophysiological principles. Surg Neurol Int 2024;15:264.

[17] Li JM, Wang SX, Wang XF, et al. Optimization of a novel mechanism for a minimally invasive surgery robot. Int J Med Robot 2010;6(1):83-90.

[18] Wang Z, Liu G, Qian S, et al. Tracking control with external force self-sensing ability based on position/force estimators and non-linear hysteresis compensation for a backdrivable cable-pulley-driven surgical robotic manipulator. Mech Mach Theory 2023;183:105259.

[19] Ramkumar SS, Rajanala R, Parameswaran S, et al. Experimental verification of failure of Amontons' law in polymeric textiles. J Appl Polym Sci 2004;91(6):3879-3885.

[20] Madhani AJ. Design of teleoperated surgical instruments for minimally invasive surgery. Massachusetts Institute of Technology 1998.

[21] Mucksavage P, Kerbl DC, Pick DL, et al. Differences in Grip Forces Among Various Robotic Instruments and da Vinci Surgical Platforms. J Endourol 2011;25(3):523-528.

[22] Bao X, Li W, Lu M, et al. Experiment study on puncture force between MIS suture needle and soft tissue. Biosurf Biotribol 2016;2(2):49-58.

Progress in Medical Devices
ISSN: 2957-5478

Zentime

Technical Review https://doi.org/10.61189/338398acvbvn

Interpretation of IEC TR 62926-2019 (medical electrical system–guidelines for safe integration and operation of adaptive external beam radiotherapy systems)

Chunying Jiao[1], Rongguo Yan[2], Yueling Li[1], Baolin Liu[2]

[1]Beijing Institute of Medical Device Testing, Beijing 101111, China. [2]School of Health Science and Engineering, University of Shanghai for Science and Technology, Shanghai 200093, China.

Corresponding author: Rongguo Yan.

Acknowledgement: This work was supported by the National Key Research and Development Program (Research and establishment of testing standards for radiation therapy equipment and new technologies 2022YFC2409502).

Declaration of conflict of interest: None.

Received March 13, 2025; Accepted April 10, 2025; Published June 30, 2025

Highlights

- This paper provides an in-depth interpretation of the international technical report IEC TR 62926-2019.
- It offers reference guidance for the development and application of adaptive radiotherapy technology in China, promoting its integration into domestic equipment.

Abstract

This paper presents an interpretation of the latest international technical report, IEC TR 62926-2019 (Medical Electrical System- Guidelines for Safe Integration and Operation of Adaptive External Beam Radiotherapy Systems) for real-time adaptive radiotherapy. It outlines the background for the development of this report, analyzes general safety guidelines for adaptive radiotherapy systems, and discusses the key design elements required for the integration of such systems. Additionally, the paper reviews and summarizes two typical reference models for adaptive external beam radiotherapy systems. The aim is to enhance the understanding and implementation of this technical report and to support its potential adaptation into a national standardized guidance document in China.

Keywords: Adaptive radiotherapy, adaptive external beam radiotherapy systems, external beam equipment, motion coordination function, motion detection equipment

Introduction

Adaptive radiotherapy (ART) is a novel radiotherapy technology evolved from image-guided radiotherapy. It enables the modification of treatment according to the anatomical and/or physiological changes (e.g. the size, the shape, and the position of tumors) in the patients throughout the course of treatment. According to Clause 3.2 of IEC TR 62926-2019, adaptive radiotherapy is defined as radiotherapy that monitors patient anatomy or physiology and, based upon the monitored information, allows changes to treatment parameters throughout the course of treatment [1].

Recent technological developments in radiotherapy have made it possible for external beam equipment (EBE) to deliver radiation doses to target volumes more accurately while better protecting surrounding critical structures. One of the major challenges in modern radiotherapy is accounting for changes in anatomical structure or physiological characteristics across different treatment fractions. For example, lung tumors may undergo translation and rotational

Address correspondence to: Rongguo Yan, School of Health Science and Engineering, University of Shanghai for Science and Technology, No.516 Jungong Road, Shanghai 200093, China. E-mail: yanrongguo@usst.edu.cn.

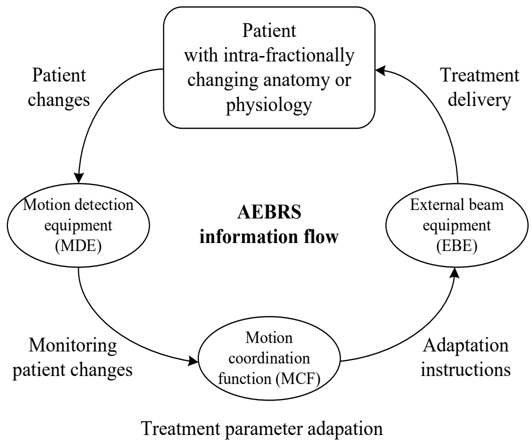


Figure 1. Concept of AEBRS with information flow. AEBRS, Adaptive External Beam Radiotherapy Systems. This figure is cited from [1].

motion, potentially leading to under-dosing of the target volume and over-dosing of nearby organs. To mitigate these risks, real-time automatic adjustment of treatment delivery is required. This can be realized through several approaches, such as: pausing the radiation beam during target volume motion, repositioning the patient by tilting treatment couch or moving the radiation head, dynamically adjusting the multileaf collimator of EBE, or modifying the scanning field in light ion beam therapy systems operating in scanning mode.

During ART, the anatomical structure or physiological characteristics of patients are monitored, allowing for real-time adjustment of treatment parameters. ART is increasingly used to ensure accurate prescription doses and optimal dose distribution despite intra-fractional changes in the target volume. Various types of motion detection equipment (MDE) are used to monitor these intra-fractional changes. Some systems rely on imaging technologies, such as X-ray, ultrasound, and magnetic resonance devices based on image-guided radiotherapy. Others use other alternative devices, including infrared sensors and optical surface tracking systems. When ART involves monitoring the target volume during treatment using MDE, effective coordination between the MDE and EBE is crucial. The Motion Coordination Function (MCF) plays a critical role in this process. It ensures that information about the target's position and shape is correctly interpreted in the context of the treatment plan, selects treatment parameters, and sends adaptive instructions to the EBE.

Figure 1 illustrates the information interaction among the MDE, EBE, MCF, and the patient [1]. During each treatment fraction, the MDE monitors changes in the patient's anatomical structure and physiology; the MCF, typically equipping prediction models, adaptive instruction generation, and validation modules, evaluates input from one or more MDEs to determine and adjust treatment parameters; these adaptive instructions are then transmitted to the EBE, such as electron accelerators, light ion beam systems, or radionuclide-based radiotherapy device; the EBE then delivers radiation to the patient based on the adjusted parameters. In summary, the integration and coordinated operation of the MDE, EBE, and MCF are essential for the safe and effective execution of ART in response to target volume variations during treatment fractions.

Currently, ART technology is implemented only in a limited number of high-end radiotherapy systems and has not been widely integrated into domestically produced radiotherapy equipment. The main reason is the inherent complexity of ART technology, which requires seamless integration of the radiotherapy machine, image guidance equipment, treatment delivery software, treatment planning software, and other supporting software and hardware components. Most domestic manufacturers remain in the research and development phase, and establishing a complete clinical workflow for adaptive radiotherapy will require further time and effort. Additionally, there are currently no relevant standards or guiding documents on adaptive radiotherapy technology in China.

Brief introduction to IEC TR 62926-2019

IEC TR 62926-2019 (Medical Electrical Systems – Guidelines for the Safe Integration and Operation of Adaptive External Beam Radiotherapy Systems for Real-Time Adaptive Radiotherapy), was formulated by the Technical Committee on Medical Electrical Equipment, Software and Systems of the International Electricity Commission (IEC TC62/SC62C), specifically under the Subcommittee for Radiotherapy, Nuclear Medicine, and Radiation Dosimetry. The technical report was released on May 20, 2019.

IEC TR 62926:2019 (hereinafter referred to as "this technical report") provides safety guidelines for integration and operation of the Adaptive External Beam Radiotherapy Systems (AEBRS). These systems are designed to manage intra-fractional motion of rigid target volumes, and their components may be sourced from one or multiple manufacturers. This technical report aims to guide stakeholders in ensuring the safe integration and operation of AEBRS for patients, operators, other personnel, and nearby sensitive equipment. This technical re-

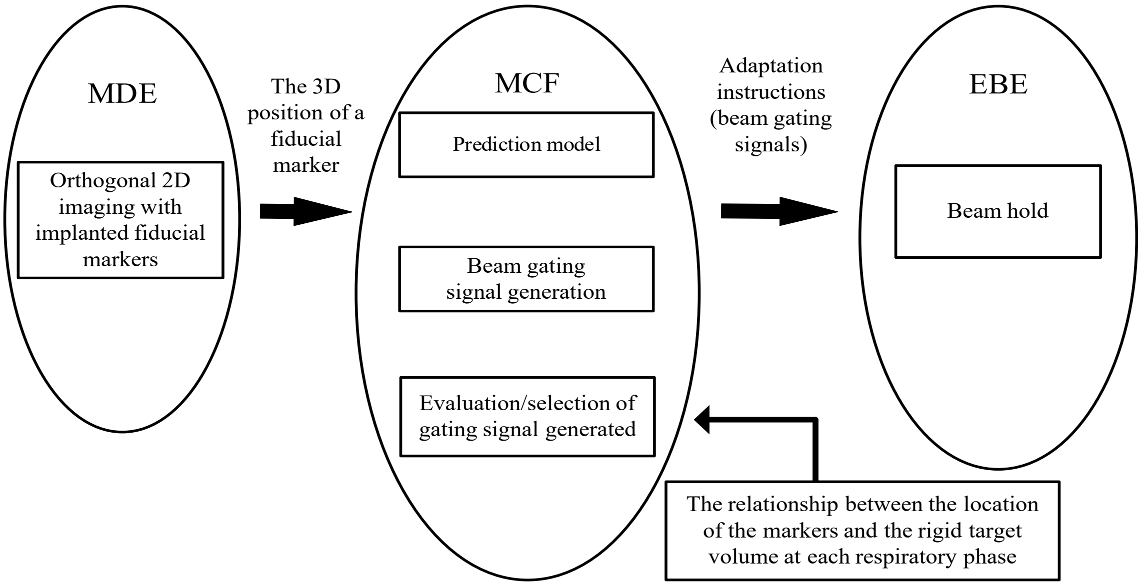


Figure 2. Beam gating system for intra-fractionally moving rigid target volumes. MDE, motion detection equipment; MCF, Motion Coordination Function; EBE, external beam equipment. This figure is cited from [1].

port outlines key safety design principles that should be followed by manufacturers or other responsible entities. In addition, it provides reference system models and highlights potential hazards that must be addressed during the risk analysis process for AEBRS dealing with intra-fractional target motion. The purpose of this technical report is to minimize the risks associated with integration and operation of medical electrical equipment and the combined use of various devices within an AEBRS.

Interpretation of main clauses in IEC TR 62926-2019

Terms and definitions

Clause 3 of this technical report gives definitions for 17 key terms commonly used throughout the document, including adaptive instruction, adaptive radiotherapy, beam gating, external beam equipment, radiation, latency, motion coordination function, motion detection equipment, prediction model, real-time adaptive radiotherapy, radiation head, and X-ray image-guided radiotherapy delay. Among them, nine items (adaptive instruction, adaptive radiotherapy, beam gating, delay, MCF, MDE, prediction model, real-time adaptive radiotherapy, and treatment parameter adaptation) are newly defined in this technical report, and other terms and definitions are derived from the existing IEC standards [1-8].

General safety guidelines for AEBRS managing intra-fractional motion rigid target volumes

Clause 4 of this technical report outlines general safety guidelines for AEBRS managing intra-fractionally moving rigid target volume (IMRTVs).

Clauses 4.2.1-4.2.2 emphasize that manufacturers should consider risk reduction when integrating various devices within an AEBRS. It is recommended that the design of these systems comply with applicable existing IEC standards. Specifically:

• IEC60601-1 (corresponding to Chinese national standard GB9706.1) and its collateral standards are generally applicable to mainstream equipment.

• Specialized safety standards should be selected according to the type of equipment. For example, the design of EBE should conform to IEC60601-2-1 (corresponding to Chinese national standard GB9706.201) or IEC60601-2-64 (corresponding to Chinese industry standard YY 9706.264); the combination of MCF and MDE should be designed in accordance with IEC60601-2-68 (corresponding to Chinese industry standard YY 9706.268); MDE shall be designed in accordance with IEC60601-2-33 (corresponding to Chinese industry standard YY9706.233) and IEC60601-2-44 (corresponding to Chinese national standard GB 9706.244).

This section also presents a reference design model for an AEBRS handling IMRTVs. The system generally consists of an EBE, a MDE and a MCF. The MDE monitors the motion of rigid target volume or equivalent structures; the MCF processes MDE data, adjusts treatment

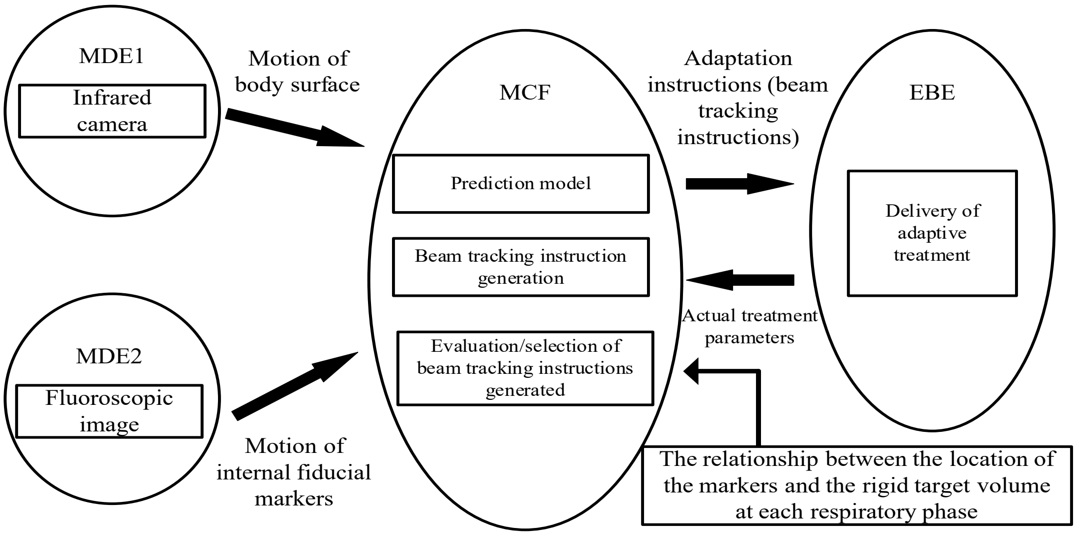


Figure 3. Beam tracking system for rigid target volumes with intra-fractional motion. MDE, motion detection equipment; MCF, Motion Coordination Function; EBE, external beam equipment. This figure is cited from [1].

parameters, and issues adaptive instructions to the EBE; and the EBE delivers radiation to the patient. The MCF can be either software or a programmable electronic medical system and should conform to the requirements of IEC 62304 (Chinese standard YY 0664) or IEC 60601-1. It is important to note that MDEs, depending on their configuration, may not necessarily fall under IEC 60601-1 requirements [9-12].

Clause 4.2.3 further provides two reference models for MCF operations in an AEBRS, namely gating model and tracking model. These models are described in **Figure 2** and **Figure 3**, respectively.

Figure 2 gives a reference model of a beam gating system designed to manage IMRTVs [1]. Before treatment, configuration parameters, describing the spatial relationship between the internal fiducial markers and the rigid target area across different respiratory phases, are preloaded into the MCF. During treatment, the MDE obtains orthogonal 2D images of the fiducial markers and calculates their 3D positions. This positional data is transmitted to the MCF, where a prediction model estimates the motion trajectory of the rigid target volume based on the input from the MDE. The predicted position is used to generate an adaptive instruction, which in this model takes the form of a beam gating signal. This signal is evaluated against a predefined positional threshold (a limit value representing the allowable 3D displacement of the reference marker). If the predicted position falls within the permissible range, a valid beam gating signal is issued to the EBE. In response, the EBE either maintains or resumes radiation delivery.

Figure 3 presents a reference model of a beam tracking system for managing IMRTVs [1]. Before treatment, configuration parameters describing the positional relationship between internal fiducial markers and the rigid target area during each respiratory phase are loaded into the MCF. The system uses two types of MDE: MDE 1 equipped with an infrared camera, monitors body surface motion to determine the respiratory phase; MDE 2, using X-ray imaging, tracks the internal fiducial markers to determine the precise position of the rigid target. Data from both MDE 1 and MDE 2 are transmitted to the MCF. A prediction model computes a correlation model of the target's position changes, including rotation and translation, based on the synchronized input from both MDEs. The output from the prediction model is used to generate an adaptive instruction, specifically a beam tracking instruction. After evaluation, this effective beam tracking instruction is sent to the EBE, which applies the appropriate therapeutic parameters to deliver treatment. Treatment parameters, such as beam direction, are sent back from the EBE to the MCF for verification.

Clause 4.3 of the technical report outlines the requirements for risk management of AEBRS. The objective of AEBRS risk analysis is to identify hazards and evaluate risks arising from the integration of MDE, MCF and EBE components. The risk management process for AEBRS should include risk analysis, risk evaluation, risk control, and acceptability evaluation of residual risks. A comprehensive risk management report must be prepared for the whole system according to the guidelines defined by ISO14971 [13-16]. Commonly identified hazards in AEBRS risk analyses include radiation hazards, electromagnetic hazards, mechanical hazards, and electrical hazards. Detailed examples of AEBRS risk management are given in Appendix B and Appendix C of this technical report.

Design elements to be considered in integrating AEBRS for the treatment of IMRTVs

Clause 5 of this technical report provides guidelines on the design elements to be considered for the safe integration of AEBRS addressing IMRTVs, which mainly includes 10 parts from Clauses 5.2.1-5.2.10.

• Clause 5.2.1 requires the manufacturer of integrated AEBRS to define the clinical performance requirements and select appropriate equipment components accordingly. The system design must ensure that irradiation is performed according to the expected clinical needs, providing a clear basis for component selection and system design.

• Clause 5.2.2 specifies the design requirements for interlocking system. It mandates that manufacturers should implement an interlocking system when uncontrolled equipment functions could lead to unacceptable risks. The description and inspection method for interlocking shall be documented in the accompanying documents. Examples of hazardous conditions requiring interlocking signals are also provided.

• Clause 5.2.3 illustrates the basic requirements for the design of AEBRS coordinate system. The coordinate system used for data input and output across all AEBRS components should be consistent to ensure that the EBE can correctly deliver irradiation. The coordinate system should conform to IEC61217 (Chinese national standard GB/T 18987) and be clearly declared in the accompanying documents [17].

• Clause 5.2.4 details communication design requirements between AEBRS devices, mainly including connection, MDE data acquisition, synchronous data handling, timing, redundancy, and data transmission protocol. The manufacturer must specify the frequency and conditions for MDE data acquisition and target position calculation before and during treatment. Data synchronization mainly requires a common timing system shared by MDE, MCF and EBE. The synchronization methods are documented. The data transmission protocol mainly requires the manufacturer to define the communication protocol for each signal transmitted to each interface, including system control command signal, status check and interlock signal.

• Clause 5.2.5 specifies the requirements for identification and management of the electrical, magnetic, and electromagnetic emission and interactions among MDE, MCF and EBE to ensure the safety of AEBRS.

• Clause 5.2.6 mandates that the specification of condition check signals for each piece of equipment, including the required frequencies, should be stated in the accompanying documents.

• Clause 5.2.7 mandates that descriptions of failure standards, the number of retry attempts, and failure notification requirements should be included in the accompanying documents.

• Clause 5.2.8 gives the requirements on delay design in AEBRS. Delay refers to the time interval between the initiation of an event and its effect, which may affect the safety of AEBRS due to intra-fractional anatomical and physiological changes. The delay design should consider two situations: during irradiation and during interruption of irradiation, including signals from the MCF to the EBE to stop irradiation and subsequent irradiation restarts.

• Clause 5.2.9-5.2.10 provide examples of typical verification items and prototype considerations for AEBRS. Typical verification items include comparing the dose distribution in the treatment plan with the dose delivered during treatment implementation [18-21]; evaluating the accuracy of prediction models for anatomical and physiological movement during treatment; testing motion detection systems, adaptive treatment parameters, and performing end-to-end verification of treatment planning and delivery.

Appendix D gives an example of a simple dynamic model.

Conclusions

IEC TR 62926-2019 provides comprehensive guidance on integrating AEBRS for real-time adaptive radiotherapy. There are two key aspects that users should pay attention to. First, although the target volumes and organs at risk may undergo deformation during movement, this technical report does not cover adaptation to target deformation. The main reason is that the real-time monitoring technology for target deformation is still under development and has not reached standardized clinical applicability. Therefore, this technical report only deals with rigid targets exhibiting intra-fractional translation and rotation; second, while this report discusses technical hazards and safety issues related to AEBRS and provides typical risk examples, it does not specify detailed clinical procedures for AEBRS. Instead, it emphasizes that responsible parties (AEBRS manufacturers or users integrating AEBRS) must not rely solely on the guidance in this report. They must also consider the actual clinical application context and incorporate specific clinical workflows and practical conditions into the design, testing, and implementation of AEBRS to ensure safety and effectiveness.

At present, there are no relevant standards and guidance documents about adaptive radiotherapy technology in China. Studying and transforming this technical report as a national standard or guidance document would provide a critical safety design and evaluation basis for both domestic and international manufacturers, promoting the development and application of ART technology in domestic equipment market.

Author contributions: Rongguo Yan contributed significantly to references and manuscript preparation; Chunying Jiao performed the interpretation of the technical report and wrote the manuscript; Yueling Li helped perform the interpretation of the technical report; Baolin Liu helped perform the outline of the manuscript.

References

[1] Medical electrical system - Guidelines for safe integration and operation of adaptive external-beam radiotherapy systems for real-time adaptive radiotherapy: IEC TR 62926-2019. International Electrotechnical Commission, 2019-05-20.

[2] Medical electrical equipment - Part 1: General requirements for basic safety and essential performance: IEC 60601-1:2005+AMD1:2012+AMD2:2020. International Electrotechnical Commission, 2020-08-20.

[3] Medical electrical equipment - Part 2-44: Particular requirements for the basic safety and essential performance of X-ray equipment for computed tomography: IEC 60601-2-44:2009+AMD1:2012+AMD2:2016. International Electrotechnical Commission, 2016-03-31.

[4] Medical electrical equipment - Glossary of defined terms. International Electrotechnical Commission: IEC TR 60788:2004. International Electrotechnical Commission, 2004-02-17.

[5] Medical electrical equipment - Part 2-1: Particular requirements for the basic safety and essential performance of electron accelerators in the range 1 MeV to 50 MeV: IEC 60601-2-1:2020. International Electrotechnical Commission, 2020-10-28.

[6] Medical electrical equipment - Part 2-68: Particular requirements for the basic safety and essential performance of X-ray-based image-guided radiotherapy equipment for use with electron accelerators, light ion beam therapy equipment and radionuclide beam therapy equipment: IEC 60601-2-68:2014+AMD1:2025. International Electrotechnical Commission, 2025-02-04.

[7] Medical electrical equipment - Part 2-33: Particular requirements for the basic safety and essential performance of magnetic resonance equipment for medical diagnosis: IEC 60601-2-33:2010+AMD1:2022. International Electrotechnical Commission, 2022-08-04.

[8] Medical electrical equipment - Part 2-64: Particular requirements for the basic safety and essential performance of light ion beam medical electrical equipment: IEC 60601-2-64:2014. International Electrotechnical Commission, 2014-09-03.

[9] Medical electrical equipment - Part 2-1: Particular requirements for the basic safety and essential performance of electron accelerators in the range 1 MeV to 50 MeV: GB 9706.201-2020. Standardization Administration of the People's Republic of China, 2020-12-24.

[10] Medical electrical equipment - Part 2-64: Particular requirements for the basic safety and essential performance of light ion beam medical electrical equipment: YY 9706.264-2022. National Medical Products Administration, 2022-05-18.

[11] Medical electrical equipment - Part 2-68: Particular requirements for the basic safety and essential performance of X-ray-based image-guided radiotherapy equipment: YY 9706.268-2022. National Medical Products Administration, 2022-05-18.

[12] Medical device software - Software life cycle processes: IEC 62304:2006+AMD1:2015. International Electrotechnical Commission, 2015-06-26.

[13] Medical devices - Application of risk management to medical devices: ISO 14971:2019. International Electrotechnical Commission, 2019-12-10.

[14] Huq MS, Fraass BA, Dunscombe PB, et al. The report of Task Group 100 of the AAPM: Application of risk analysis methods to radiation therapy quality management. Med Phys 2016;43(7):4209.

[15] Medical devices - Guidance on the application of ISO 14971: ISO/TR 24971:2020. International Organization for Standardization, 2020-06-16.
[16] Medical devices - Application of risk management to medical devices: GB/T 42061-2022. Standards press of China, 2022-10-12.
[17] Radiotherapy equipment - Coordinates, movements and scales: GB/T 18987-2015. Standardization Administration of the People's Republic of China, 2015-12-10.
[18] Standard for Medical Devices and Medical Systems - Essential Safety Requirements for Equipment Comprising the Patient-Centric Integrated Clinical Environment (ICE) - Part 1: General Requirements and Conceptual Model: ASTM F2761-13. ASTM International, 2013.
[19] Hanley J, Dresser S, Simon W, et al. AAPM Task Group 198 Report: An implementation guide for TG 142 quality assurance of medical accelerators. Med Phys 2021;48(10):e830-e885.
[20] Medical electrical equipment - Medical electron accelerators - Functional performance characteristics: IEC 60976:2007. International Electrotechnical Commission, 2007-10-16.
[21] Medical electrical equipment - Medical electron accelerators - Guidelines for functional performance characteristics: IEC 60977:2008. International Electrotechnical Commission, 2008-07-09.