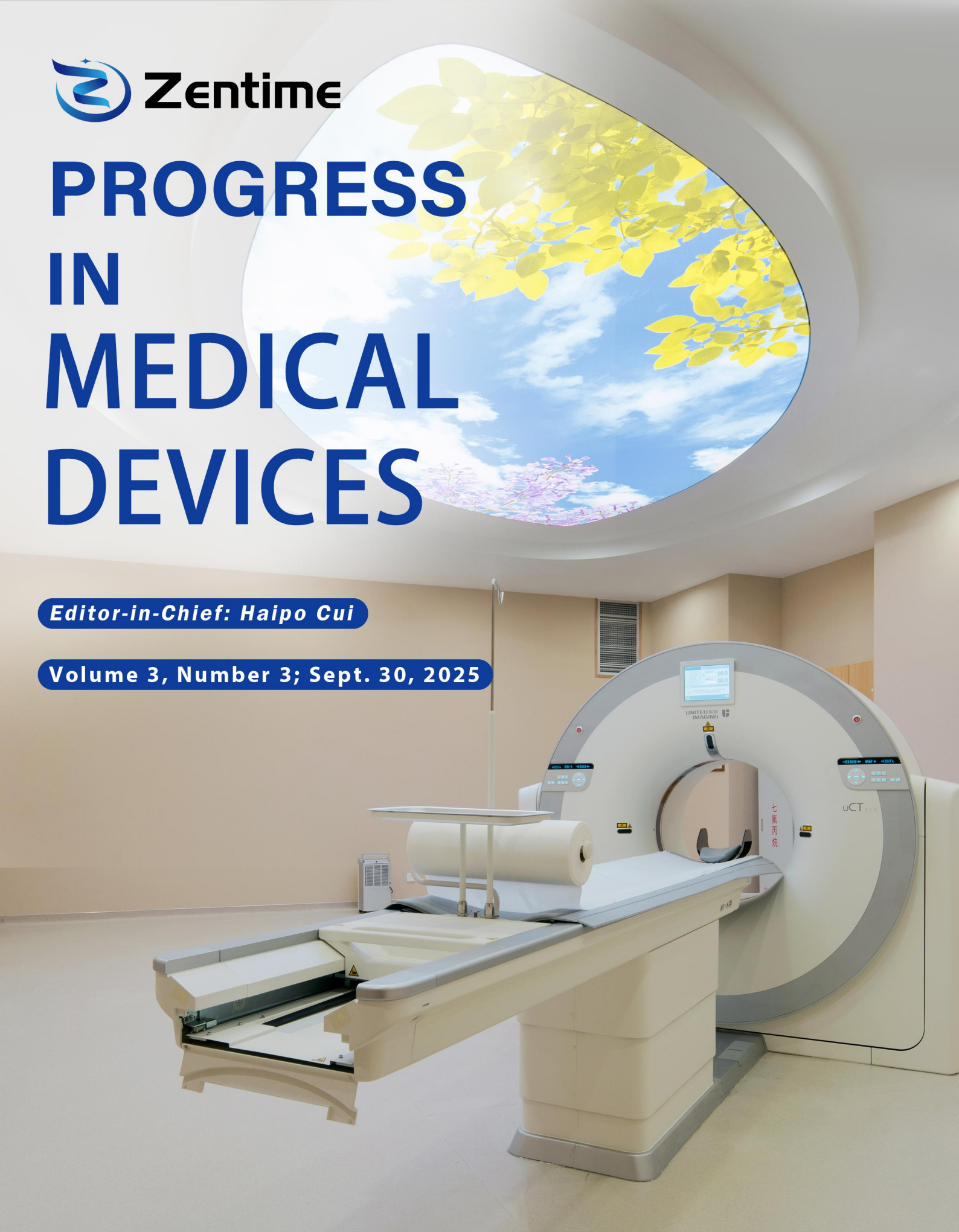
Zentime

PROGRESS IN MEDICAL DEVICES



Editor-in-Chief: Haipo Cui

Volume 3, Number 3; Sept. 30, 2025

ISSN: 2957-5478

Published by ZENTIME PUBLISHING CORPORATION

Progress in Medical Devices
ISSN: 2957-5478

Zentime

Research Article https://doi.org/10.61189/143103ncllxn

Application and accuracy enhancement of an improved spatial registration method in electromagnetic navigation for tumor ablation surgery

Xinyao Li, Gaoshuai Chen, Xin Yang, Yuchao Liu, Yu Xia

The School of Mechanical Engineering, University of Shanghai for Science and Technology, Shanghai 200093, China.

Corresponding author: Xinyao Li.

Declaration of conflict of interest: None.

Received November 15, 2024; Accepted July 11, 2025; Published September 30, 2025

Highlights

- Proposed an improved SVD algorithm to reduce errors caused by electromagnetic interference, thereby enhancing the precision of surgical navigation systems.
- Validated the algorithm through puncture experiments on biomimetic tissue of angles 0°, 5°, and 10° , demonstrating high accuracy in simulating tumor localization.

Abstract

Objective: To enhance the accuracy and efficiency of tumor ablation procedures, this study integrated tumor ablation technology with a surgical navigation system and addressed errors induced by electromagnetic interference in clinical settings. **Methods:** An improved singular value decomposition (SVD) algorithm was proposed, which iteratively traversed the point set by selecting three corresponding points at a time to compute transformation matrices. Each matrix was applied to the original dataset, and the one yielding the minimum error between the transformed and target point sets was selected as the optimal solution. This approach mitigated the influence of individual outlier points and significantly reduced navigation errors caused by electromagnetic interference, thereby enhancing registration accuracy and robustness. **Results:** Experimental results demonstrated that the improved algorithm achieved high accuracy in simulating lesion localization. The root mean square error (RMSE) was employed as a quantitative metric. RMSE values were 0.85, 0.82, and 0.75 at rotation angles of 0°, 5°, and 10°, respectively, indicating minimal deviation between the transformed and target point clouds. **Conclusion:** Compared with the conventional SVD algorithm, the improved method yielded consistently lower RMSE values, confirming its effectiveness in minimizing electromagnetic navigation errors and enhancing registration accuracy.

Keywords: Tumor Ablation, surgical navigation system, electromagnetic interference and spatial registration

Introduction

With the global aging population, cancer treatment faces increasing new challenges and complexities. As a locoregional therapeutic approach, tumor ablation has gained widespread clinical adoption due to its minimally invasive nature, cost-effectiveness, and association with faster recovery. Recent developments have seen the integration of tumor ablation techniques with surgical navigation systems, forming a novel therapeutic paradigm in oncology. Surgical navigation systems reconstruct a three-dimensional spatial model based on preoperative imaging data to accurately localize tumors during surgery. These systems enable precise visualization of the tumor and surrounding critical structures. Surgeons can preoperatively plan the optimal surgical path using dedicated software and receive real-time feedback on instrument positioning throughout the procedure. This allows for effective avoidance of vital structures and accurate targeting of lesions for ablation or excision [1].

As these systems continue to evolve, they are

Address correspondence to: Xinyao Li, The School of Mechanical Engineering, University of Shanghai for Science and Technology, No. 516 Jungong Road, Yangpu District, Shanghai 200093, China. Tel: +86-15156575887; E-mail: 15156575887@163.com.

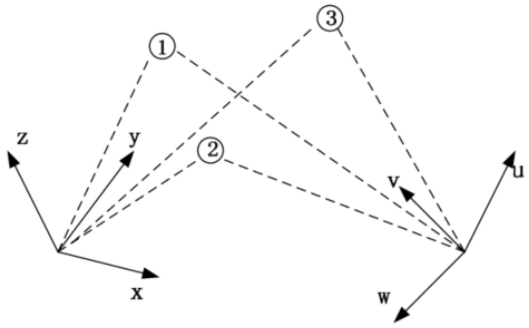


Figure 1. Schematic diagram of paired point-set registration.

increasingly used in minimally invasive procedures such as tumor thermal ablation, prostate surgery, and other laparoscopic interventions. Surgical navigation enhances anatomical clarity, optimizes intraoperative decision-making, and reduces the risk of complications. Several commercial optical navigation systems—such as the NDI Polaris, StealthStation, and Stryker ADAPT, are already in clinical use [2-4].

However, electromagnetic navigation systems remain susceptible to errors due to environmental interference [5]. These inaccuracies may lead to incomplete tumor ablation or unnecessary damage to healthy tissue, ultimately compromising therapeutic outcomes and increasing recurrence risk [6]. Clinical equipment and metallic instruments can reduce the accuracy and reliability of electromagnetic tracking systems, while limitations in real-time performance may further contribute to intraoperative errors [7, 8].

This study focuses on an improved spatial registration method based on electromagnetic navigation, which links magnetic field pose data with medical imaging to provide real-time visualization of the ablation needle's trajectory relative to the tumor. Specifically, we propose an enhanced singular value decomposition (SVD) algorithm to mitigate electromagnetic-induced errors and improve the precision and efficiency of the ablation procedure.

Surgical navigation systems are conceptually composed of two core components: medical imaging modalities and stereotactic positioning systems. The imaging module typically includes radiographic technologies such as X-ray, computed tomography, positron emission tomography, and magnetic resonance imaging. The positioning subsystem encompasses optical tracking, electromagnetic sensors, acoustic guidance, and articulated robotic arms. A critical function of these systems is spatial co-registration, which accurately aligns the patient's physical anatomy with virtual reconstructions derived from medical imaging, thereby enabling real-time visualization of surgical instruments relative to the target lesion during intervention procedures [9].

Despite advancements in medical imaging achieving submillimeter spatial resolution and stereotactic devices reaching micron-level positioning accuracy through sophisticated calibration protocols, the overall efficacy of surgical navigation platforms remains limited by deficiencies in spatial registration algorithms—particularly in multimodal image fusion and compensation for intraoperative tissue deformation. This technological bottleneck often results in registration errors exceeding 2–3 mm in clinical practice, significantly undermining the translation of theoretical precision into practical surgical accuracy.

To address registration inaccuracies, Zhou et al. employed a markerless registration method combined with spatial drift compensation techniques [10]. Li et al. applied the Coherent Point Drift algorithm based on Gaussian Mixture Models, to register abdominal point clouds in scenarios lacking prominent surface features [6]. Xiong et al. incorporated deformable self- and cross-attention mechanisms into point cloud registration, leveraging spatial local relational information as positional embeddings within a Spatial Deformable Transformer framework [11].

Currently, most surgical navigation systems rely on paired point-set registration, as illustrated in **Figure 1**. This method computes a spatial transformation that aligns two sets of points captured from different coordinate systems. It is favored for its stability, computational simplicity, and efficiency. Reference points are typically derived from external markers placed on the patient's skin or anatomical landmarks. These markers are localized in imaging data and linked to real-world coordinates through stereotactic calibration, thereby completing the co-registration between the patient's anatomy and the imaging dataset.

Paired point-set registration algorithms include the three-point method, quaternion-based approaches, and the SVD matrix decomposition method. Among them, the three-point method offers simplicity, rapid computation, and ease of implementation, requiring only three corresponding points to complete registration. However, its low accuracy and sensitivity to measurement errors limit its utility in high-pre-

cision contexts. Therefore, it is generally used for initial alignment rather than in scenarios requiring fine surgical navigation accuracy.

Materials and methods

SVD matrix factorization algorithm

In spatial transformation, the coordinates of corresponding reference points in different coordinate systems are often influenced by multiple factors and may not align precisely. Sources of error include device inaccuracies, stereotactic localization errors, and mechanical tolerances of surgical instruments. Therefore, the goal of spatial registration algorithms is to estimate the optimal transformation—specifically, the rotation matrix R and translation vector T—that minimizes the alignment error [12]. This optimization problem is defined as:

$$\mathrm{E(R,T)} = \sum_{i=1}^{m} \|q_i - (R^* p_i + T)\|^2 \quad (1)$$

In the equation, q_i represents the coordinates of the i-th control point in the target coordinate system; p_i denotes the coordinates of the i-th control point in the source coordinate system.

The transformation relationship is considered optimal, when $\mathrm{E(R,T)}$ is minimized. To reduce the impact of reference point errors, m (where m≥3) reference points are used to mitigate the influence of erroneous data on the overall result. Before solving the transformation relationship between the two-point sets, their centroids are first calculated as follows:

$$p_{\text{center}} = \frac{1}{\mathrm{m}} \sum_{i=1}^{\mathrm{m}} p_i \quad (i = 1,2,3 \ldots m) \quad (2)$$

$$q_{\text{center}} = \frac{1}{\mathrm{m}} \sum_{i=1}^{\mathrm{m}} q_i \quad (i = 1,2,3 \ldots m) \quad (3)$$

In the equation, p_{center} denotes the centroid of the control point set in the source coordinate system; q_{center} is the centroid in the target coordinate system.

Subsequently, the point sets are centered, as shown below.

$$p_i' = p_i - p_{\text{center}} \quad (4)$$

$$q_i' = q_i - q_{\text{center}} \quad (5)$$

where p'_i and q'_i are the centered coordinates of the i-th control point in the source and target coordinate system, respectively.

With this, Equation (1) simplifies to:

$$\mathrm{E(R,T)} = \sum_{i=1}^{\mathrm{m}} \|q_i' - \mathrm{R}p_i'\|^2 \quad (6)$$

From equation (6), minimizing $\mathrm{E(R,T)}$ reduces to determining the optimal rotation matrix R. To this end, the covariance matrix H is constructed:

$$\mathrm{H} = \sum_{i=1}^{\mathrm{m}} p_i' q_i'^{t} \quad (7)$$

SVD is performed on H:

$$\mathrm{H} = \mathrm{USV} \quad (8)$$

In the equation, both U and V are orthogonal matrices, and S is a diagonal matrix, with diagonal elements called singular values. This decomposition retains full covariance information and provides an important foundation for subsequent steps.

The matrix X is then defined as:

$$X = VU^t \quad (9)$$

The determinant of X is computed to detect potential inversion.

1) If $\det(\mathrm{X})>0$, then $\mathrm{R} = \mathrm{X}$.

2) If $\det(\mathrm{X})<0$, X is a reflection matrix (indicating inversion). In this case, V is adjusted:

If $\mathrm{V}=[\mathrm{v_1 v_2 v_3}]$, then let $\mathrm{V}'=[\mathrm{v_1 v_2 - v_3}]$. In this case:

$$\mathrm{X}' = \mathrm{V}'\mathrm{U}^{\mathrm{t}} \quad (10)$$

Thus, $\mathrm{R}=\mathrm{X}'$. In stereotactic navigation, however, mirror inversion does not occur; therefore, a detected $\det(\mathrm{X})<0$ suggests data inconsistencies that warrant further inspection.

Once R is obtained, the translation vector T is derived from the centroids:

$$\mathrm{T} = q_{\text{center}} - \mathrm{R}p_{\text{center}} \quad (11)$$

The R and T obtained via SVD thus represent the optimal solution, minimizing $\mathrm{E(R,T)}$.

SVD algorithm and registration improvement

SVD registration principle

To eliminate the influence of translational differences between point sets, the data should first be centered so that their centroids coincide with the origin. This preprocessing step allows the registration process to focus solely on the rotational relationship between the point sets, thereby avoiding biases introduced by translation [13].

Once the point sets are centered, the covariance matrix is computed. The covariance matrix captures the linear relationship between the two-point sets and quantifies the degree of correlation between corresponding elements [14]. It is defined as:

$$\sigma(p_i, q_i) = \frac{1}{n-1}\sum_{i=1}^{n} (p_i - \bar{p})(q_i - \bar{q}) \tag{12}$$

Where p_i and q_i represent the i-th observation points in the two-point sets, the covariance matrix Σ can be obtained as follows:

$$H = \begin{bmatrix} \omega(p_1, q_1) & \cdots & \omega(p_1, q_n) \\ \vdots & \ddots & \vdots \\ \omega(p_n, q_1) & \cdots & \omega(p_n, q_n) \end{bmatrix} \in R^{n\times n} \tag{13}$$

Each element in the covariance matrix represents the covariance between the i-th coordinate of point set p and the i-th coordinate in point set q. Specifically, the diagonal elements $\sigma(p_i, q_i)$ indicate the correlation between the i-th dimension of p and the i-th dimension of q. A value close to 0 suggests no significant correlation, a positive value indicates a positive correlation, and a negative value implies a negative correlation.

Thus, the diagonal elements represent the variance of each dimension, providing insights into the data distribution and the extent of transformation. The off-diagonal elements of the covariance matrix represent the correlation between different dimensions. Analyzing the covariance matrix helps reveal linear relationships between dimensions can and the underlying data structure. Therefore, SVD is applied to extract key information from the covariance matrix.

Performing SVD on the covariance matrix yields the matrices U、Σ、V^T, which are used to determine the optimal rotation matrix R. For the rotation matrix, the least square's objective function is:

$$\mathrm{E}^2 = \sum_{i=1}^{N} (q_i - Rp_i)^T (q_i - Rp_i) \tag{14}$$

By simplifying the above equation, we obtain:

$$\mathrm{E}^2 = \sum_{i=1}^{N} (q_i^T q_i + p_i^T p_i - 2q_i^T R p_i) \tag{15}$$

Because the first two terms are constants, minimizing E^2 is equivalent to maximizing the last term. We define:

$$F = \sum_{i=1}^{N} q_i^T R p_i \tag{16}$$

The objective then becomes finding the rotation matrix R that maximizes F. Simplifying F gives:

$$F = \mathrm{Trace}\left(\sum_{i=1}^{N} R p_i q_i^T\right) = \mathrm{Trace}\left(R \sum_{i=1}^{N} p_i q_i^T\right) = \mathrm{Trace}(RH) \tag{17}$$

In the equation, $H = \sum_{i=1}^{N} p_i q_i^T$.

Let A be a symmetric positive-definite matrix. For any orthogonal matrix B, the following holds:

$$\mathrm{Trace}(A) \geq \mathrm{Trace}(BA) \tag{18}$$

Thus, to maximize F, it is sufficient to find the rotation matrix R such that RH is a symmetric positive-definite matrix. Given that $H=U\Sigma V^T$, let $R=VU^T$.

Substituting this into the expression for RH gives:

$$RH = (VU^{\mathrm{T}})(U\Sigma V^{\mathrm{T}}) = V\Sigma V^{\mathrm{T}} \tag{19}$$

At this point, RH is a symmetric positive-definite so $R=VU^T$ maximizes F.

Improvement of SVD-based spatial registration

As discussed in Introduction, three non-collinear points are sufficient to uniquely determine a rigid transformation, including both rotation and translation [15]. If fewer than three points are available, the transformation cannot be uniquely defined.

Conversely, using more than three points introduces redundancy, potentially leading to an over-constrained system. Since the goal is to minimize the registration error $\mathrm{E(R,T)}$, selecting

Table 1. Parameters of random transformations

Coordinate Axes	Euler Angles (°)	Translation (mm)
X Axis	42.59	9.96
Y Axis	40.53	7.72
Z Axis	-32.12	3.97

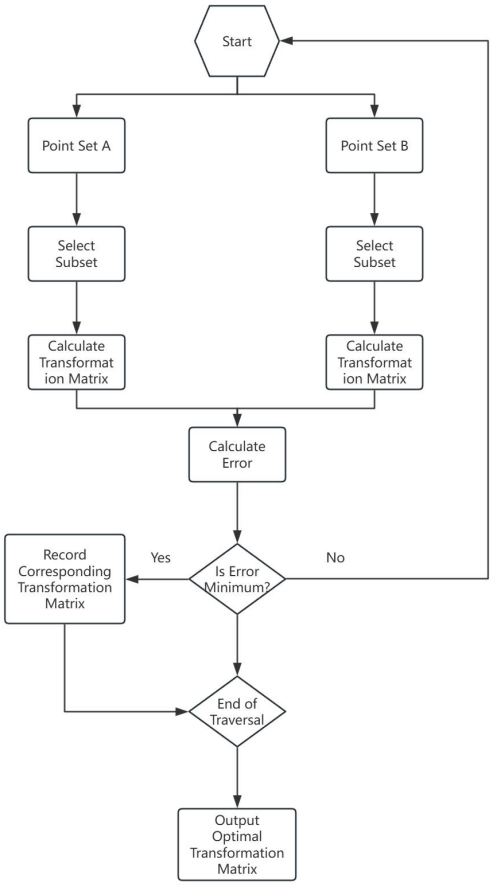


Figure 2. Flowchart of the improved SVD algorithm. SVD, singular value decomposition.

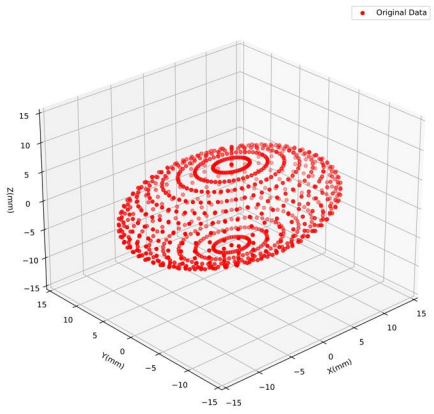


Figure 3. Three-dimensional virtual elliptical point cloud data.

three corresponding points is a practical choice to compute the transformation [16]. However, in real-world scenarios, visualization errors, noise, and electromagnetic interference may distort the point data. To reduce the influence of any single point on the transformation outcome, additional points are included in the point sets. Theoretically, with low noise levels, increasing the number of points improves the robustness and accuracy of the estimated transformation matrix [17].

In practice, however, the accuracy of the newly added points cannot be guaranteed. If points with significant errors are introduced, the overall transformation accuracy may deteriorate. To address this issue, we propose an improved spatial registration algorithm. This method iteratively selects subsets of three corresponding points from the full point set to compute candidate transformation matrices. Each matrix was applied to the original source point set to generate a transformed set, which was then compared to the target set. The transformation matrix yielding the smallest registration error was selected as the optimal solution. The overall workflow is illustrated in **Figure 2**.

Experiment and results analysis

Simulation experiment

A three-dimensional point set was first generated as the source dataset. Random rotation and translation transformations were then applied to this point set, with the corresponding parameters recorded as the ground truth. To simulate real-world conditions, random noise was added during the transformation process, producing a noisy target point set.

Both the standard SVD algorithm and the proposed improved algorithm were used to register the two-point sets. The resulting transformation matrices were used to extract Euler angles and translation vectors, which were compared against the ground truth to evaluate registration accuracy.

The synthetic point cloud was generated using scientific libraries such as NumPy and Pandas, forming a three-dimensional virtual elliptical distribution, as shown in **Figure 3**. A series of random transformations was then applied to obtain the target point cloud. The parameters of the random transformation are summarized in **Table 1**.

Table 2. Calculated random transformation parameters

Coordinate Axes	Original Transformation	SVD Algorithm	Improved Algorithm
X Axis	42.59	42.47	42.70
Y Axis	40.53	40.63	40.52
Z Axis	-32.12	-32.54	-32.38

Note: SVD, singular value decomposition.

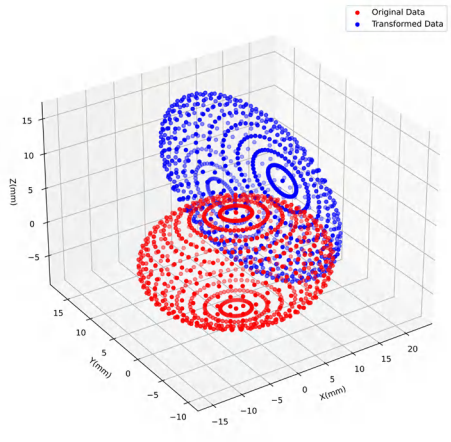


Figure 4. Comparison of point cloud data before and after transformation.

Rotation matrix and transformation matrix construction

The rotation matrix is derived from sequential rotations around the X-, Y-, and Z-axes, as defined by the following

$$\mathbf{R}_{xyz}(\alpha,\beta,\gamma) = \mathbf{R}_x(\alpha)\mathbf{R}_y(\beta)\mathbf{R}_z(\gamma) \tag{20}$$

In the formula,

$$\mathbf{R}_x(\alpha) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\alpha & -\sin\alpha & 0 \\ 0 & \sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$\mathbf{R}_y(\beta) = \begin{pmatrix} \cos\beta & 0 & \sin\beta & 0 \\ 0 & 1 & 0 & 0 \\ -\sin\beta & 0 & \cos\beta & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$\mathbf{R}_z(\gamma) = \begin{pmatrix} \cos\gamma & -\sin\gamma & 0 & 0 \\ \sin\gamma & \cos\gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

The specific expression is given by:

$$\mathbf{R}_{xyz}(\alpha,\beta,\gamma) = \begin{bmatrix} \cos\alpha\cos\gamma - \cos\beta\sin\alpha\sin\gamma & -\cos\beta\cos\gamma\sin\alpha - \cos\alpha\sin\gamma & \sin\alpha\sin\beta \\ \cos\gamma\sin\alpha + \cos\alpha\cos\beta\sin\gamma & \cos\alpha\cos\beta\cos\gamma - \sin\alpha\sin\gamma & -\cos\alpha\sin\beta \\ \sin\beta\sin\gamma & \cos\gamma\sin\beta & \cos\beta \end{bmatrix} \tag{21}$$

Substituting the data, the rotation matrix is solved as:

$$\mathbf{R}_{xyz}(\alpha,\beta,\gamma) = \begin{bmatrix} 0.64372277 & 0.76390755 & 0.04545592 \\ -0.40408034 & 0.38974887 & -0.82753544 \\ -0.64987696 & 0.51433556 & 0.55957025 \end{bmatrix} \tag{22}$$

The translation vector is constructed based on the translation distances:

$$t = [9.9572333 \quad 7.72319645 \quad 3.97196656]^t \tag{23}$$

By combining the rotation matrix and the translation vector, the 4×4 transformation matrix T is formed:

$$T = \begin{bmatrix} R & t \\ 0 & 1 \end{bmatrix} \tag{24}$$

By substituting the data:

$$T = \begin{bmatrix} 0.64372277 & 0.76390755 & 0.04545592 & 9.9572333 \\ -0.40408034 & 0.38974887 & -0.827535440 & 7.72319645 \\ -0.64987696 & 0.51433556 & 0.55957025 & 3.97196656 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{25}$$

Following established practices in registration and point cloud matching research, Gaussian noise was introduced during coordinate transformation to simulate real-world uncertainties, such as electromagnetic interference and sensor measurement errors [18-20]. Specifically, the transformation matrix was applied to the original point cloud data, and random noise was added during this process to mimic real-world conditions, generating a new set of point cloud data. This data is visualized in **Figure 4**, where blue points represent the new point cloud data.

Both the standard SVD algorithm and the improved version were applied to the original and noisy point clouds. Euler angles and translation vectors were computed from the resulting transformation matrices and compared with the ground truth parameters (**Table 2**).

The translation errors were minimal and are therefore omitted. Squared differences for each point set are shown in **Figure 5**.

As shown, both algorithms achieved high spatial transformation accuracy. However, in practical applications—such as surgeries involving

Table 3. Transformation parameters after adding high-error points

Coordinate Axes	Original Transformation	SVD Algorithm	Improved Algorithm
X Axis	42.59	39.97	42.82
Y Axis	40.53	43.64	40.03
Z Axis	-32.12	-34.17	-36.93

Note: SVD, singular value decomposition.

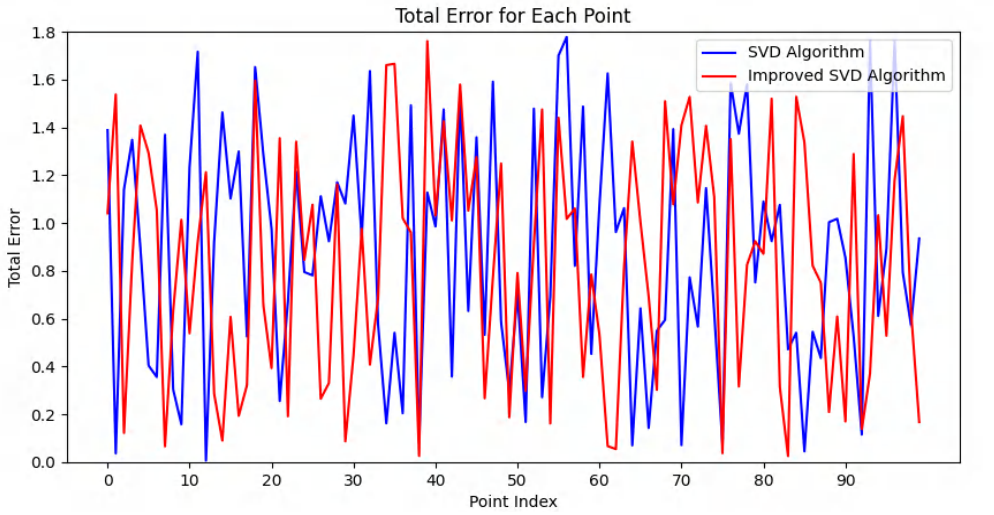


Figure 5. Squared differences for each point set.

electromagnetic tracking—metallic interference could introduce substantial error to individual data points. To simulate this, controlled perturbations were introduced into the dataset.

In such cases, traditional SVD cannot identify or exclude outliers, as it gives equal weight to all correspondences. Consequently, registration accuracy declines when outliers are present.

The improved algorithm addresses this limitation by iteratively selecting subsets of three corresponding points and identifying the transformation matrix that minimizes registration error. This approach effectively reduces the influence of erroneous data points.

The experiment is repeated with a small number of deliberately introduced high-error points. The performance of both algorithms is shown in the **Table 3**. The average squared differences for each method are shown in **Figure 6**.

The results clearly demonstrated that the conventional SVD algorithm is vulnerable to outliers. Because it incorporates all point correspondences—including those with large deviations—into its computation, it suffers from decreased registration accuracy.

In contrast, the improved SVD algorithm employed a robust outlier rejection mechanism prior to transformation estimation. By selectively retaining geometrically consistent point pairs through iterative consistency evaluation, the algorithm enhanced registration precision and minimizes the impact of noisy or erroneous data points.

Physical model experiment

In the physical model experiment, two coordinate systems were established: the electromagnetic coordinate system of external fiducial markers and their corresponding 3D coordinates in the medical image space. Specifically, electromagnetic sensors are affixed to the patient's skin as fiducial points. CT scanning was then performed to acquire medical images containing these markers. Simultaneously, the electromagnetic navigation system continuously tracked the pose of each fiducial point in real time.

First, the coordinates of the fiducial points in both systems were registered to derive a transformation matrix between the electromagnetic and medical image coordinate systems. Subsequently, an electromagnetic sensor was mounted on the surgical instrument. Using the transformation matrix, the instrument's position was displayed in the medical image coordinate space, thereby enabling real-time image-guided navigation to the target lesion.

Since the physical model lacks CT data, its dimensions and lesion coordinates were predefined. Based on the principles of im-

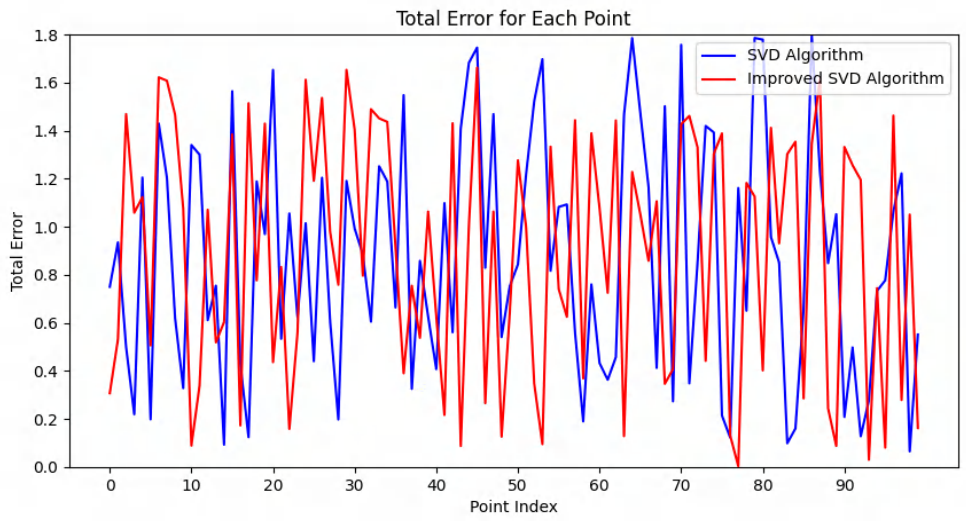


Figure 6. Squared differences after adding high-error points.

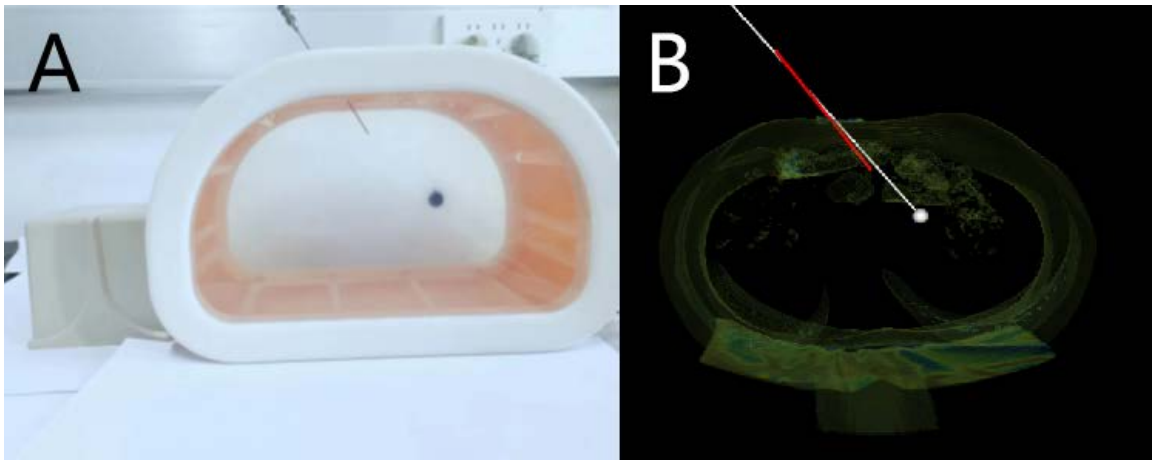


Figure 7. Navigation system integrated with physical model. (A) Navigation System Overview; (B) Integration with Phantom.

age-based coordinates, a virtual image space is constructed. Fiducial sensors were positioned at designated locations, and their coordinates in the virtual environment were obtained. Given the known location of the lesion in the physical model, its image-space coordinates can also be derived.

By aligning the electromagnetic and image coordinates, the system guided the ablation needle to the target, as illustrated in **Figure 7**. The white lines indicate the planned navigation path, and the red lines show the real-time position of the ablation needle.

An electromagnetic sensor was affixed to the tip of the ablation needle to capture its real-time pose. A local coordinate system was established with the sensor's center as the origin. Using the known geometry and mounting location of the ablation needle, the needle tip's position in the local coordinate system was calculated. This is then transformed into the electromagnetic coordinate system. Applying the precomputed transformation matrix, the needle tip's position was finally represented in the medical image space. Since the coordinates of the lesion are known, the needle tip can be precisely guided into the target under real-time image navigation.

To validate the feasibility of the proposed algorithm, a series of physical experiments were conducted. In the first experiment, six sensors were placed on a simulated tissue to mimic the location of potential model to emulate lesions. Guided by the navigation software interface, a biopsy needle was inserted into the model based on system instructions.

During the procedure, the clinician adjusted the needle in real time via the software interface, ensuring that the virtual needle and its tip aligned with the preplanned puncture path.

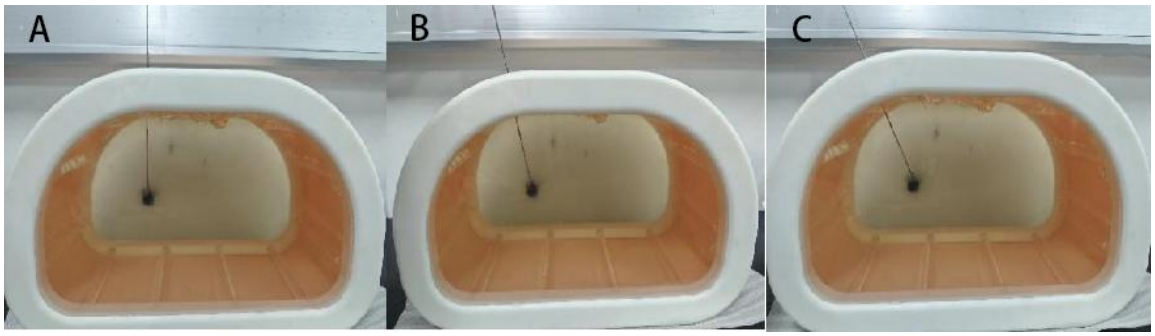


Figure 8. Needle insertion into simulated tissue. (A) Insertion at 0°; (B) Insertion at 5°; (C) Insertion at 10°.

Table 4. Root mean square error for needle insertions at different angles using the improved SVD algorithm

	0	5	10
Maximum Value	1.21	1.2	0.98
Standard Deviation	1.01	1.03	0.73

Note: SVD, singular value decomposition.

Table 5. Root mean square error for needle insertions at different angles using the traditional SVD algorithm

	0	5	10
Maximum Value	1.38	1.25	1.19
Standard Deviation	1.22	1.19	0.84

Note: SVD, singular value decomposition.

As shown in **Figure 8**, needle insertions were performed at three angles: 0°, 5°, and 10°. To minimize electromagnetic interference, all potential sources were kept at least 2 meters away from the setup (FG ≤ 2 m).

First, a target point was defined within the simulated tissue. The navigation system was used to identify its coordinates, and the relative positioning between the needle tip and the target was calculated. The path was then planned and executed accordingly. Once the needle tip reached the target, its final position was recorded using the navigation system. These data were used to evaluate localization accuracy. All participating clinicians had prior experience operating the navigation system.

To assess the precision of the experiment, the standard deviation of multiple insertions was calculated:

$$SD = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n} |p_i p|^2} \tag{26}$$

In the formula, p_i represents the needle tip position in the i-th insertion, and the centroid position of the point cloud formed by the needle tip's position in the n-th insertion.

In this experiment, the accuracy was validated by measuring the root mean square error between the actual needle tip positions and the target lesion location. The formula is as follows:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} | p_{\text{obs},i} p_{\text{tru},i} |^2}{n}} \tag{27}$$

In the formula, $p_{\text{obs},i}$ represents the actual needle tip position in the i-th insertion, and ptru, $p_{\text{tru},i}$ denotes the target lesion position for the i-th insertion.

The experimental results are presented in **Table 4**. Each needle insertion angle (0°, 5°, and 10°) was tested 100 times. For every five insertions, the centroid of the resulting point cloud was calculated, and its root mean square error relative to the target lesion location was computed.

Similarly, using the traditional SVD algorithm, needle insertions at angles of 0°, 5°, and 10° were repeated 100 times into the same lesion. The centroid of the point cloud formed by the actual needle tip positions from every five insertions was calculated. The root mean square error between each centroid and the target lesion position was then computed (**Table 5**).

The comparison results indicate that the improved SVD algorithm consistently achieved lower maximum errors and smaller standard deviations across all insertion angles. These

findings highlight the enhanced accuracy and stability of the improved algorithm compared to the traditional approach.

Notably, at angles of 10° and 15°, the reductions in both maximum error and variability were particularly substantial, suggesting that the improved algorithm can more effectively limit deviation in needle trajectory. In clinical scenarios, this capability translates into more accurate targeting of the lesion, improved surgical success rates, and enhanced postoperative recovery for patients.

Conclusion

In conclusion, the improved SVD algorithm demonstrates not only theoretical advantages but also excellent performance in practical applications. By minimizing registration errors caused by variations in needle insertion angles, the algorithm significantly improves the precision and safety of tumor ablation procedures. Its high clinical applicability and potential for widespread implementation are evident.

Future work should focus on further optimizing the algorithm by incorporating diverse clinical datasets and applying it to various tumor types, thereby strengthening its robustness and broadening its scope of application in interventional oncology.

Author contributions: This study was conceived and co-designed by Xinyao Li and Gaoshuai Chen. Xinyao Li, as the project leader and corresponding author, was responsible for determining the research protocol, supervising project progress, acquiring funding, writing the original draft, and reviewing and finalizing the manuscript. Gaoshuai Chen contributed to the development of the core research ideas and was primarily responsible for developing analytical tools, data analysis, and results visualization. Xin Yang provided key contributions in terms of theoretical guidance and in-depth analysis of the results. Yuchao Liu was responsible for data collection, curation, and validation. Yu Xia provided experimental resources, technical support for the research, and contributed to discussion of findings. All authors have read and approved the final manuscript.

References

[1] Hering A, Hansen L, Mok TCW, et al. Learn2Reg: Comprehensive Multi-Task Medical Image Registration Challenge, Dataset and Evaluation in the Era of Deep Learning. IEEE Trans Med Imaging 2023;42(3):697-712.

[2] Takai H, Murayama M, Kii S, et al. Accuracy analysis of computer-assisted surgery for femoral trochanteric fracture using a fluoroscopic navigation system: Stryker ADAPT® system. Injury 2018;49(6):1149-1154.

[3] Takeba J, Umakoshi K, Kikuchi S, et al. Accuracy of screw fixation using the O-arm® and StealthStation® navigation system for unstable pelvic ring fractures. Eur J Orthop Surg Traumatol 2018;28(3):431-438.

[4] Jackson P, Simon R, Linte C. Surgical Tracking, Registration, and Navigation Characterization for Image-guided Renal Interventions. Annu Int Conf IEEE Eng Med Biol Soc 2020;2020:5081-5084.

[5] Sun M, Rao L, Zhang C, et al. Analysis of influence of surgical instruments on accuracy of magnetic navigation system for craniofacial surgery robots. Comput Assist Surg (Abingdon) 2023;28(1):2210744.

[6] Li J, Deng Z, Shen N, et al. A fully automatic surgical registration method for percutaneous abdominal puncture surgical navigation. Comput Biol Med 2021;136:104663.

[7] Wagner A, Schicho K, Birkfellner W, et al. Quantitative analysis of factors affecting intraoperative precision and stability of optoelectronic and electromagnetic tracking systems. Med Phys 2002;29(5):905-912.

[8] Stevens F, Conditt MA, Kulkarni N, et al. Minimizing electromagnetic interference from surgical instruments on electromagnetic surgical navigation. Clin Orthop Relat Res 2010;468(8):2244-2250.

[9] Lian C, Li X, Kong L, et al. CoCycleReg: Collaborative cycle-consistency method for multi-modal medical image registration. Neurocomputing 2022;500:799-808.

[10] Zhou Z, Yang Z, Jiang S, et al. Augmented reality surgical navigation system based on the spatial drift compensation method for glioma resection surgery. Med Phys 2022;49(6):3963-3979.

[11] Xiong F, Kong Y, Xie S, et al. Spatial deformable transformer for 3D point cloud registration. Sci Rep 2024;14(1):5560.

[12] Agostinho S, Ošep A, Bue AD, et al. (Just) A Spoonful of Refinements Helps the Registration Error Go Down. 2021 IEEE/CVF International Conference on Computer Vision (ICCV) 2021;6088-6097.

[13] Li Z, Wang C, Ma J, et al. Point set registration via rigid transformation consensus. Comput Electr Eng 2022;101:108098.

[14] Chatrasingh M, Wiratkapun C, Suthakorn J. A generalized closed-form solution for 3D registration of two-point sets under isotropic and anisotropic scaling. Results Phys

2023;51:106746.

[15] Horn B. Closed-form solution of absolute orientation using unit quaternions. J Opt Soc Am A 1987;4:629-642.

[16] Wang GH, Liu T, Tang ZW, et al. Model assembly error analysis method based on local feature registration. Aeronaut Manuf Technol 2021;(03):33-6+41.

[17] Huang X, Mei G, Zhang J. Feature-Metric Registration: A Fast Semi-Supervised Approach for Robust Point Cloud Registration Without Correspondences. 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) 2020;11363-11371.

[18] Ma JY, Zhao J, Tian JW, et al. Robust point matching via vector field consensus. IEEE Trans Image Process 2014;23(4):1706-1721.

[19] Charles RQ, Su H, Kaichun M, et al. PointNet: Deep Learning on Point Sets for 3D Classification and Segmentation. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) 2017;77-85.

[20] Besl PJ, McKay ND. A method for registration of 3-D shapes. IEEE Trans Pattern Anal Mach Intell 1992;14(2):239-256.

Progress in Medical Devices
ISSN: 2957-5478

Zentime

Research Article https://doi.org/10.61189/063815kuibzu

Design and verification of the testing device for thoracic aortic stent grafts

Yu Zhou[1], Shiju Yan[1], Ailing Zhang[2]

[1]School of Health Science and Engineering, University of Shanghai for Science and Technology, Shanghai 200082, China. [2]College of Health Management, Shanghai Jian Qiao University, Shanghai 201306, China.

Corresponding author: Ailing Zhang.

Declaration of conflict of interest: None.

Received January 22, 2025; Accepted March 13, 2025; Published September 30, 2025

Highlights

- A novel testing device for measuring the radial support force and bending spring-back force of stent grafts has been developed.
- A theoretical geometric model for the bending behavior of stent grafts was derived and applied in numerical simulations, significantly enhancing the simulation results.
- The device's reliability was validated by comparing numerical simulation outcomes with physical experimental results under varying compression diameters and bending angles.

Abstract

Objective: To design a testing device for measuring the radial support force and bending spring back force of stent grafts and evaluate its effectiveness. **Methods:** A radial force and spring-back force testing device was designed to integrate with a tensile testing machine. The radial compression and bending characteristics of stent grafts for thoracic aorta applications were analyzed, and the corresponding conversion formulas were derived. A custom stent ring fixture was fabricated, and a five-wave gradient stent was sewn. Both physical experiments and finite element simulations were conducted. Radial support forces were measured by gripping 20%, 40%, and 60% of the stent's diameter, and bending tests were performed at angles of 60°, 90°, and 180°. The stability of the testing device was analyzed through comparative tests across different compression diameters and bending angles. **Results:** The device demonstrated high detection precision, stability, and accuracy, with minimal deviation across multiple measurements. The mechanical behavior of the stent observed in both finite element simulations and physical experiments showed consistent results. **Conclusions:** The testing device developed in this study effectively measures the mechanical changes in large-diameter stent grafts, providing a new reference for testing large-diameter stents.

Keywords: Stent grafts, testing device, bending spring back force, radial support force

I Introduction

Since Dake et al. first successfully applied thoracic endovascular aortic repair (TEVAR) to treat thoracic aortic aneurysms, the technique has continuously evolved [1]. Currently, large-diameter self-expanding stents are commonly used for treating such conditions. Once implanted, these stents effectively expand the vessel and block abnormal blood flow. In addition to biocompatibility, the mechanical properties of stent grafts is closley associated with the effectiveness of TEVAR and the occurrence of postoperative complications [2].

Radial support force and spring-back force are key indicators for evaluating the mechanical performance of stents [3-5]. Insufficient radial support force can result in poor fixation of the stent graft within the vessel, leading to displacement or endoleaks due to blood flow [6]. Conversely, excessive radial support force may damage the vessel intima [7]. Spring-back force refers to the inherent tendency of self-expanding endografts to return to their original shape when passively bent, such as when positioned across the aortic arch [8]. When spring-back force is too high, the stent graft's apposition

Address correspondence to: Ailing Zhang, College of Health Management, Shanghai Jian Qiao University, NO. 1111 Hucheng Ring Road, Pudong New Area, Shanghai 201306, China, Tel: +86-18616022898. E-mail: zhangailing@gench.edu.cn.

to the vessel wall is compromised, potentially causing the "bird beak phenomenon" and endoleaks [9]. Moreover, excessive stress at the contact points between the stent and the vessel wall can lead to intimal tearing, resulting in the formation of new entry tears [10, 11]. To address these issues, prior studies have focused on enhancing stent support and compliance through modifications in stent materials, peak designs, or by adding connecting struts [12-14]. However, assessing the effectiveness of these modifications requires the development of reliable and scientific testing devices.

Radial support force is commonly measured using methods such as planar compression, V-groove compression, and iris rolling compression [15]. In planar compression, parallel pressing plates are moved vertically to compress the stent, yielding a single radial support force value, but its reference value is limited [16, 17]. The V-groove compression method improves upon planar compression, by subjecting the stent to three-directional forces during compression [18-20]. However, different stent outer diameters require corresponding V-groove sizes, limiting the equipment's applicability. The iris rolling compression method applies radial compression using fan-shaped blocks arranged circumferentially. As the number of fan-shaped blocks increases, the contact surface approximates a circular shape, providing a more accurate reflection of the stent's radial support performance [21]. However, the complex structure of the equipment makes it difficult to avoid interference and friction among various mechanisms during production, resulting in high manufacturing and processing difficulty.

Methods such as cantilever beam bending, three-point bending, and four-point bending are commonly used to measure the spring-back force of stents. In the cantilever beam method, one end of the stent is fixed, and deformation force is applied to the other end, causing the stent to bend [20]. This method avoids the influence of stent movement on the spring-back force test and allows for continuous measurement at different angles. However, the direction of the dynamometer may not be perpendicular to the testing interface, and the result is a component force. In 2000, Ormiston et al. suspended a vascular stent with a steel hook and conducted a three-point bending experiment [22]. While the loading method was simple, it caused uneven bending of the stent. The four-point bending method places the sample on two loading points and applies pressure on two points above them, ensuring even distribution of the bending moment and yielding more accurate results [23]. However, its complex structure limits its application.

The thoracic aortic stent graft, a type of large vascular stent, poses significant challenges in measuring its radial mechanical properties due to its cylindrical structure. During bending, the spring-back force of the stent changes continuously along the vertical direction of the bending arc, making accurate measurement difficult. Based on these two key characteristics and current advances in radial support force and spring-back force detection, this study proposes a new testing device that is simple to manufacture and can be used in conjunction with a tensile testing machine. The feasibility of this device is verified through physical experiments and finite element simulations, and the reliability of the test results is analyzed.

II Materials and methods

Conversion formulas

The testing device for the stent graft is used in conjunction with the tensile testing machine. Through this device, the tensile force from the tensile testing machine is converted into radial compression and bending forces for the stent graft. This allows for the precise force measurement system of the tensile testing machine to be utilized for detecting the radial support force and spring-back force of the stent.

Compression of the stent graft requires the application of circumferential pressure. By winding silk threads around the stent and applying tensile force, the tensile force is converted into circumferential pressure, causing the ring of the stent graft to compress [24]. The relationship between the radial force from the silk threads and the circumferential tension borne by the stent is given by:

$$F_R = 2\pi F_{circle} \tag{1}$$

where F_R represents the radial support force, and F_{circle} represents the circumferential tension. Since the tensile force F_l of the tensile testing machine in this experiment is twice that of F_{circle}, the relationship between the radial force and the tensile force measured by the tensile testing machine is:

$$F_R = \pi F_l \tag{2}$$

The model of the testing device for the spring back force is shown in **Figure 1**. By applying tensile forces at both ends, the stent graft is made to bend. Assuming that the target bending angle of the stent graft is α, the stent reach-

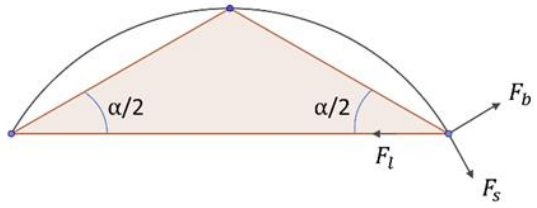


Figure 1. Mechanical model of the spring back force detection device.

es this angle when one end of it is bent by α/2. Let F_l be the tensile force at one end of the stent, F_s be the axial support force along the stent, and F_b be the spring-back force perpendicular to the axial direction of the stent. From this, the relationship between the spring back force and the tensile force can be derived as follows:

$$F_b = \frac{F_l}{sin\alpha} \tag{3}$$

Fabrication

As shown in **Figure 2A**, the compression device consists of an aluminum alloy support frame, black sliding rails, black sliders, stainless steel fixed pulleys, nylon threads, and white stent sleeves. The support frame provides overall stability. The sleeves are used to fix the two ends of the stent graft, while the sliders are employed to position and secure the sleeves. Four small metal pulleys are mounted on the support frame. Two pulleys are positioned on either side of the tensile axis of the tensile testing machine to ensure that the tensile force is transmitted vertically downward through the threads. The other two pulleys are aligned parallel to the top surface of the stent graft, ensuring that the force is applied horizontally along the tangent of the stent after being transmitted. The magnitude of the radial support force can then be accurately calculated using Equation (2).

As shown in **Figure 2B**, the bending and spring-back force testing device retains rotational and axial translational freedom when the two ends of the stent are bent. Compared to the radial compression device, this device replaces the aluminum alloy support frame and adjusts the positions of the two metal pulleys. White pulleys are installed on the two sliders to allow for the release of rotational freedom. The positions of the pulleys along the tensile axis of the tensile testing machine remain unchanged. Two metal pulleys are installed in parallel on the bottom surfaces of the sliders through angled stents. The thread on the left pulley is fixed to the right slider, and the installation method for the thread on the other side is the same. This configuration allows the two sliders to move inward as the threads are stretched horizontally, thereby bending the stent graft.

Physical experiments

Fabrication of the stent graft samples

A five-peak, Z-shaped, gradually varying stent graft was used in this experiment, as shown in **Figure 3A** [25]. The stent ring has a diameter of 33 mm, with a total length of 130 mm. Based on commercially available stents, the selected design includes seven stent rings with peak heights ranging from 10 mm to 16 mm. Other parameters include a stent wire diameter of 0.5 mm, a ring spacing of 18 mm, a coating film thickness of 0.1 mm, and a fillet radius of 0.5 mm at the peak.

Since each ring in the Z-shaped stent graft is identical, only a single-ring mold is required. As shown in **Figure 3D**, the mold is a hollow cylindrical structure with an outer diameter of φ= 33 mm [26]. Positioning holes corresponding to the fillets at the ring peaks are drilled into the mold, and adjustable pins are installed. After winding the stent wire around the pins, a pre-formed ring is shaped, and its ends are joined and fixed. Finally, the ring is sewn onto the coating film to assemble the complete stent graft.

Experimental platform and setup

The experimental platform utilizes the INSTRON 5965 dual-column tabletop electronic universal testing machine, which provides high-precision load measurements with an accuracy of±0.5% of the indicated value.

Physical experiment on radial compression

The radial support force was tested using the previously fabricated five-peak stent graft and the setup illustrated in **Figure 2A**. This device converts the tensile force generated by the testing machine into circumferential pressure applied to the stent, inducing ring compression. The machine simultaneously records displacement and force output at the terminal.

Compression displacements were incrementally increased to 20%, 40%, and 60% of the original stent diameter. Let D_0 be the original diameter, Dn the target diameter, and X the tensile displacement. According to the following formula:

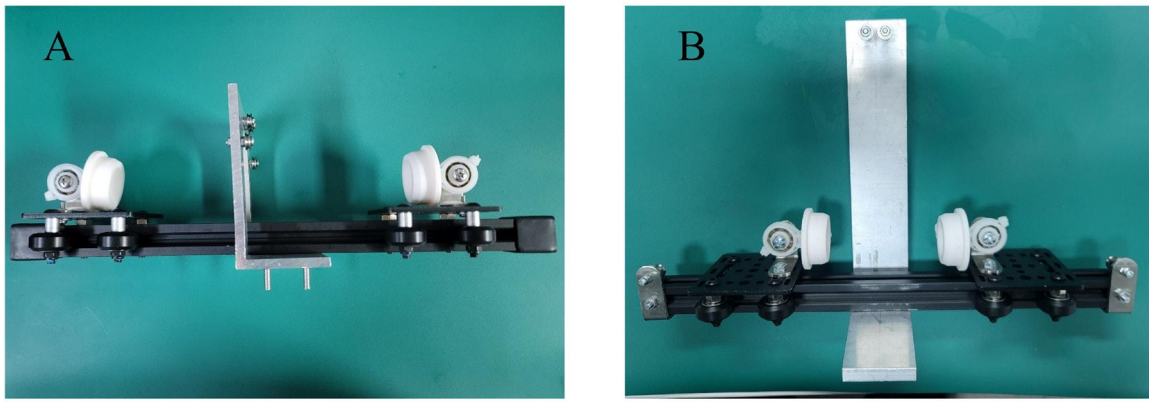


Figure 2. Physical mechanical testing device. (A) Radial support force testing device; (B) Spring back force testing device.

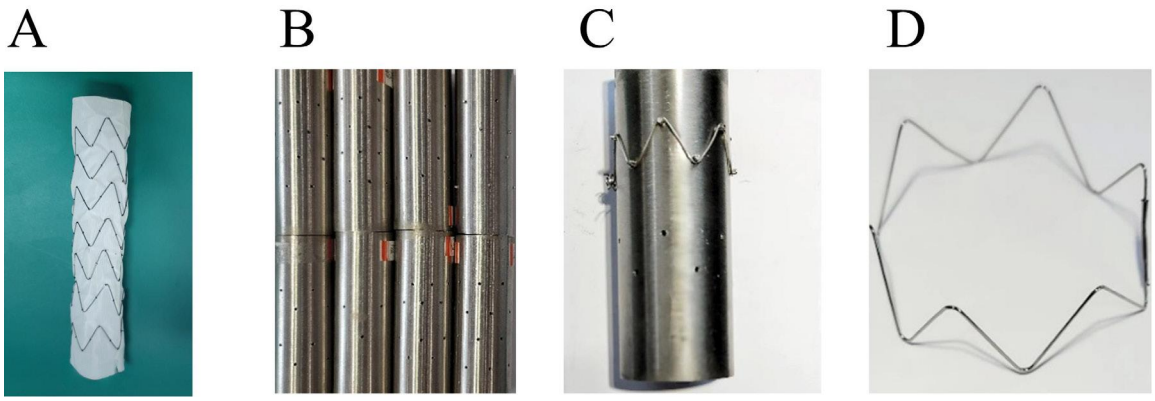


Figure 3. Mold and process for stent graft production. (A) Stent ring production mold set; (B) Support ring processing; (C) Stent ring; (D) Stent graft.

$$2*X=\pi*Do - \pi*Dn \quad (4)$$

The corresponding parameter changes are shown in **Table 1**.

During testing, all threads must remain in the same plane, and slippage at the pulleys must be avoided to ensure measurement accuracy. The tensile testing machine continuously monitors tensile force and thread displacement in real time, outputting data accordingly. The recorded tensile force is converted into radial support force using Equation (2).

Bending Experiment

The bending test employed the device illustrated in **Figure 2B**, with preset bending angles of 60°, 90°, and 180°. In this configuration, threads pull opposing sliders inward, inducing bending in the stent graft. Prior to testing, horizontal displacement must be calculated to ensure that the stent bends to the specified angles.

Assuming that the arc length along the outer curve of the stent remains constant during bending, let e_1 be the horizontal displacement and e_2 the vertical displacement of endpoints A and B, α the bending angle, and D_0 the diameter. Based on the geometric diagram in **Figure 4**, the displacements can be derived as follows:

$$e_1 = \frac{l}{2} - \left(\frac{l}{\alpha} - \frac{D_0}{2}\right) * \sin\left(\frac{\alpha}{2}\right) \quad (5)$$

$$e_2 = \left(\frac{l}{\alpha} - \frac{D_0}{2}\right) * \left(1 - \cos\left(\frac{\alpha}{2}\right)\right) \quad (6)$$

By substituting the structural parameters of the stent graft and the specified angle α the displacement values for different bending angles can be calculated. **Table 2** lists the values of e_1 and e_2 corresponding to 60°, 90°, and 180° bending.

Before testing, the stent should be pre-bent by 3–5° to prevent axial compression without bending. During the bending process, the threads and the middle plane of the stent must remain coplanar, and thread slippage on pulleys must be avoided to maintain measurement

Table 1. Corresponding relationship between tensile displacement of tensile machine and support diameter

Percentage of compression diameter	D_0/mm	D_n/mm	X/mm
20%	34	27.2	10.68
40%	34	20.4	21.36
60%	34	13.6	32.04

Table 2. The relationship between bending angle and tensile displacement of tensile machine

Stent length/mm	Rotation angle/°	Horizontal displacement e_1/mm	Vertical displacement e_2/mm
130	60	11.18	14.42
130	90	18.15	19.41
130	180	40.12	24.88

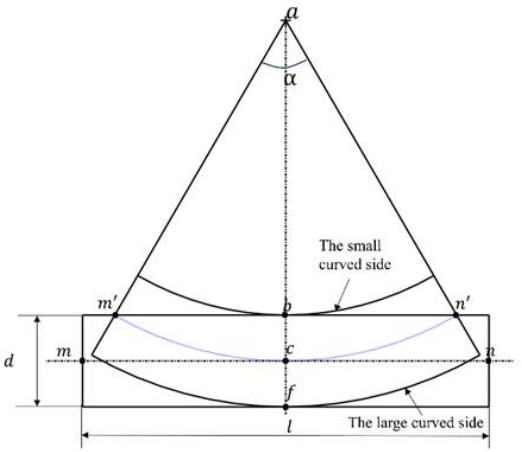


Figure 4. Geometric schematic diagram of stent bending.

accuracy. The tensile testing machine captures and outputs tensile force and thread displacement data in real time. The tensile force is then converted into the spring-back force using Equation (3).

Numerical simulation

The simulation platform uses ABAQUS 2021 for the finite element analysis. To ensure comparability between the simulation and physical experiments, the simulation model is kept consistent with the experimental setup.

The stent material is 316L stainless steel, with a density of 8.03×10^{-8} ton/mm³, a Young's modulus of 200 GPa, and a Poisson's ratio of 0.3. The coating film is made of expanded polytetrafluoroethylene (ePTFE), a medical polymer material, with a density of 2.2×10^{-12} ton/mm³, a Young's modulus of 55.2 MPa, and a Poisson's ratio of 0.46.

Explicit dynamics analysis is employed, with geometric nonlinearity enabled. The time length for the simulation is set to 0.01. The stent ring and coating film are connected by a Tie constraint, with the stent ring as the master surface and the coating film as the slave surface. The stent ring is meshed using eight-node linear hexahedral elements (C3D8R), with reduced integration and hourglass control, and a mesh density of 0.2 mm. The coating film is meshed with four-node curved thin-shell elements, also using reduced integration, hourglass control, and finite membrane strain, with a mesh density of 1 mm.

Simulation verification of radial compression

The thread curve was modeled using SOLIDWORKS 2022. A circular sketch with a diameter of 40.4 mm was created, and a helix was inserted based on the sketch. The helix has a starting angle of 180°, a pitch of 0.2 mm, and one full turn. Truss elements were used to simulate the flexible rope. The thread was meshed with T3D2 elements, and the mesh density was set to 0.5 mm. The thread curve model was imported into ABAQUS CAE 2021, assigned as a truss section, and material properties were set as follows: density 7.98E-09 ton/mm³, Young's modulus 200,000 MPa, Poisson's ratio 0.3.

The simulation is consistent with **Figure 5A**. Axial rotation and rotational degrees of freedom of the stent ring to which the load is applied were constrained, as were the axial and rotational degrees of freedom of the stretching thread. Displacement loads with equal magnitudes and opposite directions were applied at both ends of the thread. The experiment was conducted for compression displacements of 20%, 40%, and 60% of the stent's diameter.

The reaction force in the moving direction of the thread's endpoints during compression was monitored by setting the RF2 history output based on the global coordinate system at the helix endpoints.

Bending simulation verification

Reference points RP1 and RP2 were defined at the centers of the stent graft's two end faces. MPC (Multi-Point Constraint) constraints were applied between the control points and

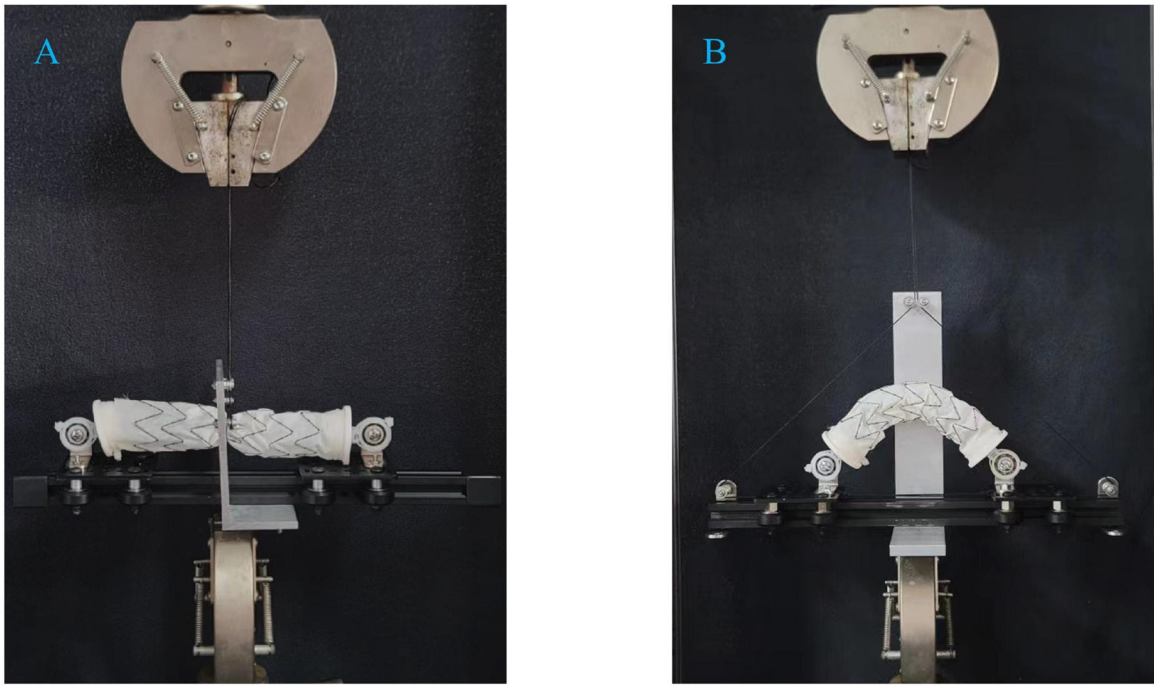


Figure 5. Mechanical performance testing of stent graft. (A) Physical experiment on compression; (B) Physical experiment on bending.

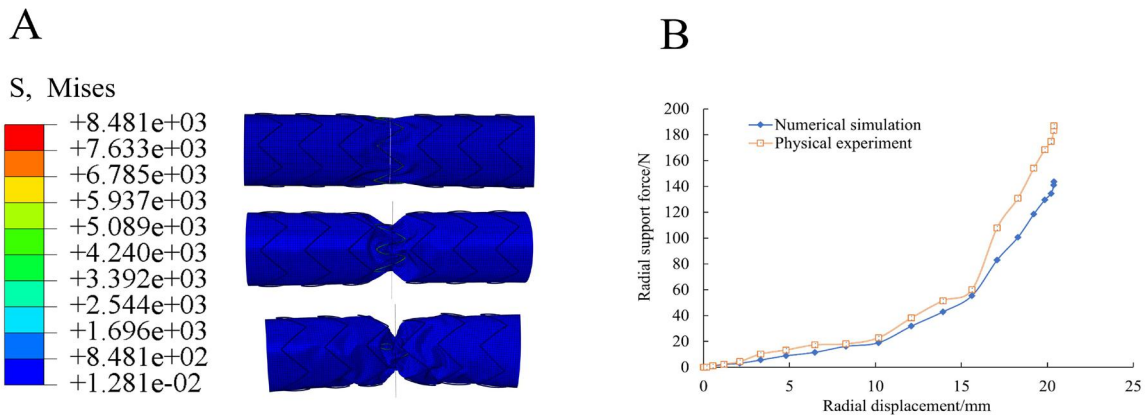


Figure 6. Stress distribution from numerical simulation. (A) Stress nephograms of the stent graft under 20%, 40%, and 60% compression; (B) Stress nephograms of the stent graft bent at 60°, 90°, and 180°.

the nodes at the top of the stent, binding their degrees of freedom together [27]. Equal and opposite angular loads were applied at the two reference points to ensure that the stent reached the desired bending angle. The change in the spring-back force during the bending process was tracked by outputting the reaction forces at the reference points. The simulation is consistent with **Figure 5B**.

III Results

Results of radial compression experiment

The simulation results are shown in **Figure 6**. During the experiment, a specially designed radial loading device, in conjunction with a tensile testing machine, was used to apply a uniform radial displacement load to the stent. The experiment measured the radial support performance of the thoracic aortic stent graft. The results showed the changes in stent morphology and radial support force values under 20%, 40%, and 60% compression displacements, both in physical and simulation experiments.

The simulation results indicate that the radial support force increases gradually with the applied load. The force increases slowly within the 0–15 mm range and rises sharply after 15 mm. In comparison, the physical experiment data largely agree with the finite element simulation results up to 17 mm of compression. However, beyond 17 mm, the increase in radial support force in the physical experiment becomes minimal. On the testing machine's display, the val-

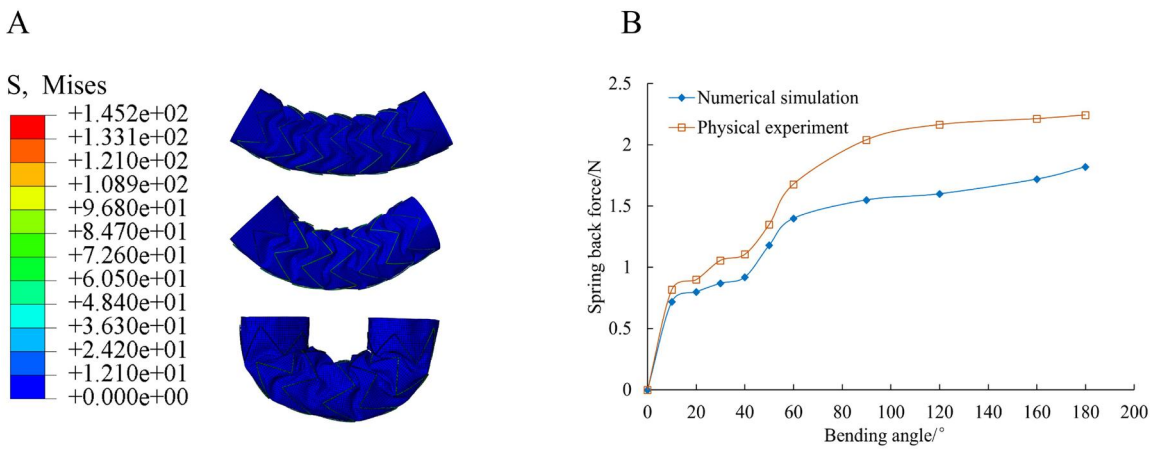


Figure 7. Comparison between Physical Experiment and Numerical Simulation. (A) Curves showing the variation of radial force in the compression experiment and simulation; (B) Curves showing the variation of spring-back force in the physical experiment and the simulation experiment.

ues show only slight fluctuating increases.

Bending experiment results

The simulation results are shown in **Figure 7**. In the physical bending and spring-back force experiment, the stent was placed in the device and bent to angles of 60°, 90°, and 180°. The mechanical parameters of the stent during bending were measured. The experimental results showed that the stent's mechanical performance during bending and recoil was in good agreement with the finite element simulation results.

In the physical experiment, the spring-back force increased rapidly during bending. After reaching 90°, the spring-back force gradually increased. Similarly, in the simulation, the radial support force remained constant when the stent bent within 0–15° but increased rapidly thereafter, showing a trend consistent with the physical experiment.

IV Discussion

The study on the radial support force and bending spring-back force of the thoracic aortic stent graft, conducted through finite element simulations and physical experiments, shows good agreement between the two methods. This consistency effectively validates the effectiveness of the force-measuring device for the thoracic aortic stent graft designed in this study.

The design and construction of the device are straightforward, achieving an effective balance between simplicity and functionality. Core components such as sliders, sliding rails, and pulleys work in coordination, ensuring the stability and reliability of the detection process. The overall structural design meets the required detection standards.

The device's performance test results indicate excellent detection accuracy. When paired with commonly used tensile testing machines, its stability and accuracy surpass those of standard dynamometers. Additionally, repeatability tests show minimal deviations across multiple measurements, demonstrating good stability and ensuring the consistency and reliability of the detection data.

While the results of the two methods align well in most cases, certain discrepancies exist under specific conditions. Notably, in the later stages of stent compression, some stent rings in the physical experiment exhibit yielding behavior, resulting in almost no increase in the radial support force. In contrast, the simulation neglects the stent material's plastic behavior, as yield stress and hardening modulus are not incorporated, which keeps the stent in the elastic phase during compression. Additionally, as the compression diameter increases, self-contact behavior occurs in the stent rings, causing the radial support force curve to become steeper in the simulation.

To improve the accuracy of future research, the following improvements are recommended: first, use a different thread material to reduce friction; second, optimize the finite element simulation model to more precisely account for the nonlinear material characteristics and complex boundary conditions; third, increase the number of experimental samples to minimize the impact of individual variations on the results.

V Conclusions

This study developed a mechanical property testing device for large-vessel stents, which can be used in conjunction with a tensile testing machine. The device is capable of performing mechanical property tests on radial support force and bending spring-back force. Physical experiments were conducted to test the support and compliance properties of the self-designed 5-peak gradient stent grafts. The results demonstrated that the stents have good support and excellent compliance, and the corresponding mechanical properties were identified. Finite element simulations were used to model the radial compression and bending of the stent grafts, and the corresponding mechanical change curves were obtained. These curves were consistent with the physical experiment results, which further validated the reliability of the stent graft testing device designed in this study.

Author contributions: Ailing Zhang conceived the project and formulated the overall experimental plan. Yu Zhou was responsible for the general assembly of the device, conducting the experiments, and processing the data. Shiju Yan completed the design of the device. The manuscript was jointly completed by all the authors.

References

[1] Dake MD, Kato N, Mitchell RS, et al. Endovascular stent-graft placement for the treatment of acute aortic dissection. N Engl J Med 1999;340(20):1546-1552.

[2] Perrin D, Demanget N, Badel P, et al. Deployment of stent grafts in curved aneurysmal arteries: toward a predictive numerical tool. Int J Numer Method Biomed Eng 2015;31(1):e02698.

[3] Kumar A, Bhatnagar N. Finite element simulation and testing of cobalt-chromium stent: a parametric study on radial strength, recoil, foreshortening, and dogboning. Comput Methods Biomech Biomed Engin 2021;24(3):245-259.

[4] Pan C, Han Y, Lu J. Structural Design of Vascular Stents: A Review. Micromachines (Basel) 2021;12(7):770.

[5] Matsumoto T, Matsubara Y, Aoyagi Y, et al. Radial force measurement of endovascular stents: Influence of stent design and diameter. Vascular 2016;24(2):171-176.

[6] Li Z, Kleinstreuer C. Analysis of biomechanical factors affecting stent-graft migration in an abdominal aortic aneurysm model. J Biomech 2006;39(12):2264-2273.

[7] Aras T, Tayeh M, Aswad A, et al. Exploring Type IIIb Endoleaks: A Literature Review to Identify Possible Physical Mechanisms and Implications. J Clin Med 2024;13(15):4293.

[8] Dong Z, Fu W, Wang Y, et al. Stent graft-induced new entry after endovascular repair for Stanford type B aortic dissection. J Vasc Surg 2010;52(6):1450-1457.

[9] Shahbazian N, Romero DA, Forbes TL, et al. Identification of geometric and mechanical factors predictive of bird-beak configuration in thoracic endovascular aortic repair using computational models of stent graft deployment. JVS Vasc Sci 2022;3:259-273.

[10] Li M, Ma T, Cai Y, et al. Numerical simulation of the distal stent graft-induced new entry after TEVAR. Int J Numer Method Biomed Eng 2024;40(5):e3819.

[11] Joo HC, Youn YN, Ko YG, et al. Comparison of open surgical versus hybrid endovascular repair for descending thoracic aortic aneurysms with distal arch involvement. J Thorac Dis 2018;10(6):3548-3557.

[12] Tang F, Hu C, Huang S, et al. An Innovative Customized Stent Graft Manufacture System Assisted by Three-Dimensional Printing Technology. Ann Thorac Surg 2021;112(1):308-314.

[13] Ryu DS, Won DS, Kim JW, et al. Efficacy of thermoplastic polyurethane and gelatin blended nanofibers covered stent graft in the porcine iliac artery. Sci Rep 2022;12(1):16524.

[14] Vellaparambil R, Han WS, Di Giovanni P, et al. Potential of auxetic designs in endovascular aortic repair: A computational study of their mechanical performance. J Mech Behav Biomed Mater 2023;138:105644.

[15] Schmidt W, Behrens P, Brandt-Wunderlich C, et al. In vitro performance investigation of bioresorbable scaffolds - Standard tests for vascular stents and beyond. Cardiovasc Revasc Med 2016;17(6):375-383.

[16] Barragan P, Garitey V, Mouneimne K, et al. Longitudinal compression behaviour of coronary stents: a bench-top comparative study. EuroIntervention 2014;9(12):1454-1462.

[17] Xue W, Lin J, Wang F, et al. Design and mechanical behaviors of sutureless prototype composite nitinol stent-grafts. Text Res J 2018;88(13):1517-1530.

[18] Zhuravleva IY, Timchenko TP, Vladimirov SV, et al. Ab ovo: Factors Affecting the Radial Stiffness of Thoracic Aorta Stent-Grafts. Sovrem Tekhnologii Med 2021;13(1):17-25.

[19] Maleckis K, Deegan P, Poulson W, et al. Comparison of femoropopliteal artery stents under axial and radial compression, axial tension, bending, and torsion deformations. J Mech Behav Biomed Mater 2017;75:160-

168.

[20] Kwiecinski J, Cheng CP, Uberoi R, et al. Thoracic aortic parallel stent-graft behaviour when subjected to radial loading. J Mech Behav Biomed Mater 2021;118:104407.

[21] Voûte MT, Hendriks JM, van Laanen JH, et al. Radial force measurements in carotid stents: influence of stent design and length of the lesion. J Vasc Interv Radiol 2011;22(5):661-666.

[22] Ormiston JA, Dixon SR, Webster MW, et al. Stent longitudinal flexibility: a comparison of 13 stent designs before and after balloon expansion. Catheter Cardiovasc Interv 2000;50(1):120-124.

[23] Demanget N, Latil P, Orgéas L, et al. Severe bending of two aortic stent-grafts: an experimental and numerical mechanical analysis. Ann Biomed Eng 2012;40(12):2674-2686.

[24] Yan S, Song C, Yang Y, et al. Design and implementation of mechanical property testing equipment for thoracic aortic stent grafts. Minim Invasive Ther Allied Technol 2024;33(2):120-128.

[25] Yan S, Song C, Si Y, et al. Design of non-equal-strut stent hoops for structural optimization of thoracic aortic stent-grafts. Minim Invasive Ther Allied Technol 2022;31(1):58-71.

[26] Zhang L, Chen X, Liu M. Research of Customized Aortic Stent Graft Manufacture. IOP Conf Ser: Mater Sci Eng 2017;187(1):012027.

[27] Vellaparambil R, Han WS, Di Giovanni P, et al. Computational Comparison of the Mechanical Behavior of Aortic Stent-Grafts Derived from Auxetic Unit Cells. Cardiovasc Eng Technol 2024;15(2):199-210.

Progress in Medical Devices
ISSN: 2957-5478

Zentime

Research Article https://doi.org/10.61189/804304ntkyxb

Investigation of water-assisted colonoscopy using a constant-temperature water infusion system

Hongsheng Li[1*], Chunhua Zhou[2*], Taojing Ran[2*], Yao Zhang[2], Xiaonan Shen[2], Shiju Yan[1], Duowu Zou[2]

[1]School of Health Science and Engineering, University of Shanghai for Science and Technology, Shanghai 200093, China. [2]Department of Gastroenterology, Ruijin Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai 200025, China.
[*]The authors contribute equally.

Corresponding authors: Duowu Zou and Shiju Yan.

Acknowledgments: This work was supported by the Medical Engineering Cross-disciplinary Project Special Fund of Ruijin Hospital, Shanghai Jiao Tong University School of Medicine, and the University of Shanghai for Science and Technology.

Declaration of conflict of interest: None.

Received February 15, 2025; Accepted June 18, 2025; Published September 30, 2025

Highlights

- Developed a 37 °C constant-temperature electronic control module, including circuit board design and selection of electronic components, ensuring stable and reliable temperature control for the water infusion system.
- Evaluated and optimized the layout and selection of pipelines and valves, ensuring secure, leak-proof water connections, precise valve flow directions, and rapid, reliable valve operation.
- Designed an efficient tailored to the dimensions and shape of the water reservoir, determining optimal parameters including heating power, structural configuration, size, and placement.

Abstract

Background: Colonoscopy is a key technique for the prevention and early detection of colorectal cancer. Water-assisted colonoscopy is increasingly adopted due to its potential to reduce patient discomfort. However, the temperature of the infused water plays a crucial role in both procedural quality and patient experience. This study aimed to optimize water-assisted colonoscopy by developing a constant-temperature water infusion system. **Methods:** A two-dimensional finite element model was established using COMSOL Multiphysics to simulate the heat transfer process between the heating base and the liquid container. The system consisted of a medical-grade 304 stainless steel container, a nichrome heating wire embedded in rubber, and an integrated piping network. Quadrilateral meshing was applied to short-range solid–liquid interfaces and triangular meshing elsewhere, resulting in detailed modeling for both natural heating (27,801 elements) and circulation heating (43,998 elements). Based on simulation results, a hardware platform was developed to deliver sterile water at a constant temperature of 37 °C for digestive endoscopic procedures. **Results:** Circulation heating demonstrated superior thermal efficiency and more uniform temperature distribution than natural heating. Under ambient conditions (25 °C), the system reliably maintained water temperature at (37±1)°C . Partitioned meshing enhanced computational precision with a minimum element size of 0.1 mm. Solid-liquid coupling analysis confirmed stable heat conduction during dynamic infusion. The device allows for independent temperature presetting and stepless flow rate adjustment via a control panel. It is also compatible with standard endoscopic systems, thereby enhancing procedural efficiency and safety. **Conclusion:** The proposed constant-temperature water infusion system model offers a reliable and adaptable solution for water-assisted colonoscopy, improving both diagnostic performance and patient comfort through precise thermal regulation.

Keywords: Water-assisted colonoscopy, constant-temperature control, finite element analysis, thermodynamic simulation, medical device design

Address correspondence to: Duowu Zou, Department of Gastroenterology, Ruijin Hospital, Shanghai Jiao Tong University School of Medicine, No. 227 South Chongqing Road, Huangpu District, Shanghai 200025, China. E-mail: zdw_pi@163.com. **Shiju Yan**, 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: yanshiju@usst.edu.cn.

Introduction

Colorectal cancer remains a significant global public health concern, especially in developing countries. Early detection and prevention through screening colonoscopy are essential strategies for reducing both the incidence and mortality associated with colorectal cancer [1, 2]. The detection and removal of precancerous polyps during colonoscopy have been proven to be effective in preventing the progression to colorectal cancer [3]. However, conventional colonoscopy techniques that utilize air insufflation often cause considerable patient discomfort, leading to decreased adherence to screening recommendations. Such discomfort may include abdominal pain, bloating, and overall intolerance to the procedure, potentially discouraging patients from undergoing necessary follow-up examinations [4, 5].

Although sedation and analgesia are commonly employed during colonoscopy, the use of propofol and other anesthetic agents is generally not recommended for patients with significant comorbidities or high-risk profiles. Additionally, some patients require the presence of a family member during anesthetized procedures, followed by a recovery period, which may pose logistical challenges. In recent years, water-assisted colonoscopy (WAC) has emerged as a promising alternative to conventional air insufflation techniques [6-8]. WAC methods are generally classified into two main types: water immersion and water exchange [9]. The efficacy of WAC has been validated in multiple randomized controlled trials [10, 11]. A recent meta-analysis reported that the water exchange method achieved the highest adenoma detection rate, particularly in the right colon and in colorectal cancer screening cases [12].

Additionally, the Boston Bowel Preparation Scale scores were significantly higher with water exchange. This technique was also associated with a higher proportion of unsedated examinations and reduced real-time insertion pain, making it the least painful insertion method and potentially improving procedural accessibility and cost-effectiveness. Investigations have examined the impact of water temperature on procedural outcomes. Instilling room-temperature water into the rectosigmoid colon has been linked to increased mucus secretion, necessitating additional cleansing during withdrawal [13]. A double-blind randomized trial comparing cool versus warm water immersion demonstrated that, while both approaches were effective, warm water provided advantages in reducing the need for auxiliary maneuvers [14]. These findings suggest that maintaining infused water at body temperature (37 °C) may enhance both patient comfort and procedural efficiency.

The challenge of maintaining constant water temperature during colonoscopy has generated increasing interest in the development of automated temperature control systems. These systems need to be precise, reliable, and capable of sustaining stable temperatures throughout the entire procedure [15, 16]. The integration of modern technologies-such as embedded controllers and high-precision temperature monitoring systems-offers promising solutions to address this challenge.

Numerical simulation

System modeling

The interaction between the heating chassis and the liquid container was investigated using finite element simulations conducted in COMSOL Multiphysics. The objective was to generate simulation data to support the design of a constant-temperature water infusion system. Considering the complexity of the solid-liquid coupling problem-particularly regarding real-time performance and computational efficiency-a two-dimensional model was employed to simplify the heating system. This approach effectively reduced model complexity and significantly shortened computation time, while preserving the reliability of simulation outcomes. Given the minimal wall thickness of the hose (only 1-1.2 mm), its impact on the overall thermodynamic behavior of the solid-liquid system was considered negligible. Therefore, hose wall was further simplified as a thin boundary layer. Additionally, the container shell was modeled as a heat transfer interface to simulate thermal conduction.

The finite element model used in this simulation, as illustrated in **Figure 1**, is divided into two sections. The left panel shows the key components of the thermostatic heating system, including the liquid container, heating plate, and piping network. The right panel presents the geometric dimensions of the model.

By removing the piping system, a natural heating model was developed to facilitate a comparative analysis of heat conduction and fluid dynamics between the two configurations. The liquid container is cylindrical, with a bottom diameter labeled as R, an overall height of H,

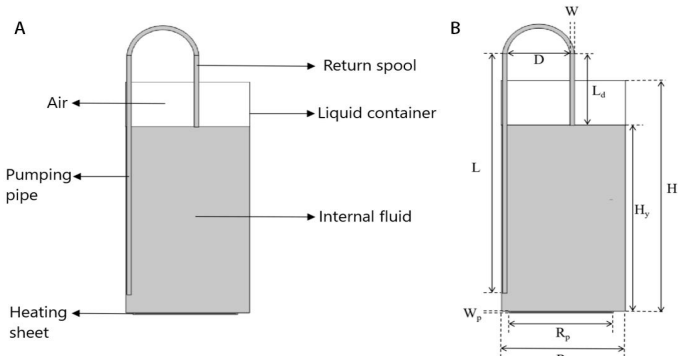


Figure 1. Finite element model used for thermal simulation. (A) Structural layout of the model; (B) Geometric dimensions of the model.

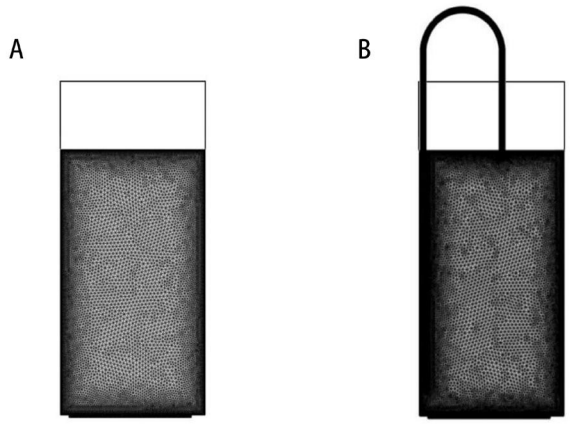


Figure 2. Mesh division of the water infusion container model. (A) Natural heating mode; (B) Circulating heating mode.

and a liquid height of H_y. The bottom heating plate is disc-shaped, with a diameter of R_p and a thickness of W_p. The integrated piping system comprises both long and short pipes, each with a diameter W, lengths L and L_d, respectively, and an inter-pipe spacing denoted as D. Detailed geometric parameters for each component were defined during model construction to ensure both accuracy and practical relevance. The specific values of these parameters are provided in **Table 1**.

To ensure precise temperature control and enhance heat transfer efficiency, the liquid container was fabricated using medical-grade 304 stainless steel. The housing of the heating element was made of rubber, with a nickel–chromium alloy heating wire embedded inside. Based on these materials and structural characteristics, simulations were conducted to analyze temperature distribution and heat flow dynamics during the heating process. Further details of these parameters are provided in **Table 2**.

The simulation analysis in this study incorporated multiple physical fields, including fluid flow and multiphysics coupling. The initial phase of the thermal boundary condition analysis involved defining the external air environment and the internal physiological saline, as well as setting the initial temperature of all system components to 25 °C. Calculations were performed using the finite element method, specifically employing the turbulence and solid-liquid coupling modules.

The finite element simulation in this study focused on heat conduction and fluid flow phenomena near the short-distance interface where solid–liquid contact occurs. A differentiated meshing strategy was employed, applying quadrilateral meshing to the contact edge regions and triangular meshing to the remainder of the model. Particular attention was given to the rubber heating sheet, silicone tubing, and normal saline, with finer mesh refinement applied to these critical components. The mesh distribution of the model is illustrated in **Figure 2**. For the natural heating mode, the model consisted of 27,801 domain elements and 1,016 boundary elements, with a minimum element size of 0.2 mm. In the circulating heating mode, the mesh comprised 43,998 domain elements and 2,133 boundary elements, with a minimum element size of 0.1 mm.

Simulation data analysis

This study examined temperature variations during natural and circulating heating processes to evaluate differences in heating efficiency and temperature distribution. Simulation results were analyzed to visualize the temperature fields at three critical time points (0.1 h, 0.14 h, 0.18 h) using thermal cloud maps, as shown in **Figures 3**, **4**, and **5**. In the natural heating mode, the heating sheet rapidly heated up and transferred heat to the bottom of the container, eventually reaching the lower layer of the liquid. However, this process resulted in an uneven temperature distribution within the liquid, with a notable temperature gradient between the bottom and upper middle layers, even when the bottom temperature approached the target value of 37 °C at 0.18 h.

In contrast, the circulating heating mode-

Table 1. Geometric parameters of the water infusion container model

Parameter	R	H	H_y	R_p	W_p	W	L	L_d	D
Value (mm)	140	280	220	120	2	6	300	100	70

Table 2. Thermophysical properties of four different materials used in the heating system

Property	304 Stainless Steel	Rubber	Nichrome	Normal Saline
Density (kg/m^3)	7,930	1,100	8,400	1,000
Melting Point (℃)	1,398	177	1,400	-0.5
Electrical Conductivity (S/m)	1.45×10^{-6}	1.0×10^{-10}	1.2×10^{-6}	1.553
Thermal Conductivity (W/m·K)	16.3	0.5	30	0.6
Specific Heat Capacity (J/kg·K)	500	1,300	440	4,200

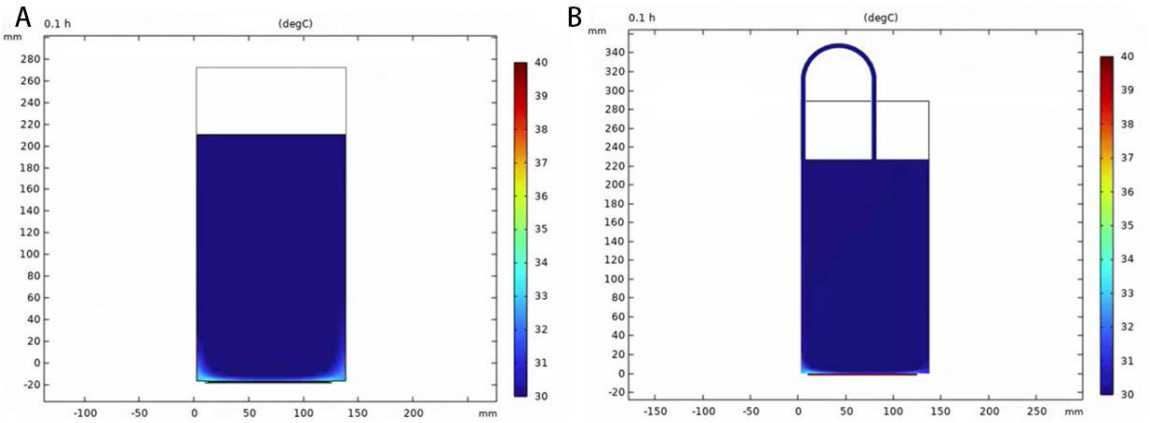


Figure 3. Heat distribution in the water infusion container model at 0.1 h. (A) Natural heating mode; (B) Circulating heating mode.

enabled by a closed-loop pipe and valve system-demonstrated more uniform heat distribution and improved temperature control. The circulating system reached the target temperature more rapidly than the natural heating mode, highlighting its superior thermal efficiency.

System development

The system developed in this study consists of three main components: the circuit control system, the external working devices, and the mechanical control system. A schematic representation of the constant temperature test platform is shown in **Figure 6**. The circuit control system includes several key components, such as a central processor for data processing, a heating system with a proportional-integral-derivative (PID) adaptive algorithm, a multi-channel real-time temperature measurement circuit, an efficient water level alarm system, a stable power supply module, and a multifunctional human-computer interaction panel. Additionally, it features various control instructions and communication modules to facilitate system operation. The external working devices include sensors and heating elements specifically designed for the liquid container. These components transmit signals to the circuit control system in real time, enabling dynamic adjustments to the heating process according to programmed instructions. The mechanical control system is responsible for regulating the valve channels and pump speeds, ensuring smooth switching between injection modes and enabling precise control of the liquid flow rate within the container and pipes.

Hardware design

Main control circuit

The STM32F103ZET6 was selected as the main control chip for this study. With improvements in key performance metrics such as data conversion rate, storage capacity, operating temperature range, power consumption, and interface functionality, this chip provides enhanced performance and flexibility, making it ideal for complex embedded applications [17].

Heater element

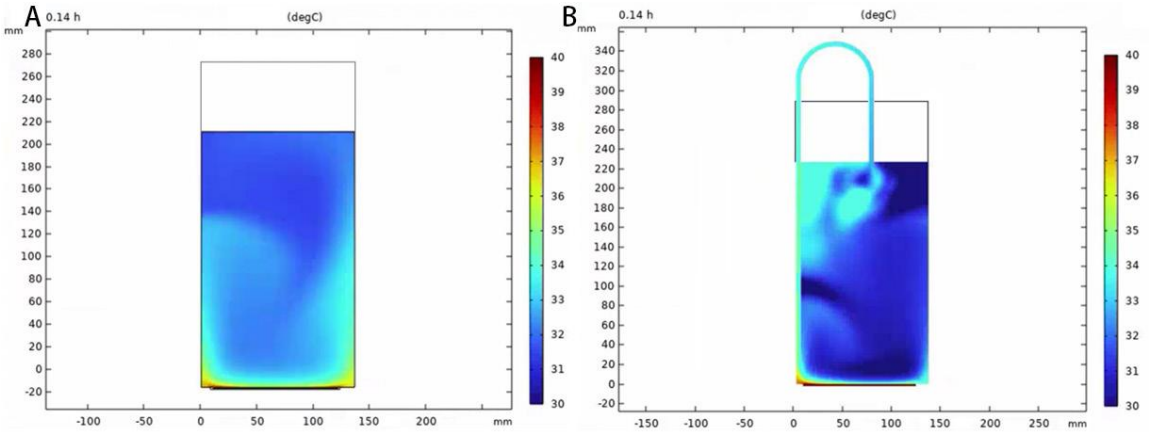


Figure 4. Heat distribution in the water infusion container model at 0.14 h. (A) Natural heating mode; (B) Circulating heating mode.

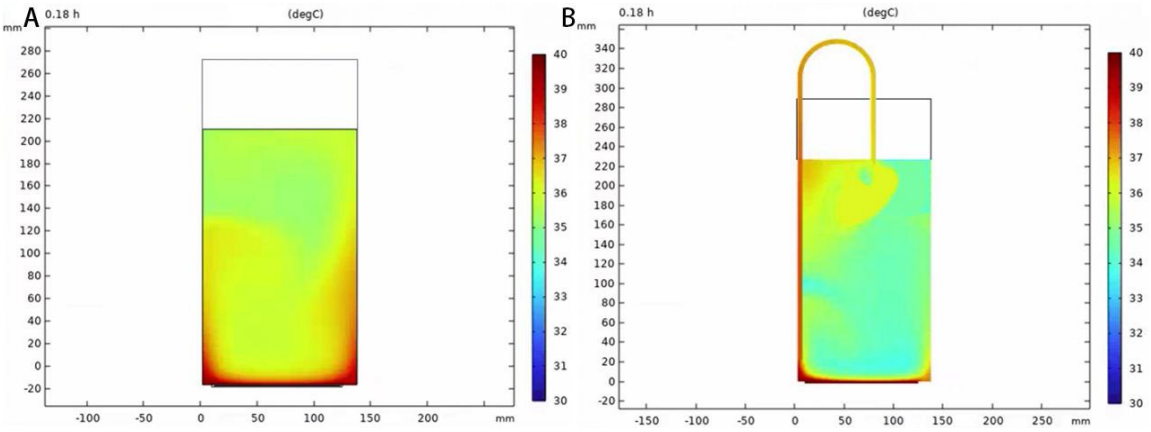


Figure 5. Heat distribution in the water infusion container model at 0.18 h. (A) Natural heating mode; (B) Circulating heating mode.

To prevent direct contact between the heater and the liquid during the diagnostic and therapeutic process, a silicone rubber heating sheet was employed in this study. The heating sheet is designed to be directly affixed to the bottom of the liquid container. It is made from a nickel-chromium alloy, a material with high resistivity, allowing it to self-generate heat when power is applied. This design offers flexibility, enabling customization to the shape and size of the container while ensuring uniform surface heating. Additionally, the heating sheet exhibits excellent thermal inertia, which enables gradual heat release, preventing local overheating and maintaining the stability and reliability of the heating process.

Temperature measuring circuit

In this study, negative temperature coefficient thermistors were primarily employed as temperature sensors to meet the measurement requirements. The sensor has a resistance value of 10K at room temperature, which varies with changes in the surrounding temperature, making it suitable for applications within the range of -55 °C to 300 °C [18]. The thermoplastic elastomer waterproof temperature probe is the actual sensor used, as it is compact enough to be integrated into a micro device and strategically positioned at a critical monitoring location within the container. The experimental setup includes a total of 12 pre-established temperature measurement points, with each sensor connected to the circuit control module via data transmission lines in the positioning bracket [19].

Temperature control circuit

The temperature control circuit developed in this study primarily consists of a PC817 optocoupler, an S8050 low-power NPN silicon transistor, and an SLA-12VC-SL-A relay. The circuit employs an electrical isolation drive scheme, as shown in **Figure 7**. It includes three main interfaces: the general purpose input/output control port on the left, which is directly connected to the main control circuit; the power

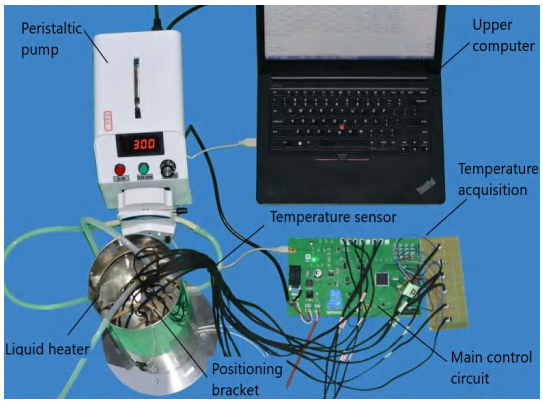


Figure 6. Prototype of the developed water infusion system.

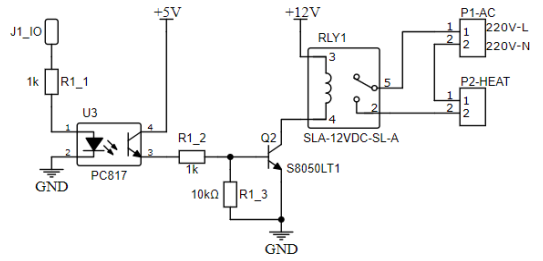


Figure 7. Temperature control circuit of the developed water infusion system. GND, Ground.

input port for 220 V AC; and the connection port for the heating component [20].

Upper computer display

This study employs LabVIEW software to develop an upper computer program for display purposes. The program facilitates real-time monitoring and management of the lower electrical circuit and its operational status. The system interface and component block diagram are shown in **Figure 8** and **Figure 9**, respectively. The software platform includes indicators, drawing tools, and user components. Various communication protocols supported by LabVIEW are utilized to configure parameters, collect temperature data, plot real-time change curves, and perform other core functions [21].

PID algorithm

To achieve precise temperature control, this study primarily employs the PID algorithm within the control system. The PID equation (using the coefficient method) is applied for this purpose.

$$u(k) = Kp * (e(k) - e(k-1)) + Ki * e(k) + Kd * (e(k) - 2e(k-1) + e(k-2))$$

In the equation, u (k) represents the controller output, while e (k) denotes the control deviation. The numerical signal is transmitted to the main control chip through the general purpose input/output port via the temperature acquisition circuit, and the PID calculation is performed in milliseconds [22]. Subsequently, the pulse width modulation (PWM) generator is used to adjust the heating power of the heating sheet. Once the proportional coefficient Kp, integral coefficient Ki, and derivative coefficient Kd are determined through calibration, the system is able to regulate the temperature of the static liquid with an error margin of less than 1 °C [23].

Functional verification and performance testing

Experimental setup

To implement the system, the following steps should be taken: First, activate the power module switch to initiate the system and bring the device to a ready state. Second, activate the pressure sensors to monitor the water injection process. The control panel will immediately display the amount of water injected into the liquid container, while the system will maintain minimum and maximum alarm limits for the water level. If the liquid level drops below the minimum threshold, a continuous beep will sound until the level rises above this threshold.

The system also includes a maximum water level alarm to prevent overflow, with the maximum water level set at 4 L for the container in use. Once the liquid reservoir is filled, the heating temperature can be preset using the control panel button, and the pump speed can be adjusted. After completing the preparation, open the second valve to initiate the internal circulation function. Once the device is started, the heating system should be activated to begin the heating process. This will cause the bottom heating element to rapidly heat the liquid to the desired temperature. Temperature distribution within the liquid can be closely monitored using the temperature measurement points, numbered from No. 1 to No. 12, located on the positioning bracket, as shown in **Figure 10**.

During the experiment, the core components of the temperature control system operated continuously to monitor the liquid temperature and transmit data to the central control unit in real time. The system automatically adjusts the operation of the heating device using an advanced control algorithm, which compares the real-time temperature with the preset target value. To ensure the precision of the

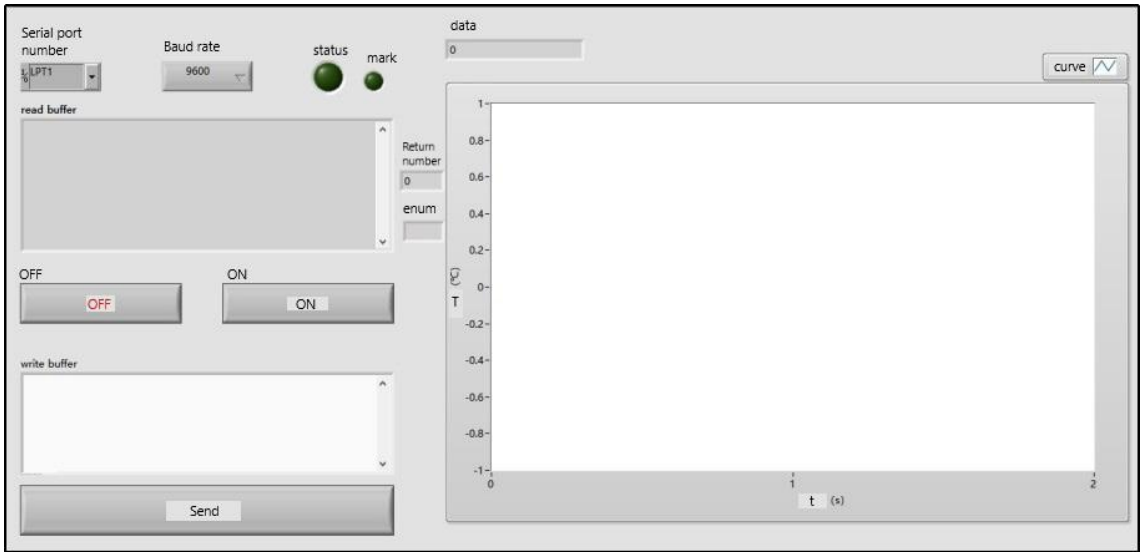


Figure 8. System interface of the developed water infusion system.

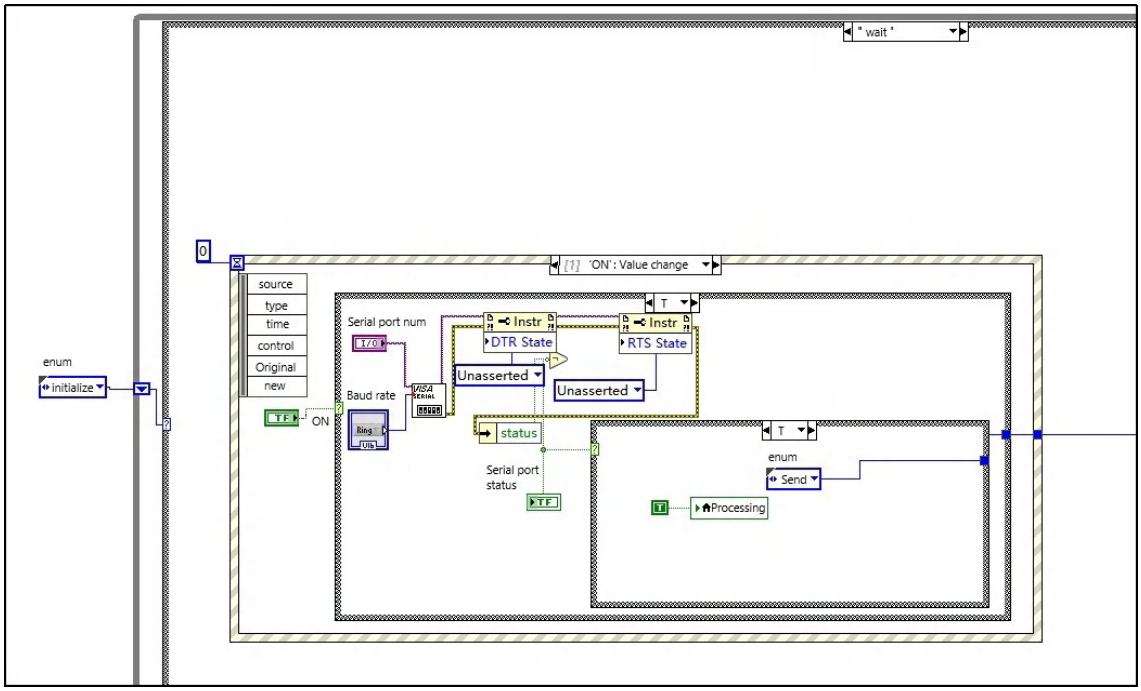


Figure 9. Component block diagram of the developed water infusion system.

experiment, the No. 1 valve remains closed until the liquid temperature reaches the required conditions. Once the liquid temperature reaches the target value, a progress indicator appears on the control panel, signaling the operator to manually initiate the water injection process. This involves adjusting the No. 1 valve and the pump speed control to accurately inject water into the simulated intestine through the injection port.

Results

Finite element simulation results

After conducting a comprehensive evaluation of the thermostatic heating system's performance, significant conclusions were drawn, particularly regarding the heating mode with constant temperature control. When relying solely on control for heating, the slow natural convection within the liquid leads to inefficient heat transfer, resulting in uneven temperature distribution. The maximum temperature difference can reach up to 5 °C, highlighting the limitations of a single thermostatic control mode in maintaining a uniform internal liquid temperature.

To address this issue, a proposal was made to

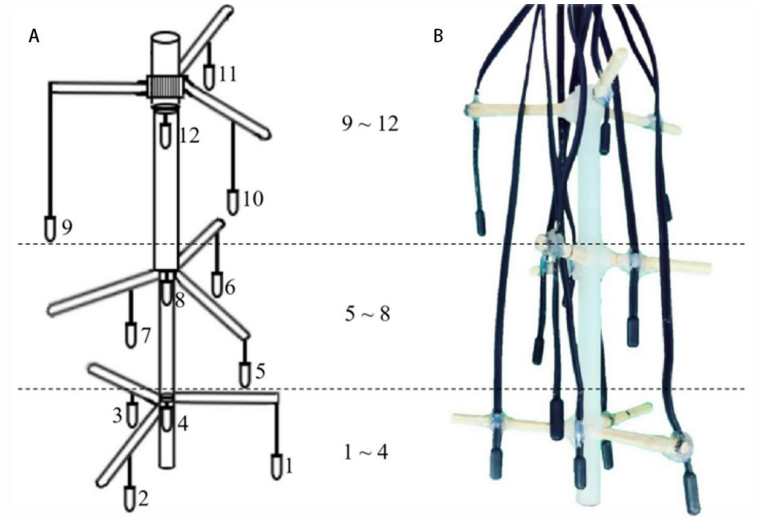


Figure 10. Multi-point temperature measurement stent for prototype performance testing. (A) Layout drawing; (B) Physical connection.

enhance heat transfer by increasing the return flow of the pipeline as an auxiliary measure. This approach accelerates temperature equalization by enhancing liquid heat circulation during heating, reducing the maximum temperature difference to approximately 1 °C. This not only visually illustrates the temperature variation between the upper and lower levels in the natural heating mode but also highlights the importance of pipeline reflux in improving the accuracy of liquid temperature control. Furthermore, the circulating method results in significantly more consistent liquid heat distribution compared to a simulation model without the circulating pipe and valve system.

Experimental platform results

In this study, the initial temperature of the liquid was set to approximately 25.5 °C, reflecting the ambient room temperature. The liquid volume was maintained at a constant 3.5 L, with a target temperature of 37 °C, designed to simulate the internal temperature of an organism. To ensure data accuracy and system stability, any extreme outliers were excluded from the temperature measurements. The temperature difference ($\Delta t = t_{max} - t_{min}$) was then calculated based on the highest and lowest values from the remaining data.

The experiment identifies two key time points: t1, which marks the onset of significant temperature fluctuations (ΔT exceeding approximately 0.5 °C), indicating the responsiveness of the heating system; and t2, which represents the time required for the system to reach the target temperature (T_{set}), determined by the mean of the remaining 10 temperature readings. This time point serves as an indicator of the system's efficiency in raising the water temperature to the desired level. This analysis aims to provide a comprehensive understanding of the dynamic response of the temperature control system and its ability to adapt to varying experimental conditions, thereby informing enhancements in the design and operational strategies of the experimental setup.

The study focuses on monitoring the maximum and minimum temperatures during the heating process, as well as the variations in the maximum temperature difference between them. Dynamic temperature changes at 200 W without pump-assisted convection are shown in **Figure 11**. This figure highlights the natural temperature distribution pattern in the absence of external forces and emphasizes the potential maximum temperature gradient under these conditions. The optimal configuration, identified after screening, involves circulating heating at 200 W power and 300 rpm, as shown in **Figure 12**. The temperature fluctuations observed in this setup demonstrate how precise control of heater power and pump speed can enhance the uniformity and efficiency of the heating process.

The system's thermal dynamic response characteristics were analyzed during initial startup under specific experimental environmental conditions, including normal atmospheric pressure and room temperature (25 °C). Analysis of the experimental data revealed an initial lack of significant temperature change in both liquid groups, which can be attributed to inherent system response time delays and heat transfer buffering effects. This indicated limited heat conduction efficiency at the onset of system operation, leading to a transient temperature stability phase. However, after approximately 50 seconds, a noticeable temperature increase was observed in both experimental conditions. Subsequently, the liquid temperature exhibited a trend toward stabilization, eventually reaching a relatively constant state. The temperature fluctuations around the designated value indicated that the system had achieved thermal equilibrium.

The pump body was not employed for heating in the experimental group. As a result, the gradual

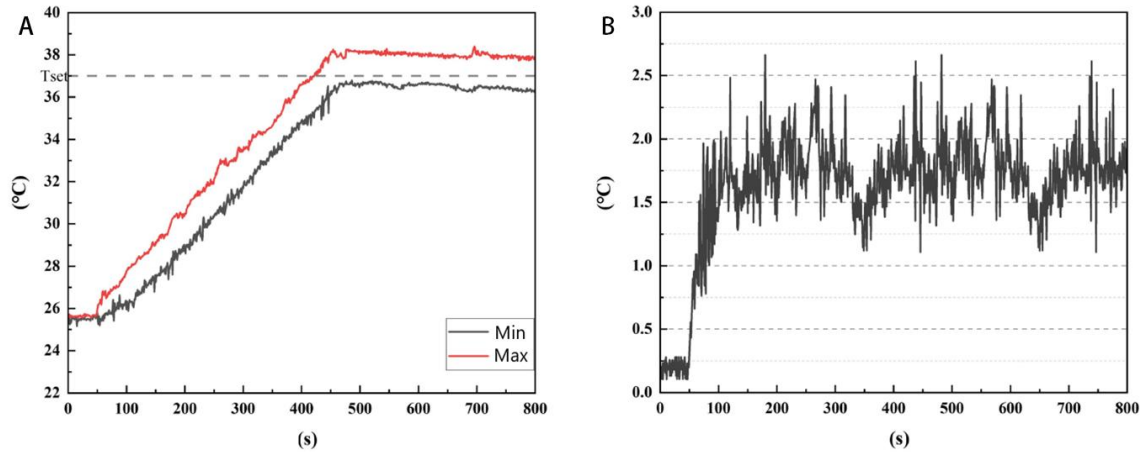


Figure 11. Temperature variation of the prototype water infusion device at 200 W, 0 rpm. (A) Minimum maximum temperature; (B) Maximum temperature difference.

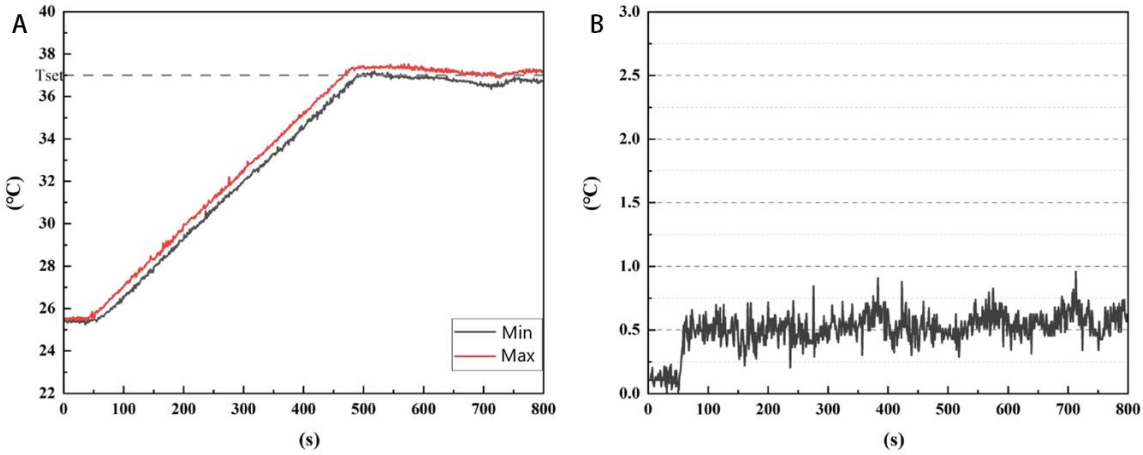


Figure 12. Temperature variation of the prototype water infusion device at 200 w, 300 rpm. (A) Minimum maximum temperature; (B) Maximum temperature difference.

increase in liquid temperature led to increasing inhomogeneity in temperature distribution. This was evident from the significant rise in the temperature difference between the highest and lowest recorded points, which reached a maximum of approximately 1.8 °C, with isolated instances exceeding 2.6 °C. The fluctuation underscores how the absence of the pump body significantly affects liquid temperature distribution. Around 8 minutes into the experiment, a notable inflection point was observed in the temperature data from the two measurement points, indicating that the liquid temperature was approaching the target temperature (T_set), eventually achieving the preset value of 37 °C. Although the liquid temperature stabilized thereafter, a noticeable disparity between the highest and lowest temperatures remained.

Analysis of the two temperature measurement points revealed a significant reduction in internal temperature variation after heating, at approximately 0.4 °C, underscoring the circulating heating system's effectiveness in enhancing heating uniformity. Once the target temperature (T_set) was reached, the temperature difference stabilized at approximately 0.5 °C, despite a brief overshoot in the maximum reading. As the experiment progressed, the temperature difference between the two measurement points gradually diminished, ultimately converging to near uniformity.

Discussion

This study focused on the development of a 37 °C water infusion device for colonoscopy diagnosis and treatment, addressing several key research and development issues. Firstly, it was demonstrated that the circulating heating mode, as opposed to natural convection heating, enhances heating efficiency and ensures uniform temperature distribution of the infusion water. Secondly, the device effectively

maintains a steady 37 °C temperature for the infusion water in room temperature settings.

In the finite element simulation section of this study, two distinct heating modes were compared for liquid containers: natural convection heating mode and cyclic uniform temperature control heating mode. The comparison aimed to assess heating efficiency and temperature distribution uniformity between the two modes.

The numerical simulation results indicated that the circulating uniform-temperature control heating mode exhibited significant advantages. A temperature control scheme combining a PID control algorithm with PWM technology was utilized in the prototype development. This scheme integrates the precise adjustment capabilities of the PID algorithm with the dynamic power output control enabled by PWM technology. It not only allows for efficient regulation of system heating power but also optimizes energy utilization and reduces system energy consumption. Compared to traditional PID technology, this combined approach offers more precise control by linking the PID output with the PWM duty cycle. This integration allows the system to respond more quickly to changes, thereby enhancing the dynamic performance of the system.

By appropriately adjusting the pump flow rate and implementing this temperature control scheme, the continuity and uniformity of the infusion water heating process can be maintained, ensuring safety, quality, and improved patient tolerance during endoscopic diagnostic and treatment procedures.

Conclusion

In this study, a device was developed to provide constant-temperature water for WAC via water infusion. The device underwent functional verification and performance evaluation through both numerical simulation and prototype testing. Simulation results demonstrated that the cyclic uniform heating mode, as opposed to natural convection heating, significantly enhances heating efficiency and ensures more uniform infusion water temperature. Prototype testing further confirmed that the temperature control scheme-integrating PID control with PWM technology-offers superior accuracy, rapid response, and stability. The implementation of 37 °C constant-temperature water infusion in WAC is expected to significantly improve the efficiency, safety, and patient satisfaction associated with endoscopic procedures.

However, several aspects of this study remain to be fully explored. To address these limitations, future work will focus on the following areas for improvement and further discussion. Subsequent research may consider incorporating fuzzy control technology to optimize the control process and enhance the system's resistance to external disturbances. Additionally, the adoption of multiple control schemes is proposed, allowing operators to control functions-such as starting and stopping the water infusion system-without the need for manual operation.

Author contributions: Hongsheng Li: Writing-original draft, Methodology, Investigation, Formal analysis, Data curation. Chunhua Zhou: Conceptualization, Writing-review and editing. Taojing Ran: Writing-review and editing. Yao Zhang: Investigation, Data curation. Xiaonan Shen: Investigation, Data curation. Shiju Yan: Conceptualization, Supervision, Writing-review and editing. Duowu Zou: Conceptualization, Supervision, Funding acquisition.

References

[1] Bray F, Ferlay J, Soerjomataram I, 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] Davidson KW, Barry MJ, Mangione CM, et al. Screening for Colorectal Cancer: US Preventive Services Task Force Recommendation Statement. Jama 2021;325(19):1965-1977.

[3] Issaka RB, Chan AT, Gupta S. AGA Clinical Practice Update on Risk Stratification for Colorectal Cancer Screening and Post-Polypectomy Surveillance: Expert Review. Gastroenterology 2023;165(5):1280-1291.

[4] Leung CW, Kaltenbach T, Soetikno R, et al. Water immersion versus standard colonoscopy insertion technique: randomized trial shows promise for minimal sedation. Endoscopy 2010;42(7):557-563.

[5] Hsieh YH, Tseng CW, Hu CT, et al. Prospective multicenter randomized controlled trial comparing adenoma detection rate in colonoscopy using water exchange, water immersion, and air insufflation. Gastrointest Endosc 2017;86(1):192-201.

[6] Cadoni S, Ishaq S. How to perform water-aided colonoscopy, with differences between water immersion and water exchange: a teaching video demonstration. VideoGIE 2018;3(5):169-170.

[7] Leung FW. Methods of reducing discomfort during colonoscopy. Dig Dis Sci

2008;53(6):1462-1467.
[8] Terruzzi V, Paggi S, Amato A, et al. Unsedated colonoscopy: A neverending story. World J Gastrointest Endosc 2012;4(4):137-141.
[9] Chen Z, Li Z, Yu X, et al. Is water exchange superior to water immersion for colonoscopy? A systematic review and meta-analysis. Saudi J Gastroenterol 2018;24(5):259-267.
[10] Hsieh Y, Tseng C, Koo M, et al. Feasibility of sedation on demand in Taiwan using water exchange and air insufflation: A randomized controlled trial. J Gastroenterol Hepatol 2020;35(2):256-262.
[11] Hsieh YH, Koo M, Leung FW. A patient-blinded randomized, controlled trial comparing air insufflation, water immersion, and water exchange during minimally sedated colonoscopy. Am J Gastroenterol 2014;109(9):1390-1400.
[12] Khan R, Ruan Y, Yuan Y, et al. Relative Efficacies of Interventions to Improve the Quality of Screening-Related Colonoscopy: A Systematic Review and Network Meta-Analysis of Randomized Controlled Trials. Gastroenterology 2024;167(3):560-590.
[13] El Rahyel A, McWhinney CD, Parsa N, et al. Room temperature water infusion during colonoscopy insertion induces rectosigmoid colon mucus production. Endoscopy 2020;52(12):1118-1121.
[14] Falt P, Šmajstrla V, Fojtík P, et al. Cool water vs warm water immersion for minimal sedation colonoscopy: a double-blind randomized trial. Colorectal Dis 2013;15(10):e612-617.
[15] Calcara C, Aseni P, Siau K, et al. Water immersion sigmoidoscopy versus standard insufflation for colorectal cancer screening: A cohort study. Saudi J Gastroenterol 2022;28(1):39-45.
[16] Leung FW, Leung JW, Mann SK, et al. The water method significantly enhances patient-centered outcomes in sedated and unsedated colonoscopy. Endoscopy 2011;43(9):816-821.
[17] Oyetoke OO. A practical application of ARM Cortex-M3 processor core in embedded system engineering. Int J Intell Syst Appl 2017;9(7):70-88.
[18] Liu G, Guo L, Liu C, et al. Evaluation of different calibration equations for NTC thermistor applied to high-precision temperature measurement. Measurement 2018;120:21-27.
[19] Kharitonov OV, Firsova LA. Separation Features of TPE and REE by Complex Displacement Chromatography from Highly Active Solutions Formed after SNF Processing. Radiochemistry 2022;64(6):721-727.
[20] Zong X, Zhang J-F, Zheng J, et al. Design of boundary pulse width modulation controller for a class of heat equations. Appl Math Comput 2023;437:127511.
[21] Hao H, Zhang HL. Design of flow measurement system based on LabVIEW. IOP Conference Series: Materials Science and Engineering 2020;789:012027.
[22] Meng J, Gao H, Ruan M, et al. Design of vacuum annealing furnace temperature control system based on GA-Fuzzy-PID algorithm. Plos One 2023;18(11):e0293823.
[23] Singgih H, Siswoko S, Komarudin A. Study of PID control implementation in the process of cheap breaking using Peltier cooling elements. IOP Conference Series: Materials Science and Engineering 2020;732:012061.

Progress in Medical Devices
ISSN: 2957-5478

Zentime

Research Article https://doi.org/10.61189/941872mmikqi

Transformer network–based disease subtyping from multidimensional lesion-layer features

Linrong Yuan*, Yutong Xie*, Danhong Li*, Jianghui Li, Miao Yu, Siqi Wang, Yu Wang, He Ren

Faculty of Medical Instrumentation, Shanghai University of Medicine and Health Sciences, Shanghai 201318, China.
*The authors contribute equally.

Corresponding author: He Ren.

Declaration of conflict of interest: None.

Received July 26, 2025; Accepted September 10, 2025; Published September 30, 2025

Highlights

- A total of 289 patient CT datasets were analyzed, and 15 optimal radiomic features were identified using ANOVA, correlation analysis, and random forest ranking, ensuring high discriminative power and clinical interpretability.
- The proposed model demonstrated excellent performance (Accuracy: 0.98, Area Under the Curve: 0.99) in training set, demonstrating robust learning capacity and the ability to distinguish lesion subtypes from multidimensional radiomic features.
- By leveraging serialized radiomic trends rather than isolated feature analysis, this study provides a new paradigm for early screening and personalized diagnosis of lung adenocarcinoma.

Abstract

Objective: To develop and validate a Transformer-based radiomics model for classifying lung adenocarcinoma subtypes from computed tomography imaging data. **Methods:** We retrospectively collected 289 computed tomography images of lung adenocarcinoma, including adenocarcinoma in situ, minimally invasive adenocarcinoma, and invasive adenocarcinoma. Correlation-based feature analysis was employed and identified 15 optimal radiomic features. A Transformer-based classification model incorporating multi-head attention and position-wise feed-forward Networks was subsequently constructed. **Results:** The proposed model achieved a training accuracy of 0.98, test accuracy of 0.914, training recall of 0.942, test recall of 0.874, training F1-score of 0.940, test F1-score of 0.871, training area under the curve of 0.99, and test area under the curve of 0.88. **Conclusion:** This Transformer-based radiomics model effectively classifies lung adenocarcinoma subtypes, aiding early screening, diagnosis, and personalized treatment strategies to improve patient prognosis.

Keywords: Computed tomography, lung adenocarcinoma subtype analysis, radiomic characteristics, deep learning, transformer network model

Introduction

Cancer remains one of the most prevalent malignancies worldwide, with lung cancer exhibiting the highest incidence and mortality rates in China, showing an upward trend over recent decades [1-4]. Lung adenocarcinoma is the predominant histological type of lung cancer. According to the World Health Organization pathological classification, lung adenocarcinoma is categorized into pre-invasive lesions [e.g., atypical adenomatous hyperplasia, adenocarcinoma in situ (AIS)], minimally invasive adenocarcinoma (MIA), and invasive adenocarcinoma (IAC) [5-8].

Lung adenocarcinoma typically manifests as pulmonary nodules on CT. With continuous advancements in CT technology and deep learning, the screening and detection of pulmonary nodules have rapidly increased [9-12]. Deep learning algorithms enable precise medical image segmentation and feature extraction. Recent studies have shown that Transformer-based and Convolutional Neural Network models significantly improve the accuracy of semantic segmentation and predictive classification in medical imaging [13-15]. Jing and

Address correspondence to: He Ren, Faculty of Medical Instrumentation, Shanghai University of Medicine and Health Sciences; No. 279, Zhouzhu Highway, Pudong New Area, Shanghai 201318, China. Tel: +86-18817581363. E-mail: renh@sumhs.edu.cn.

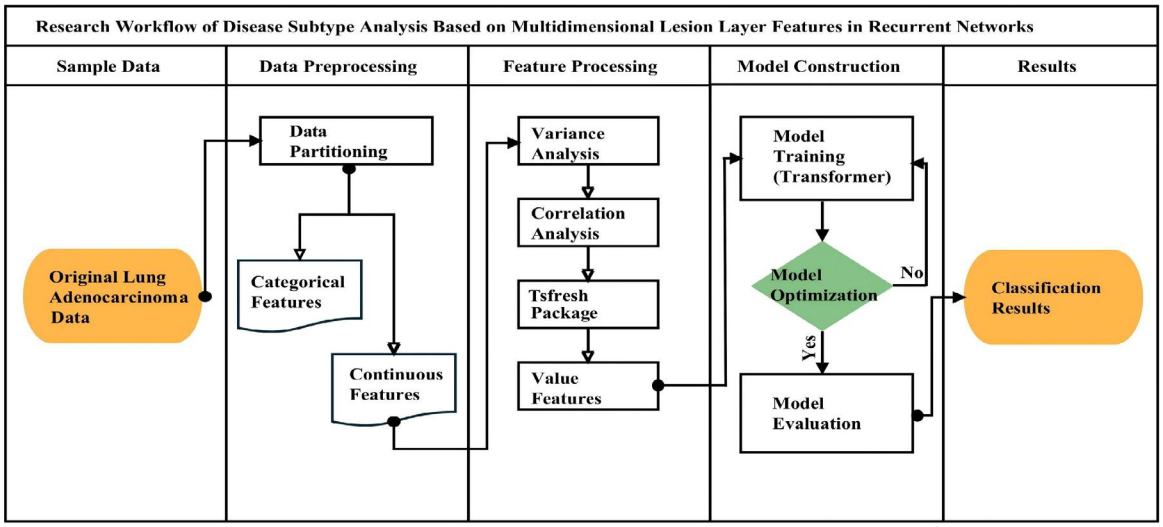


Figure 1. Flowchart of the proposed lung adenocarcinoma subtype classification framework.

Table 1. Baseline characteristics of the included patients

Parameters/Features	Contents
CT equipment	United Imaging 760 CT; Siemens SOMATOM Emotion 16 CT
Scanning parameters	Tube current: 42-126 mA/34-123 mA; Tube voltage: 120-130 kV; Slice thickness: 1 mm; Matrix: 512×512
Reconstruction algorithm	Standard reconstruction algorithm
Radiation dose	Conventional low-dose chest CT scans (42-126 mA; 34-123 mA)
Sex	113 males (39.1%); 176 females (60.9%)
Age	40-75 years (mean 56.3±9.2 years)
Smoking history	85 smokers (29.4%); 204 nonsmokers (70.6%)
Pathological classification	AIS: 50 cases (17.3%); MIA: 172 cases (59.5%); IAC: 67 cases (23.2%)

Note: CT, computed tomography; AIS, adenocarcinoma in situ; MIA, minimally invasive adenocarcinoma; IAC, invasive adenocarcinoma.

Hussein et al. analyzed CT images of pulmonary nodules using specialized software and showed that radiological and imaging characteristics can assist clinicians in differentiating disease subtypes, thereby enhancing model accuracy [16, 17] .

In this study, we adopted feature analysis, machine learning, and recurrent neural networks (**Figure 1**). By integrating imaging and radiological features from a lung adenocarcinoma dataset, we applied appropriate feature-selection algorithms to identify predictive features for disease subtype classification. Using targeted radiomics feature sequences, we developed a classification model that combined radiomics and Transformer networks. The integrated approach facilitated early screening and diagnosis of lung adenocarcinoma and supported clinicians in formulating differentiated treatment strategies for specific patient subtypes, thereby optimizing therapeutic outcomes and prognosis.

Materials and methods

Data source

This study analyzed low-dose chest CT images from 289 patients with histopathologically confirmed lung adenocarcinoma. The cohort comprised 172 cases of MIA, 50 cases of IAC, and 67 cases of AIS. All patients underwent low-dose chest CT scans using two systems: the United Imaging 760 CT (42–126 mA, 120 kV, 1 mm slice thickness, matrix 512×512) and the Siemens Emotion 16 CT (34–123 mA, 130 kV, 1 mm), both employing standard reconstruction algorithms and conventional low-dose modes. Baseline demographic and clinical variables included sex, age, smoking history, and pathological grade (**Table 1**).

The dataset comprised CT images from 289 patients, capturing multi-slice imaging features with continuous lesion layers. Lesions were initially segmented by radiologists, who placed seed points at the lesion centers. Subsequently, a region-growing algorithm was applied to automatically delineate three-dimensional lesion boundaries, and radiomics feature were extracted using specialized software. This process

Table 2. Selected radiomic features and their meanings

Raw Feature	Feature Meanings
Original_shape2D_PixelSurface	2D pixel surface shape feature of the original image.
Square_gldm_DependenceNonUniformity	Dependence nonuniformity derived from the gray level dependence matrix of the square-shaped image.
Original_firstorder_Energy	First-order energy feature of the original image.
Original_shape2D_MinorAxisLength	2D minor axis length shape feature of the original image.
Wavelet-LL_gldm_SmallDependenceHighGrayLevelEmphasis	Small dependence high gray level emphasis derived from the gray level dependence matrix of the LL sub-band of the wavelet-transformed image.
Original_firstorder_RobustMeanAbsoluteDeviation	First-order robust mean absolute deviation feature of the original image.
Wavelet-LL_firstorder_RobustMeanAbsoluteDeviation	First-order robust mean absolute deviation feature of the LL sub-band of the wavelet-transformed image.
Square_firstorder_InterquartileRange	First-order interquartile range feature of the square-shaped image.
Original_shape2D_MeshSurface	2D mesh surface shape feature of the original image.
Original_gldm_SmallDependenceHighGrayLevelEmphasis	Small dependence high gray level emphasis derived from the gray level dependence matrix of the original image.
Logarithm_glcm_Idn	Inverse Difference Normalized (Idn) derived from the gray level co-occurrence matrix of the logarithmic image.
Gradient_glrlm_ShortRunLowGrayLevelEmphasis	Short run low gray level emphasis derived from the gray level run length matrix of the gradient image.

yielded 963 radiomics features, including spatial relationships, morphological characteristics, density, and higher-order texture descriptors. Among these, 11 were categorical variables, while 952 were continuous. As the categorical features primarily described spatial and morphological characteristics with limited predictive relevance, subsequent feature processing focused on screening and dimensionality reduction of the continuous variables. High-contrast features were selected for model training to enhance objectivity and reproducibility.

Feature extraction

During the feature extraction phase, analysis of variance-was initially applied to screen variables. Columns with a variance<1 were removed to eliminate features contributing minimally to the results, leaving 515 continuous variables. Using the corr() function in Python for correlation analysis, we quantified inter-variable relationships and identified highly correlated or weakly informative features, reducing the dataset to 194 variables. Subsequently, the tsfresh package was used to generate time-series–based descriptors. The 1,667×194 feature matrix was processed with the extract_features algorithm and partitioned into 289 groups based on “group” and “group_id” columns. For each group of 194 variables, we derived maxima, autocorrelation, and other features, forming a new dataset of 289×150,336 dimensions. Subsequently, the select_features function was then applied to assess feature relevance and remove uninformative variables, retaining approximately 5% of features and producing a reduced dataset of 289×79,220.

Finally, a random forest algorithm was employed for feature ranking and optimization. Feature importance scores were computed with respect to the target outcome, enabling the identification of highly predictive variables. The top 15 most important features were selected and mapped back to the original dataset, forming the final optimized feature subset (**Table 2**).

Model architecture

A classification model was constructed based on the multi-head attention mechanism and Transformer network (**Figure 2**). Within this framework, multiple self-attention units were integrated, and their outputs were concatenated, enabling efficient parallel training.

By combining multi-head attention with position-wise feed-forward networks, the Transformer eliminated the need for additional convolutional or recurrent layers, simplifying the architecture while enhancing parallel computing efficiency without compromising model performance [18]. The parameters of the model are detailed in **Table 3**.

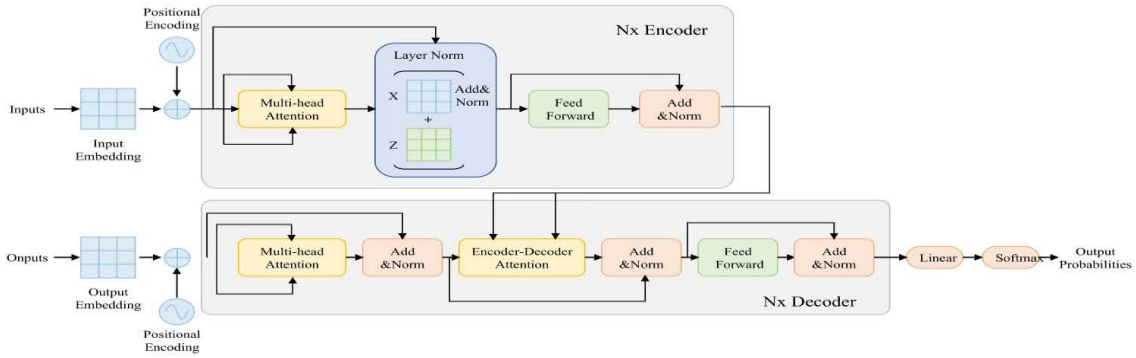


Figure 2. Multiple attention mechanisms and Transformer model diagrams.

Table 3. Model parameters and their specifications

Parameter	Description	Value
NUM_head	Number of heads in the multi-head attention mechanism	6
Dropout	Dropout rate to prevent overfitting	0.1
Hidden	Dimension of the hidden layer	512
LR	Learning rate (controls the update speed of parameters during training)	1e-3
Vocab_size	Vocabulary size	15
Embed_dim	Embedding dimension of the feature vectors	120
Epoch	Number of training epochs	150
Batch_size	Mini-batch size for training	64

The proposed model adopts an encoder-decoder architecture. The encoder integrates multi-head self-attention and position-wise feed-forward layers to extract high-dimensional dependencies from the input features. The decoder further refines these representations by incorporating encoder-decoder attention and projecting the output through a softmax classifier. Positional encoding helps preserve the sequential nature of radiomic data. This architecture effectively captures long-range dependencies and complex feature relationships derived from CT images, thereby supporting accurate subtype classification.

Assessment indicators

To comprehensively evaluate the segmentation performance of the model, the following key evaluation metrics were employed:

Accuracy

Accuracy reflects the proportion of predictions correctly classified by the model. It integrates true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN), thus indicating whether the overall predictions are correct. In lung adenocarcinoma subtype classification, higher accuracy indicates that the model correctly identifies a larger proportion of cases. The calculation formula is as follows:

$$\text{Accuracy}=\frac{1}{N}\sum_{i=1}^{N}\frac{(TP+TN)_i}{(TP+FP+TN+FN)_i} \quad (1)$$

Recall

Recall represents the proportion of samples belonging to a given positive class that are correctly identified by the model. In the deep learning-driven three-class classification of lung adenocarcinoma, when any subtype is designated as the positive class, a high recall indicates that the model can effectively capture its pathological features, thus reducing the risk of underdiagnosis. The formula is as follows:

$$\text{Recall}=\frac{1}{N}\sum_{i=1}^{N}\frac{(TP)_i}{(TP+FN)_i} \quad (2)$$

F1-Score

The F1-score is the harmonic mean of precision and recall. Since precision and recall may trade off against each other, relying on a single indicator may not fully reflect model performance. The F1-score integrates both measures, and a higher value indicates that the model achieves a better balance between precision and recall,

Table 4. Performance metrics of the Transformer model

	Accuracy	F1-score	Recall	AUC
Training	0.980	0.942	0.940	0.990
Validation	0.914	0.874	0.871	0.880

Note: AUC, area under the curve.

thereby providing a more comprehensive assessment of its ability to classify lung adenocarcinoma subtypes. The calculation formula is as follows:

$$\text{F1-score}=2\times\frac{\text{Precision}\times\text{Recall}}{\text{Precision}+\text{Recall}} \quad (3)$$

Area under the curve (AUC)

AUC refers to the area under the receiver operating characteristic curve, which plots the false positive rate (FPR) on the x-axis and the true positive rate (TPR) on the y-axis. The FPR and TPR are defined as:

$$\text{FPR}=\frac{\text{FP}}{\text{FP}+\text{TN}} \quad (4)$$

$$\text{TPR}=\frac{\text{TP}}{\text{TP}+\text{FN}} \quad (5)$$

AUC values range from 0.5 to 1.0, with larger values indicating a stronger ability to discriminate between classes. In the classification of lung adenocarcinoma subtypes, a higher AUC value indicates that the model can more effectively distinguish among the three lung adenocarcinoma subtypes.

Results

The performance metrics of the Transformer model, including accuracy (training_accuracy and val_accuracy), F1-score, recall, and AUC, are detailed in **Table 4**.

Figure 3A depicts the accuracy curves for the training and validation sets over 150 epochs, clearly illustrating the progression of model accuracy. Accuracy typically increases as training proceeds, eventually approaching a plateau. As expected, the training accuracy remains consistently higher than the validation accuracy.

Figure 3B illustrates the loss variation curve per epoch during Transformer model training. This curve tracks loss values at each epoch or batch during the training process, serving as a key metric for monitoring model performance and fine-tuning hyperparameters. In the initial phase, losses typically decrease as training iterations progress, demonstrating the model's gradual acquisition of more features and patterns. Subsequently, losses may plateau or fluctuate when the model reaches local optima and achieves optimal data fitting. If losses start to increase, this indicates either overfitting the training data or improper hyperparameter configuration.

In this study, **Figure 3C** and **Figure 3D** demonstrate the classification performance of the Transformer model across training and test datasets, as assessed by receiver operating characteristic curves and their AUC values. The receiver operating characteristic curve serves as a crucial metric for assessing model classification accuracy, and its AUC provides a comprehensive measure of model performance across various thresholds.

A higher AUC indicates stronger discriminative capability, while appropriate adjustment of the decision threshold can balance TPR and FPR, thereby improving prediction precision. The Transformer model achieved an AUC of 0.990 on the training set, demonstrating exceptional classification accuracy and strong learning capacity. On the validation set, the AUC reached 0.880-slightly lower but still indicative of satisfactory generalization to unseen data.

Overall, the Transformer model exhibited superior classification performance across both datasets, establishing a solid foundation for future research and clinical applications. Future studies could further investigate the impact of different threshold settings to optimize the model's classification effectiveness.

The Transformer model achieved a classification accuracy of 0.980 in the training set and 0.914 in the validation set. The corresponding recall rates were 0.942 and 0.874, F1-scores were 0.940 and 0.871, and AUC values were 0.990 and 0.880, respectively. Meanwhile, the loss function approached zero, indicating stable convergence. These findings demonstrate that the model combining multi-head attention and Transformer architecture delivers robust performance in classifying lung adenocarcinoma

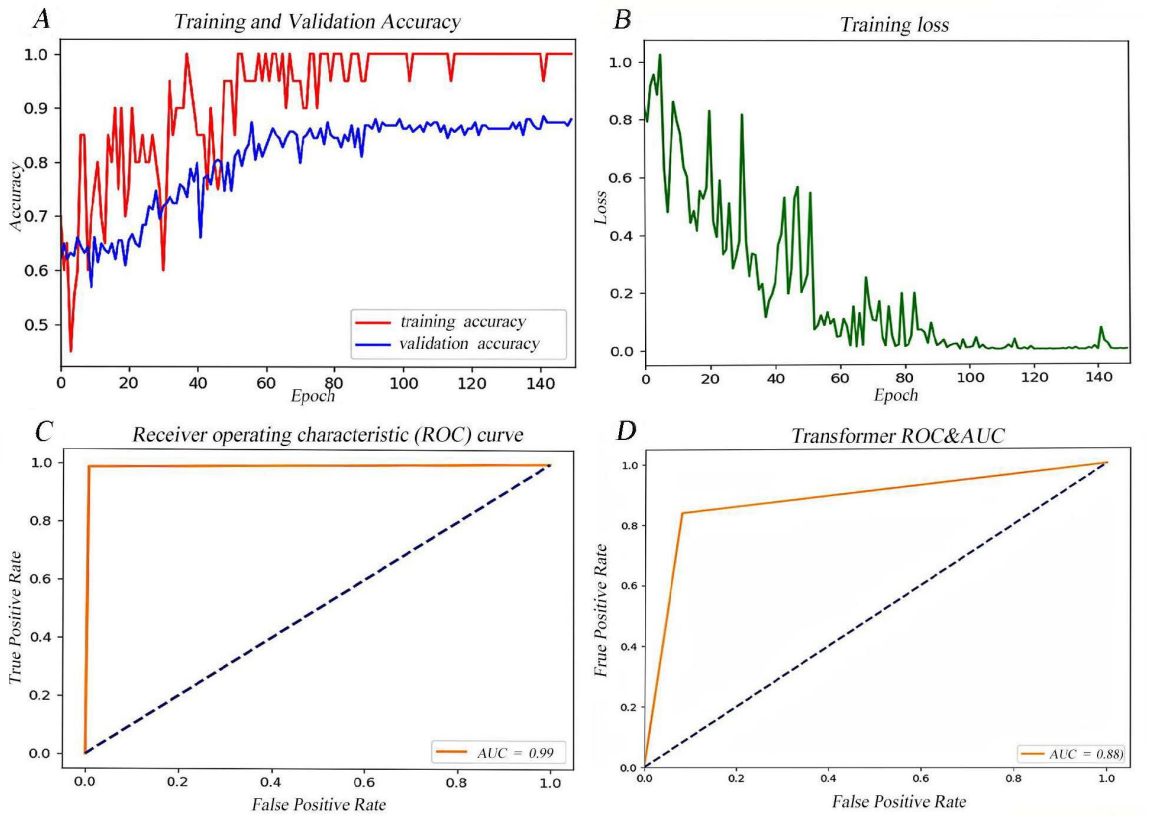


Figure 3. Performance evaluation of the Transformer model in training and validation sets. (A) Accuracy curves of training and validation sets across 150 epochs, showing steady improvement and convergence; (B) Training loss curve, indicating rapid reduction and stabilization with minimal overfitting; (C) ROC curve of the training set, demonstrating high classification capability with an AUC of 0.99; (D) ROC curve of the validation set, indicating good generalization with an AUC of 0.88. ROC, receiver operating characteristic; AUC, area under the curve.

subtypes from CT images. Moreover, the results suggest that variations in radiological features captured by CT imaging can effectively reflect lesion types, highlighting the model's potential as a valuable auxiliary diagnostic tool.

Discussion

This study presents a Transformer-based model incorporating multi-head self-attention mechanisms for classifying the three major subtypes of lung adenocarcinoma. By integrating radiomics and imaging features, the model achieved excellent performance on the training and validation sets, demonstrating high accuracy and robustness in lung adenocarcinoma subtype classification.

Previous research has reported encouraging results using deep learning for similar tasks. Ma et al. developed a deep learning model for predicting lymph node metastasis in lung adenocarcinoma, achieving AUC values of 0.948 and 0.961 in the training and validation sets, respectively [19]. Additionally, Liu et al. proposed an image-omics-based multi-subtype classification model for non-small cell lung cancer, achieving an accuracy of approximately 0.860 [20]. Although these studies highlight the potential of traditional methods and early-stage deep learning approaches, further optimization remains possible. The Transformer model introduced in this study effectively captures global contextual information and long-range dependencies, which is particularly important for identifying complex subtypes of lung adenocarcinoma. Furthermore, the multi-head self-attention mechanism enables the model to analyze feature relationships from multiple perspectives, enabling more precise differentiation among subtypes. This innovative approach demonstrates clear advantages in processing high-dimensional radiomic data, resulting in superior model performance compared to previous studies.

This study also has several limitations. First, the relatively small dataset may affect feature extraction and parameter optimization during radiological classification. Additionally, insufficient training samples are a critical issue, since high-quality predictive models require adequate data to prevent overfitting, particularly for deep learning architectures like Transformer net-

works, which typically demand massive training datasets. Thus, sample scarcity remains a key constraint on prediction accuracy. Second, the current approach relies exclusively on CT imaging for radiological feature extraction, which may compromise disease analysis precision due to limited diagnostic scope. While the model effectively classifies lesions through pattern recognition, feature selection variations could introduce bias. Furthermore, while trend analysis demonstrates potential for lesion identification, its limited applicability in clinical decision support, rule discovery, and interpretability may compromise diagnostic reliability. Collectively, these factors may reduce model accuracy and limit its effectiveness as an auxiliary diagnostic tool.

Clinical interpretability and practical significance

The radiomics model developed in this study not only achieves accurate classification of lung adenocarcinoma subtypes (AIS, MIA, IAC) but also demonstrates good clinical interpretability. In clinical practice, this model is designed to function as an auxiliary tool for radiologists rather than a replacement. By leveraging radiomics-derived risk scores and subtype predictions, radiologists can obtain quantitative information regarding lesion spatial structure, morphological features, density variations, and texture patterns, thereby enhancing their routine CT image evaluations.

By integrating the model's feature-importance rankings with imaging characteristics, physicians can more intuitively identify which radiological features most effectively distinguish pathological subtypes. For instance, higher-order texture features may indicate intra-tumor heterogeneity, while morphological parameters can reflect tumor margin infiltration. Such quantitative insights facilitate comprehensive decision-making in surgical planning, lesion grading, and follow-up strategy formulation. Furthermore, collaboration between AI models and clinicians enhances diagnostic efficiency. Physicians can utilize these models to rapidly screen suspected subtypes and verify the predictions using their expertise, fostering an "AI-assisted, physician-guided" diagnostic workflow. This approach promotes evidence-based decision-making while mitigating the risks of misdiagnosis associated with over-reliance on AI systems.

Future work

Future research will focus on addressing the current limitations of this study. In terms of data and sample size, we will expand the dataset through multi-institutional collaborations and apply data augmentation techniques to enrich training samples. These strategies will provide the model with a broader and more diverse learning foundation to effectively mitigate overfitting.

For image selection, we plan to incorporate multimodal imaging techniques such as MRI and X-ray for joint analysis, capturing lesion characteristics from multiple perspectives and improving diagnostic accuracy. Regarding feature selection, we will optimize screening algorithms and integrate clinical prior knowledge to enhance the effectiveness of feature-trend analysis in lesion identification, reducing the impact of feature variations on classification results. Additionally, we will prioritize improving the model's utility in clinical decision support by enhancing interpretability and mining decision rules, allowing the system to better assist radiologists and enhance its applicability in routine practice.

As an emerging classification method, multidimensional lesion feature analysis based on recurrent networks enhances prediction accuracy by integrating clinical and imaging features, thereby facilitating early cancer detection and precision therapy. This approach is poised for expanded applications in medicine, bioinformatics, computer vision, and related fields. With technological advancements, more efficient automated classification methods are expected to drive breakthroughs in disease diagnosis and therapeutic innovation.

Conclusion

This study developed a Transformer-based deep learning framework for classifying lung adenocarcinoma subtypes (MIA, AIS, IAC). The framework enhances sequence feature learning capabilities through multi-head attention mechanisms and fully connected layers, thereby improving model generalization and reducing overfitting. This approach enables effective training and accurate classification of radiomics features.

Experimental results demonstrate that the proposed Transformer model, integrating CT-derived radiomics features, exhibits superior performance in classifying lung adenocarcinoma subtypes. These findings validate both the diagnostic capabilities of imaging features and the clinical utility of the model. Meanwhile, radiomics provide diagnostic support from multidimensional perspectives, including morphology and

grayscale, while the sequential analysis of feature variations offers greater diagnostic value than isolated features, providing new insights for incorporating radiological characteristics into clinical decision-making.

Author contributions: Linrong Yuan, Yutong Xie, and Danhong Li jointly designed the study, developed the Transformer-based classification model, and drafted the manuscript. Jianghui Li, Miao Yu, Siqi Wang and Yu Wang were responsible for CT image preprocessing, radiomic feature extraction, model training, and performance evaluation. He Ren supervised the project, provided critical revisions, and approved the final version of the manuscript.

References

[1] Chen W, Zhang S, Zou X. Zhongguo Fei Ai Za Zhi 2010;13(5):488-493.

[2] Han B, Zheng R, Zeng H, et al. Cancer incidence and mortality in China, 2022. J Natl Cancer Cent 2024;4 (1):47-53.

[3] Li H, Zhao M, Fei G, et al. Epidemiological trends and incidence prediction of lung cancer in China based on the Global Burden of Disease study 2019. Front Med (Lausanne) 2022;9:969487.

[4] Guo L, Zhu C, Cai L, et al. Global burden of lung cancer in 2022 and projected burden in 2050. Chin Med J (Engl) 2024;137 (21):2577-2582.

[5] Li X, Zhang W, Yu Y, et al. CT features and quantitative analysis of subsolid nodule lung adenocarcinoma for pathological classification prediction. BMC Cancer 2020;20(1):60.

[6] Chen H, Carrot-Zhang J, Zhao Y, et al. Genomic and immune profiling of pre-invasive lung adenocarcinoma. Nat Commun 2019;10(1):5472.

[7] Su H, Chen L, Wu J, et al. Proteogenomic characterization reveals tumorigenesis and progression of lung cancer manifested as subsolid nodules. Nat Commun 2025;16(1):2414.

[8] Wu Z, Zhao Y, Peng Y, et al. LncRNA TMEM99 Complexes with IGF2BP2 to Inhibit Autophagy in Lung Adenocarcinoma. Adv Sci (Weinh) 2025;12(33):e202507871.

[9] Wang J, Lin L, Zhao S, et al. Sheng Wu Yi Xue Gong Cheng Xue Za Zhi 2019;36(4):670-676.

[10] Zhou J, Hu B, Feng W, et al. An ensemble deep learning model for risk stratification of invasive lung adenocarcinoma using thin-slice CT. NPJ Digit Med 2023;6(1):119.

[11] Shuvo S, Mamun TB. AutoLungDx: A Hybrid Deep Learning Approach for Early Lung Cancer Diagnosis Using 3D Res-U-Net, YOLOv5, and Vision Transformers. 2023.

[12] Yu P, Zhang H, Wang D, et al. Spatial resolution enhancement using deep learning improves chest disease diagnosis based on thick slice CT. npj Digit Med 2024;7(1):335.

[13] Kuang H, Wang Y, Liu J, et al. Hybrid CNN-Transformer Network With Circular Feature Interaction for Acute Ischemic Stroke Lesion Segmentation on Non-Contrast CT Scans. IEEE Trans Med Imaging 2024;43(6):2303-2316.

[14] Azad B, Adibfar P, Fu K. TransDAE: Dual Attention Mechanism in a Hierarchical Transformer for Efficient Medical Image Segmentation. arXiv arXiv:2409.02018, 2024.

[15] Zhang Y, Xi R, Wang W, et al. Low-Contrast Medical Image Segmentation via Transformer and Boundary Perception. IEEE Trans Emerg Top Comput Intell 2024;8(3):2297-2309.

[16] Jing R, Wang J, Li J, et al. A wavelet features derived radiomics nomogram for prediction of malignant and benign early-stage lung nodules. Sci Rep 2021;11(1):22330.

[17] Hussein S, Gillies R, Cao K, et al. TumorNet: Lung nodule characterization using multi-view Convolutional Neural Network with Gaussian Process. 2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017) 2017;1007-1010.

[18] Papagiannopoulou A, Angeli C. Social Media Text Summarization: A Survey Towards a Transformer - based System Design. Adva Sci, Technol Eng Syst J 2023;8(6):26-36.

[19] Ma X, Xia L, Chen J, et al. Development and validation of a deep learning signature for predicting lymph node metastasis in lung adenocarcinoma: comparison with radiomics signature and clinical - semantic model. Eur Radiol 2023;33(3):1949-1962.

[20] Liu J, Cui J, Liu F, et al. Multi-subtype classification model for non-small cell lung cancer based on radiomics: SLS model. Med Phys 2019;46(7):3091-3100.

Progress in Medical Devices
ISSN: 2957-5478

Zentime

Review Article https://doi.org/10.61189/220864jklusg

Application of recurrent laryngeal nerve monitoring technology in thyroid and parathyroid surgery

Yong Wang, Yu Zhou

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

Corresponding author: Yu Zhou.

Declaration of conflict of interest: None.

Received December 8, 2024; Accepted February 26, 2025; Published September 30, 2025

Highlights

- A systematic introduction to the electrophysiological principles of recurrent laryngeal nerve (RLN) monitoring, the principles of RLN monitoring equipment, and the main RLN monitoring devices currently available.
- A discussion of the standardized procedures for RLN monitoring, along with methods for managing abnormal conditions during surgery.
- An analysis of the development status and future prospects of RLN monitoring technology.

Abstract

Protecting the recurrent laryngeal nerve (RLN) and superior laryngeal nerve during thyroid and parathyroid surgery remains a significant challenge. Traditional methods primarily rely on visual identification and regional protection techniques to minimize nerve injury. However, these approaches often face challenges such as limited working space, high procedural difficulty, and incomplete tissue removal. Intraoperative nerve monitoring uses neurophysiological techniques to assess the functional integrity of nerves, aiming to prevent or reduce nerve damage. Intraoperative nerve monitoring for laryngeal nerve protection during thyroid and parathyroid surgery provides an effective means of evaluating RLN damage. For successful RLN monitoring, both monitoring personnel and surgeons need a solid understanding of nerve monitoring principles, follow of standardized surgical procedures, and be able to troubleshoot abnormal signals during surgery. With ongoing advancements in technology, nerve monitoring devices are expected to become more sensitive, offering rapid and precise waveform analysis, and enhancing user-friendliness. Additionally, minimally invasive thyroidectomy and robot-assisted surgical systems hold promising potential for the future of thyroid surgery. This paper reviews the use of Intraoperative nerve monitoring and RLN monitoring, incorporating the latest research from both domestic and international studies. It discusses the importance of RLN monitoring, the principles of monitoring technologies, current research on RLN monitoring technology, guidelines for nerve monitoring, and strategies for managing and analyzing abnormal monitoring signals.

Keywords: Intraoperative nerve monitoring, recurrent laryngeal nerve, thyroid

Introduction

Thyroid cancer has become one of the fastest-growing solid malignancies worldwide, with rising incidence and prevalence. According to a 2013 analysis of data from 255 cancer registries conducted by the National Cancer Registry Center, thyroid cancer ranks as the seventh most common cancer overall and the fifth most common among women with a significant upward trend in its incidence [1]. Recurrent laryngeal nerve (RLN) injury remains one of the most common and serious complications associated with thyroidectomy. Several factors contribute to RLN injury, including age, differentiated thyroid carcinoma, and skin color. The overall incidence of RLN injury within 30 days following thyroid surgery is reported to be 6.0% [2]. RLN injury is widely recognized as a major complication of thyroid surgery. In China, the incidence of intraoperative RLN injury ranges from 2.23% to 7.40% [3].

Intraoperative neurophysiological monitoring

Address correspondence to: Yu Zhou, School of Health Sciences and Engineering, University of Shanghai for Science and Technology, No.516 Jungong Road, Shanghai 200093, China. Tel: +86-18021042556. E-mail: zhouyu@usst. edu.cn.

(IONM) applies electrophysiological principles to assess nerve function through electrical stimulation, which induces nerve impulses that travel along the monitored nerve, triggering muscle contractions and generating electromyography (EMG) signals [4]. The advantages of IONM include enhanced nerve identification and the prevention of nerve injury. IONM is particularly valuable in protecting motor nerves during thyroid and parathyroid surgeries, especially in high-risk procedures involving nerve damage, minimally invasive laparoscopic surgeries with limited working space, and surgeries where vocal function preservation is critical [5, 6].

Although RLN monitoring offers numerous technical advantages, its effectiveness depends on proper implementation. Surgical staff must undergo specialized training in laryngeal nerve recording to ensure accurate functional assessment and provide effective assistance to the surgeon. This article aims to introduce key aspects of RLN monitoring in thyroid surgery to help clinicians better understand, standardize, and appropriately apply RLN monitoring technology in clinical practice.

Significance of RLN monitoring

Protecting the RLN has always been a key focus in thyroid surgery. Since the late 19th century, surgeons have employed various anatomical techniques to safeguard the RLN, requiring meticulous dissection and leaving a portion of thyroid tissue to protect the nerve. These techniques include the “nerve palpation protection technique” and the “anatomical exposure technique.” The nerve palpation technique involves gently identifying and isolating the RLN through tactile exploration during surgery, while the anatomical exposure technique requires detailed dissection to directly visualize and protect the nerve. Both approaches aim to minimize the risk of RLN injury during thyroidectomy, which is essential for preserving vocal function and preventing postoperative complications. However, incomplete removal of surrounding tissues poses a risk of requiring multiple surgeries to excise tumors, which in turn increases the likelihood of RLN injury.

In 1966 and 1970, Shedd and Flisberg independently proposed the use of nerve monitoring technology during thyroid surgery to help surgeons assess RLN function and detect potential damage during procedures [7, 8]. Since then, RLN monitoring has gained increasing attention from surgeons.

The RLN is a branch of the vagus nerve that emerges after the nerve enters the thoracic cavity. The terminal branches of the RLN innervate the mucosa below the glottis and all laryngeal muscles except the cricothyroid muscle, making it the primary motor nerve for these muscles. Clinically, unilateral RLN injury can result in unilateral vocal cord paralysis, while bilateral RLN damage may lead to aphonia, respiratory distress, and even suffocation [9].

In addition, a rare anatomical variation known as the nonrecurrent laryngeal nerve (NRLN) exists, where the right RLN arises directly from the main trunk of the vagus nerve in the upper neck, bypassing its usual course and entering the larynx near the lower pole of the thyroid. This variation is often associated with the right subclavian artery, which originates from the left side of the aortic arch. The incidence of this variation is higher on the right side than on the left [10].

Relevant data indicate that the incidence of NRLN ranges from 0.5% to 1.0% [11]. While the likelihood of this variation is relatively low, it remains a significant factor contributing to RLN injury. Failing to recognize the presence of an NRLN increases the risk of RLN damage during surgery [12, 13]. Therefore, a thorough preoperative assessment and understanding of the anatomical and physiological structure of the NRLN are crucial.

Principle of RLN monitoring

RLN monitoring technology operates by using a stimulation electrode to generate an electrical current of specific intensity. This electrical stimulation induces nerve impulses in the motor neurons, which are transmitted to the innervated muscles, causing muscle contractions and generating EMG signals. The monitoring electrode, in contact with the muscle, detects these EMG signals and transmits them back to the nerve-monitoring device for amplification and processing. The processed signals are displayed as an EMG trace and accompanied by an auditory alert, depending on the results [14].

The nerve monitoring system primarily consists of the monitoring device, stimulation equipment, recording equipment, and other related devices. A schematic diagram of the monitoring process is presented in **Figure 1**.

Monitoring device main unit

The main unit of the monitoring device is the core component of the nerve monitoring system. It generates electrical stimulation sig-

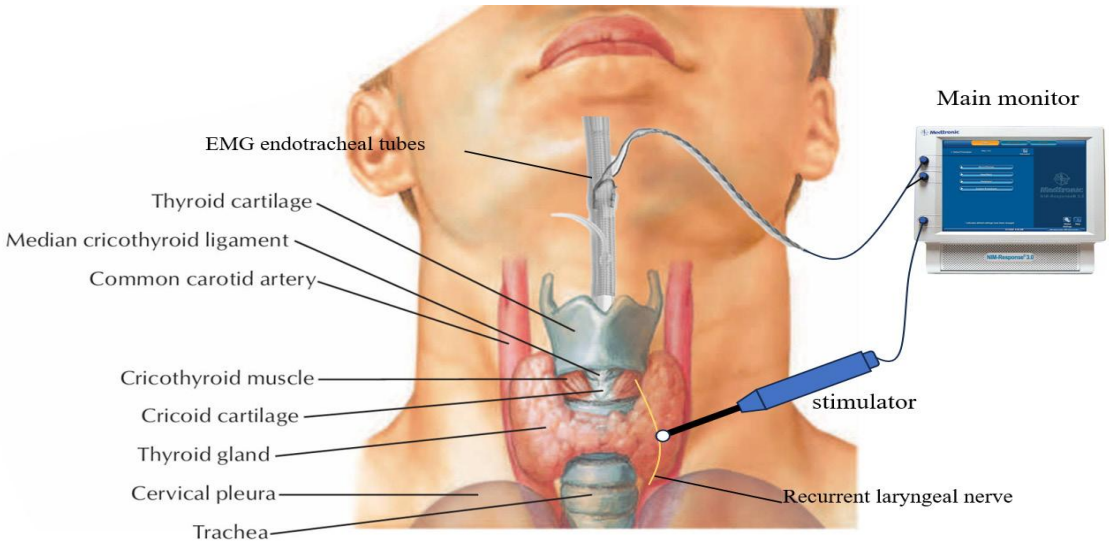


Figure 1. RLN monitoring. RLN, recurrent laryngeal nerve; EMG, electromyography.

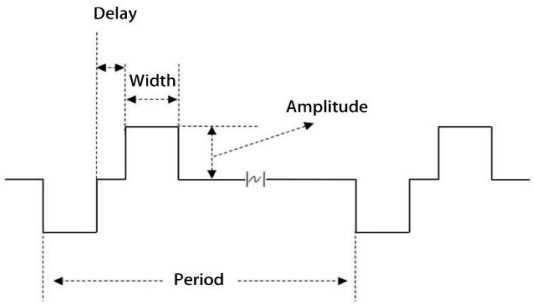


Figure 2. Pulse schematic diagram.

nals, processes EMG signals, issues alerts, and displays and stores information. Several parameters of the main unit can be adjusted during operation, including the event threshold, stimulation current intensity, stimulation frequency, and pulse width. The event threshold refers to the minimum amplitude that triggers a response from the system and can be set based on surgical requirements. Establishing an appropriate event threshold is crucial for optimizing early warning effectiveness in nerve injury detection.

The most widely used intraoperative nerve monitoring devices include the Medtronic series, Inimed series, and Dr. Langer Medical series.

(1) Medtronic Series: The Medtronic series includes three products: NIM-Response 3.0, NIM-Neuro 3.0, and NIM Vital.

NIM-Response 3.0 is specifically designed for RLN monitoring during thyroid and neck surgeries. It supports unilateral nerve monitoring and provides both audio and visual feedback on nerve function.

NIM-Neuro 3.0 offers multi-channel functionality, making it ideal for complex, high-risk surgeries.

NIM Vital (introduced in 2020) features a wireless design for both the patient interface and the console, making it portable and easy to position.

(2) Inimed Series: The Inimed series includes the C2 NerveMonitor and ISIS IOM System.

C2 NerveMonitor is designed for RLN monitoring, supporting bilateral stimulation and multi-channel recording.

ISIS IOM System is a comprehensive intraoperative nerve monitoring system that can monitor the RLN, facial nerve, and peripheral nerves simultaneously, making it suitable for complex surgeries, including neurosurgical applications.

(3) Dr. Langer Medical Series: The Dr. Langer Medical series includes the Neurosign 400 and Neurosign V4.

Neurosign 400 is specifically designed for RLN monitoring, allowing unilateral nerve monitoring with rapid nerve stimulation and feedback.

Neurosign V4 is a multifunctional nerve monitoring device with multi-channel capability, supporting various types of nerve monitoring.

Principle of electrical stimulation

The stimulating device, also known as the stimulation electrode, delivers electrical currents to depolarize motor nerve fibers, generating ac-

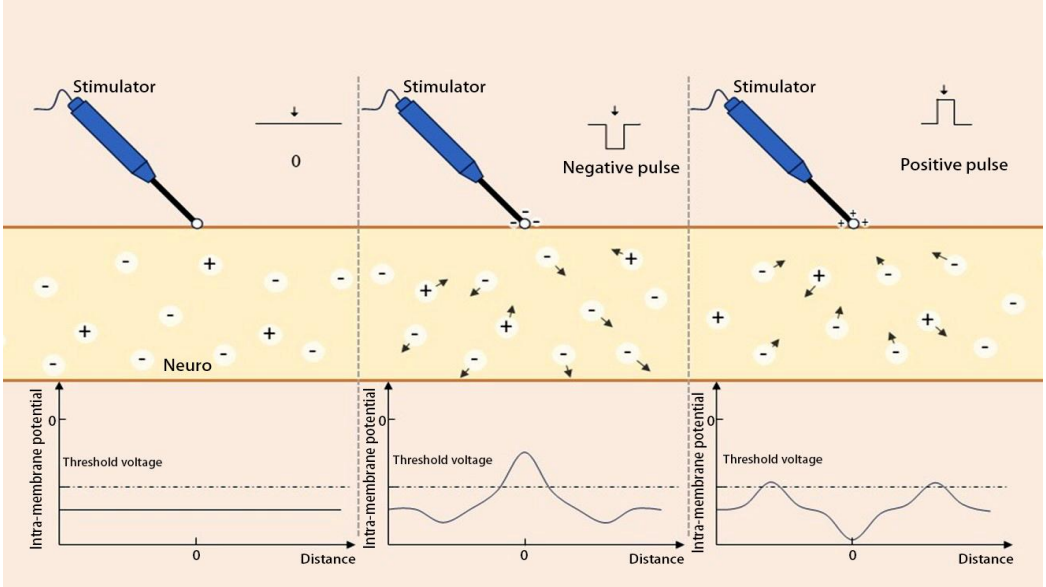


Figure 3. Schematic of a biphasic electrical stimulation pulse.

tion potentials that are transmitted to the vocal cord muscles innervated by the nerve. Stimulation electrodes come in two main types: monopolar and bipolar, with various probe designs, including spherical, needle-shaped, and hook-shaped electrodes. These electrodes differ in size, shape, electrode type, and contact form, allowing selection based on specific needs. Additionally, combining anatomical or endoscopic instruments with stimulation probes has gained attention for enhancing nerve monitoring precision during surgery. Wu et al. used ten different stimulation probes to stimulate the laryngeal nerves and monitor nerve function, demonstrating that monopolar electrodes, bipolar electrodes, and stimulating anatomical instruments are all effective in inducing nerve impulses, thus supporting RLN monitoring [15].

Electrical stimulation can be categorized into monopolar and bipolar modes based on the type of electrode used. In monopolar stimulation, the current is emitted from the stimulator, passes through tissue or nerve, and returns to the stimulation device via the return electrode plate. Due to the relatively large distance between the stimulator and the return electrode plate, current distribution in monopolar stimulation is more diffuse, making it suitable for locating the general course of nerves. In contrast, bipolar stimulation involves current flow between two probes placed a few millimeters apart, resulting in a more focused current distribution, which is particularly effective for precise nerve localization [15].

Electrical stimulation that induces nerve impulses can be generated using either a constant-voltage source or a constant-current source. The latter is preferred due to its stable current delivery, controllable stimulation effects, and reduced risk of overstimulation. Key parameters involved in electrical stimulation include pulse intensity, pulse width, pulse frequency, and stimulation waveform [16]. For RLN stimulation, the typical current range is 0-5 mA, with a pulse width of approximately 50-200 μs, a pulse interval of 100 μs, and a frequency of 1-4 Hz. These pulse parameters are illustrated in **Figure 2**.

Both positive and negative pulses can induce nerve impulses, though the mechanisms differ slightly. During negative pulse stimulation, the negative charge on the electrode causes a redistribution of charges on the axonal membrane [17]. Negative charges accumulate on the membrane's exterior beneath the electrode, while positive charges inside the membrane move toward the electrode. This leads to depolarization beneath the electrode and hyperpolarization farther from it. When the membrane potential reaches the threshold voltage, an action potential is triggered. Positive pulses induce the opposite effect: positive charges accumulate on the electrode, moving negative charges inside the membrane toward the electrode, leading to depolarization further from the electrode and hyperpolarization near it. When the membrane potential reaches threshold, an action potential is triggered. The action potentials induced by both positive and negative pulses are illustrated in **Figure 3**.

Regarding stimulation waveforms, the main types are monophasic pulse, symmetric biphasic pulse, and asymmetric biphasic pulse [16]. Domestic researchers have verified various

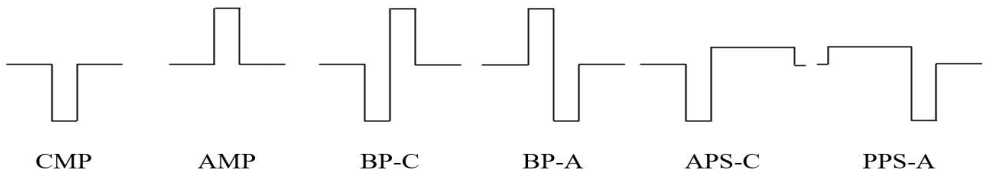


Figure 4. Schematic of stimulation waveform. CMP, monophasic pulse; AMP, anodic monophasic pulse; BP-C, cathodic first charge-balanced biphasic pulse; BP-A, anodic first charge-balanced biphasic pulse; APS-C, cathodic first charge-balanced anterior pseudomonophasic pulse; PPS-A, anodic first charge-balanced posterior pseudomonophasic pulse.

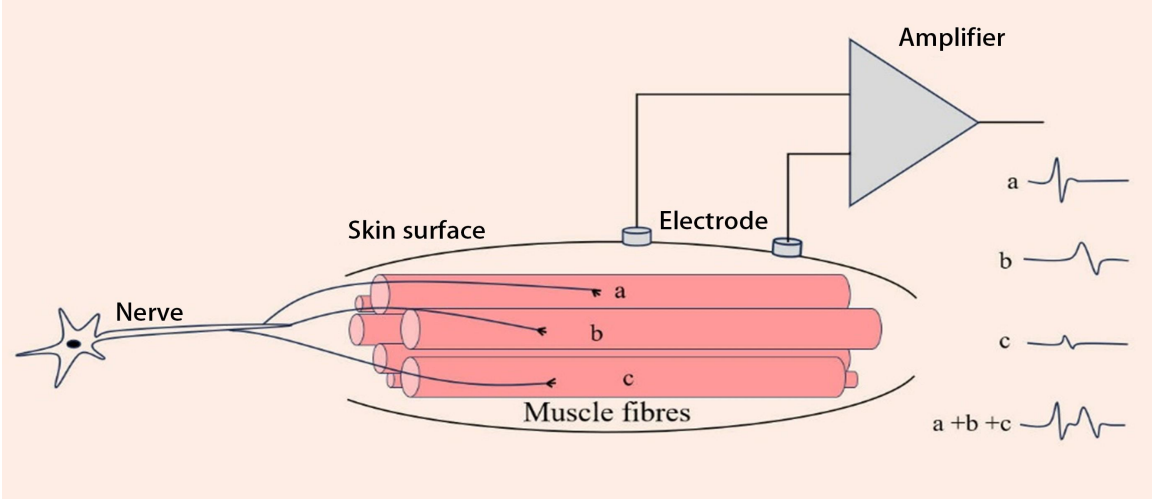


Figure 5. Schematic diagram of muscle fiber formation of EMG signal. EMG, electromyography.

pulse signals, including the Cathodic Monophasic Pulse, Anodic Monophasic Pulse, Cathodic First Charge Balanced Biphasic Pulse, Anodic First Charge Balanced Biphasic Pulse, Cathodic First Charge Balanced Anterior Pseudomonophasic Pulse, and Anodic First Charge Balanced Posterior Pseudomonophasic Pulse. The excitation induced by these waveforms, from largest to smallest, is as follows: Cathodic Monophasic Pulse, Cathodic First Charge Balanced Anterior Pseudomonophasic Pulse, Cathodic First Charge Balanced Biphasic Pulse, Anodic First Charge Balanced Posterior Pseudomonophasic Pulse, Anodic First Charge Balanced Biphasic Pulse, and Anodic Monophasic Pulse. These waveforms are illustrated in **Figure 4**. For RLN stimulation, biphasic pulse stimulation is generally preferred to avoid charge imbalance on the nerve and prevent electrode corrosion. The first phase of a biphasic pulse is a negative pulse to activate neurons, while the second phase is a positive pulse to balance the charge.

Principle of EMG signal acquisition

The purpose of the recording device is to capture EMG signals from muscles innervated by the nerve. Neuronal discharge activates numerous muscle fibers, and the cumulative activity of these fibers generates the EMG signal. As the force exerted by the body increases, more muscle fibers are recruited, resulting in a corresponding increase in the amplitude of the EMG signal [18]. The EMG signal is shown in **Figure 5**. By analyzing the waveform, amplitude, and latency of the EMG signal, nerve damage can be assessed. A detailed description of the EMG signal is provided in **Figure 6**.

The EMG signal can be captured using surface electrodes on the monitoring catheter in direct contact with the vocal cords or by inserting a needle electrode into the lower and lateral parts of the thyroid cartilage's anterior horn. Zhao et al. compared three types of recording electrodes using pigs: monitoring catheter surface electrodes, thyroid cartilage adhesive electrodes, and thyroid cartilage needle electrodes [19]. Their findings indicated that the monitoring performance of all three electrodes was nearly identical. However, non-invasive, indwelling monitoring via catheter electrodes is more commonly used in clinical practice [20].

To visualize the EMG signal, it is important to recognize that EMG signals are weak bioelectrical signals with small amplitudes (1 mV to 10 mV) and a frequency range of 0-500 Hz. These signals are highly susceptible to external interference, such as power-line noise,

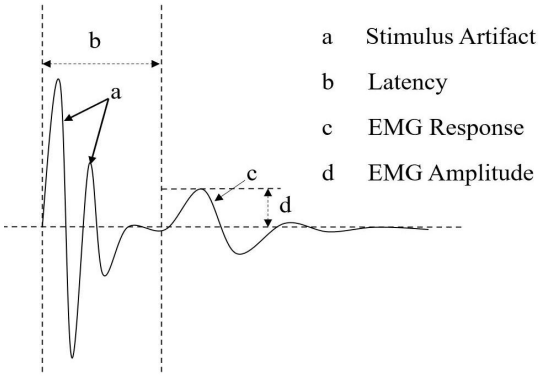


Figure 6. Schematic diagram of EMG signal description. EMG, electromyography.

equipment-generated noise, and movement artifacts. Therefore, the raw EMG signal must be amplified several thousand times and filtered to remove unwanted noise [18]. In the circuit, an instrumentation amplifier is used with a main amplifier to amplify the EMG signal. The instrumentation amplifier provides a high common-mode rejection ratio (CMRR), initially amplifying the signal by several tens of times. The main amplifier further amplifies the signal by tens to hundreds of times. A filtering circuit, including high-pass and low-pass filters, isolates the primary frequency range of the EMG signal. A notch filter is employed to eliminate power-line interference. Finally, the processed EMG data is converted into a digital signal and processed by a microcontroller [21]. The flowchart of the EMG signal acquisition process is shown in **Figure 7**. Other devices, such as line connection devices, storage equipment, and printing devices, are also used.

Current research status of RLN monitoring technology

RLN monitoring technology has been widely adopted worldwide, and decades of development have led to continuous innovation in this field. Current advancements in RLN monitoring technology focus on four main areas: the stimulation end, signal acquisition, monitoring instrument software, and application methods.

Innovation in the stimulation end involves integrating anatomical instruments with stimulation probes, resulting in devices like the stimulating dissector. This tool allows surgeons to seamlessly switch between anatomical instruments and stimulation probes, improving convenience and efficiency [22-24]. The most common signal acquisition method is using an endotracheal tube with surface electrodes, which is non-invasive and easy to use. Recent innovations include the thyroid cartilage needle electrode, which indirectly contacts the effector muscles, offering more precise signal acquisition with enhanced sensitivity [19, 25].

Innovations in monitoring instrument software aim to improve signal sensitivity, automate waveform analysis, enhance anti-interference capabilities, and provide user-friendly features.

Regarding application methods, two key innovations have emerged. The first is minimally invasive thyroidectomy assisted by endoscopy, using endoscopic forceps equipped with nerve monitoring capabilities. This procedure can be performed via axillary or bilateral axillary-breast approaches or through natural orifice surgery. Minimally invasive thyroidectomy offers high cosmetic satisfaction and comparable surgical outcomes and complication rates to traditional thyroidectomy [26, 27]. The second innovation is robot-assisted thyroidectomy, which leverages the precision and flexibility of robotic systems to overcome the limitations of traditional and endoscopic thyroidectomy, offering new directions for surgical innovation [28, 29].

Monitoring process

The monitoring process is complex and requires collaboration among various professionals, including the surgeon, anesthesiologist, technical experts, and neuro-monitoring staff. Each plays a specific role and contributes to the process. If neuro-monitoring personnel lack proper training or if neuro-monitoring procedures are not standardized, it can lead to inaccurate neurophysiological data, potentially misleading the surgeon. Therefore, adherence to standardized neuro-monitoring protocols, accurate analysis of abnormal conditions, and the identification of appropriate patient populations are essential for maximizing the benefits of this technique.

Steps of RLN monitoring technique

To ensure successful RLN monitoring, standardized operational procedures must be followed. The international guidelines for RLN monitoring, published in 2011 and 2013, outline the steps for intraoperative nerve monitoring, including preoperative laryngoscopy, preoperative vagus nerve identification, postoperative vagus nerve detection, and postoperative laryngoscopy [20, 30].

The Thyroid Surgery Committee of the Chinese Medical Association's Surgical Physician Branch has published clinical guidelines for intraoperative nerve monitoring during thyroid and parathyroid surgeries in 2013, 2017, 2019,

Table 1. RLN monitoring steps

Abbreviations	Specific steps
L1	Preoperative Laryngeal Examination
V1	Preoperative Vagus Nerve Detection
R1	After precisely locating the RLN, detect the RLN
S1	Pre-Thyroid Upper Pole Vascular Handling: Detection of the External Branch of the Superior Laryngeal Nerve
S2	Post-Thyroid Upper Pole Vascular Handling: Detection of the External Branch of the Superior Laryngeal Nerve
R2	Postoperative Detection of the RLN
V2	Postoperative Detection of the vagus nerve
L2	Postoperative Laryngeal Examination

Note: RLN, Recurrent Laryngeal Nerve.

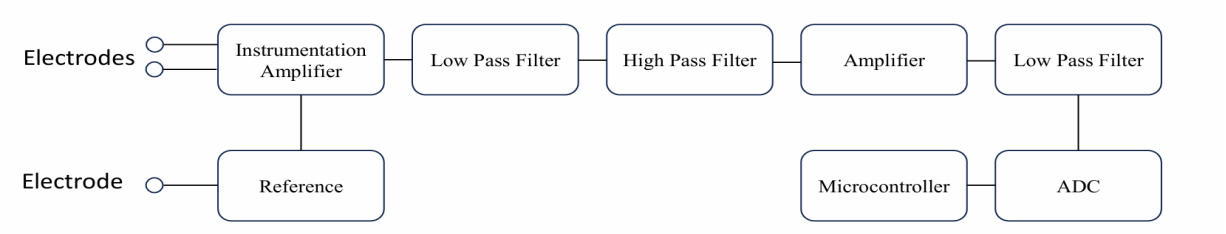


Figure 7. EMG signal acquisition. ADC, analog-to-digital converter; EMG, electromyography.

and 2023. These guidelines define standardized procedures for nerve monitoring and offer recommendations.

Since the release of the first Chinese guidelines in 2013, the procedures for RLN monitoring have gained widespread recognition in the medical community. The monitoring process has evolved from an initial six-step method to an eight-step method (**Table 1**). These steps include: preoperative laryngeal examination (L1), preoperative vagus nerve detection (V1), RLN detection (R1), detection of the external branch of the superior laryngeal nerve before handling the thyroid upper pole vessel (S1), detection of the external branch of the superior laryngeal nerve after handling the thyroid upper pole vessel (S2), postoperative RLN detection (R2), postoperative vagus nerve detection (V2), and postoperative laryngeal examination (L2).

By comparing signals before and after surgery, the function of the nerve can be assessed, providing valuable information for subsequent diagnosis and treatment.

Analysis and handling of abnormal EMG signals

Abnormal EMG signals are characterized by a significant decrease in the amplitude of the EMG signal during surgery, compared to the normal signals obtained preoperatively. In a dry environment, with an electrical stimulation intensity of 1-2 mV, if the amplitude of the EMG signal drops below 100 μV and there is no laryngeal or glottic twitching, it is considered as the loss of the EMG signal [31]. Abnormal EMG signals may be caused by issues with the monitoring system or nerve damage. When abnormal signals are detected, it is crucial to identify the cause quickly and take appropriate measures to minimize RLN damage and adjust the surgical plan if necessary.

During surgery, the first step when abnormal EMG signals are detected is for the surgeon to assess the laryngeal spasm response. This involves stimulating the ipsilateral vagus nerve and observing the laryngeal muscle tremor. The surgeon should palpate the cricothyroid muscle at the posterior aspect of the cricoid cartilage to monitor for muscle contraction. If a laryngeal spasm response is observed, it indicates that the stimulation system and neural function are intact, suggesting the abnormal signal may be due to issues with the recording equipment. The most common problem with recording equipment is electrode misplacement, which can occur due to catheter displacement, improper depth placement, or poor contact. In such cases, the monitoring catheter should be repositioned.

It is also important to check for poor contact with the grounding electrode. Sweat accumulation may cause the grounding electrode to detach or shift. If no laryngeal spasm response is observed, the issue may lie with the stimulation system. The first step is to verify whether the stimulation current is sufficient. If the current is abnormal, it could be due to a malfunc-

tion of the stimulation electrode or incorrect stimulation channel settings. In this case, the stimulation electrode should be replaced or the stimulation channel adjusted. If the stimulation current is normal, the issue could be caused by excess fluid in the surgical field, leading to current dispersion. In such cases, the surgical field should be dried.

Neural injury analysis and management

If the initial EMG signal is normal and a sufficiently strong stimulation current is applied to the nerve, but the EMG signal becomes abnormal or disappears without any laryngeal spasm response under laryngoscopy, a thorough review of the lost EMG signal is required. If no issues are found with the monitoring system, the surgeon should strongly suspect nerve injury. Neural injuries detected through IONM can be categorized into two types. Type I injury, also known as segmental injury, refers to the identification of a specific point or site of injury in the RLN. This type of injury is most commonly caused by excessive traction, clamping, or thermal injury, with traction being the most frequent cause [4, 32]. Type II injury, or complete injury, occurs when no EMG signals are detected in both the RLN and the ipsilateral vagus nerve, but signals are present in the contralateral RLN and vagus nerve. In Type II injury, no specific point or segment of injury can be identified, and the underlying mechanism remains unclear.

Traction-induced RLN injury is the most common form, with the EMG signal gradually decreasing in proportion to the extent of traction. If detected early and addressed promptly, the EMG signal may recover. If the RLN signal drops to 50% or less of its normal amplitude, this serves as an early warning of potential nerve injury [33]. In such cases, surgery should be paused, and a 20-30-minute observation period should be allowed to assess whether the signal recovers, while the possible causes are analyzed. If the EMG signal recovers, surgery can resume, but extra caution should be taken to avoid further injury. If the EMG signal does not recover, it may be necessary to adjust the surgical plan accordingly [34].

Summary

IONM technology provides an effective means of assessing RLN injury during surgery, offering valuable support to surgeons. However, challenges remain in the application of RLN monitoring technology: monitoring personnel must receive specialized training, surgeons must be skilled in standardized procedures, and they must be able to troubleshoot abnormal signals effectively. IONM technology requires rapid and precise physiological signal acquisition and response, imposing stringent demands on both hardware and software. To meet these challenges, nerve monitoring devices must improve sensitivity, automate waveform analysis, and enhance anti-interference capabilities, minimizing the risk of intraoperative nerve injury.

As technology continues to advance, nerve monitoring equipment is expected to evolve toward greater sensitivity, faster and more accurate waveform analysis, and enhanced user-friendliness. Moreover, minimally invasive thyroidectomy and robot-assisted thyroidectomy are likely to be key areas of focus for the future development of thyroid surgery.

Author contributions: Yong Wang was responsible for collecting data, writing the paper, revising the content, and proofreading the paper. Yu Zhou was responsible for providing direction for the paper and guiding the writing.

References

[1] Chen W, Zheng R, Zhang S, et al. Cancer incidence and mortality in China in 2013: an analysis based on urbanization level. Chin J Cancer Res 2017;29(1):1-10.

[2] Gunn A, Oyekunle T, Stang M, et al. Recurrent Laryngeal Nerve Injury After Thyroid Surgery: An Analysis of 11,370 Patients. J Surg Res 2020;255:42-49.

[3] Bihain F, Nomine-Criqui C, Demarquet L, et al. What is the impact of continuous neuromonitoring on the incidence of injury to the recurrent laryngeal nerve during total thyroidectomy? Surgery 2021;169(1):63-69.

[4] Wojtczak B, Sutkowska-Stępień K, Głód M, et al. Current Knowledge on the Use of Neuromonitoring in Thyroid Surgery. Biomedicines 2024;12(3):675.

[5] Liu YC, Shen CL, Fu ZY, et al. Effectiveness of the recurrent laryngeal nerve monitoring during endoscopic thyroid surgery: systematic review and meta-analysis. Int J Surg 2023;109(7):2070-2081.

[6] Wilhelm A, Conroy PC, Calthorpe L, et al. Routine use of intraoperative nerve monitoring is associated with a reduced risk of vocal cord dysfunction after thyroid cancer surgery. BMC Surg 2023;23(1):215.

[7] Shedd DP, Burget GC. Identification of the recurrent laryngeal nerve. Arch Surg 1966;92(6):861-864.

[8] Flisberg K, Lindholm T. Electrical stimulation of the human recurrent laryngeal nerve during

thyroid operation. Acta Otolaryngol Suppl 1969;263:63-67.
[9] Randolph GW, Kamani D. Intraoperative electrophysiologic monitoring of the recurrent laryngeal nerve during thyroid and parathyroid surgery: experience with 1,381 nerves at risk. Laryngoscope 2017;127(1):280-286.
[10] Mohebati A, Shaha AR. Anatomy of thyroid and parathyroid glands and neurovascular relations. Clin Anat 2012;25(1):19-31.
[11] Mayer J, van der Werf-Grohmann N, Kroll J, et al. Dysphagia after arteria lusoria dextra surgery: Anatomical considerations before redo-surgery. World J Cardiol 2017;9(2):191-195.
[12] Henry BM, Sanna S, Graves MJ, et al. The non-recurrent laryngeal nerve: a meta-analysis and clinical considerations. PeerJ 2017;5:e3012.
[13] Wang T, Dionigi G, Zhang D, et al. Diagnosis, anatomy, and electromyography profiles of 73 nonrecurrent laryngeal nerves. Head Neck 2018;40(12):2657-2663.
[14] Gavilán J, Gavilán C. Recurrent laryngeal nerve. Identification during thyroid and parathyroid surgery. Arch Otolaryngol Head Neck Surg 1986;112(12):1286-1288.
[15] Wu CW, Liu X, Barczyński M, et al. Optimal stimulation during monitored thyroid surgery: EMG response characteristics in a porcine model. Laryngoscope 2017;127(4):998-1005.
[16] Krauss JK, Lipsman N, Aziz T, et al. Technology of deep brain stimulation: current status and future directions. Nat Rev Neurol 2021;17(2):75-87.
[17] Merrill DR, Bikson M, Jefferys JG. Electrical stimulation of excitable tissue: design of efficacious and safe protocols. J Neurosci Methods 2005;141(2):171-198.
[18] Zahak M. Signal Acquisition Using Surface EMG and Circuit Design Considerations for Robotic Prosthesis [Internet]. Computational Intelligence in Electromyography Analysis - A Perspective on Current Applications and Future Challenges. InTech; 2012. Available from: http://dx.doi.org/10.5772/52556
[19] Zhao Y, Zhang D, Zhou L, et al. Proprieties of adhesive surface arrays to thyroid cartilage for recurrent laryngeal nerve monitoring. Ann Transl Med 2021;9(8):690.
[20] Randolph GW, Dralle H; International Intraoperative Monitoring Study Group, et al. Electrophysiologic recurrent laryngeal nerve monitoring during thyroid and parathyroid surgery: international standards guideline statement. Laryngoscope 2011;121 Suppl 1:S1-S16.
[21] Wang JP, Tang LQ, Bronlund JE. Surface EMG signal amplification and filtering. Int J Comput Appl 2013;82(1):15-22.
[22] Chiang FY, Lu IC, Chang PY, et al. Stimulating dissecting instruments during neuromonitoring of RLN in thyroid surgery. Laryngoscope 2015;125(12):2832-2837.
[23] Karaisli S, Gucek Haciyanli S, Haciyanli M. Comparison of stimulating dissector and intermittent stimulating probe for the identification of recurrent laryngeal nerve in reoperative setting. Eur Arch Otorhinolaryngol 2022;279(1):443-447.
[24] Zhang D, Li S, Dionigi G, et al. Stimulating and dissecting instrument for transoral endoscopic thyroidectomy: proof of concept investigation. Surg Endosc 2020;34(2):996-1005.
[25] Liu CH, Huang TY, Wu CW, et al. New developments in anterior laryngeal recording technique during neuromonitored thyroid and parathyroid surgery. Front Endocrinol (Lausanne) 2021;12:763170.
[26] Kim DH, Kim SW, Kim GJ, et al. Efficacy and Safety of Minimally Invasive Thyroid Surgery: A Network Meta-Analysis. Laryngoscope 2023;133(10):2470-2479.
[27] Ludwig B, Ludwig M, Dziekiewicz A, et al. Modern surgical techniques of thyroidectomy and advances in the prevention and treatment of perioperative complications. Cancers 2023;15(11):2931.
[28] Park J, Kim KJC. Current and Future of Robotic Surgery in Thyroid Cancer Treatment. Cancers 2024;16(13):2470.
[29] Wang B, Yu JF, Ao W, et al. Optimizing robotic thyroid surgery: lessons learned from an retrospective analysis of 104 cases. Front Endocrinol (Lausanne) 2024;15:1337322.
[30] Barczyński M, Randolph GW, Cernea CR, et al. External branch of the superior laryngeal nerve monitoring during thyroid and parathyroid surgery: International Neural Monitoring Study Group standards guideline statement. Laryngoscope 2013;123:S1-S14.
[31] Calò PG, Medas F, Conzo G, et al. Intraoperative neuromonitoring in thyroid surgery: is the two-staged thyroidectomy justified?. Int J Surg 2017;41 Suppl 1:S13-S20.
[32] Zhao Y, Li C, Wang T, et al. Translational Study to Standardize the Safe Use of Bipolar Forceps, LigaSure, Sonicision and PlasmaBlade Around the Recurrent Laryngeal Nerve in Thyroid Surgery. Surg Technol Int 2018;31:sti32/990-sti932/990.
[33] Randolph GW, Kamani D, Wu CW, et al. Surgical anatomy and monitoring of the recurrent laryngeal nerve. In: editors. Surgery of the thyroid and parathyroid glands. Elsevier, 2021. p. 326-359. e310.
[34] Schneider R, Randolph GW, Dionigi G, et al. International neural monitoring study group guideline 2018 part I: staging bilateral thyroid surgery with monitoring loss of signal. Laryngoscope 2018;128:S1-S17.

Progress in Medical Devices
ISSN: 2957-5478

Zentime

Review Article https://doi.org/10.61189/417936cenngx

Leadless pacemakers: A review of communication methods, energy management, and clinical applications

Yundi Zhao[1,2], Liping Du[2], Wei Chen[2], Ping Guo[2], Chunsheng Wu[1,2]

[1]College of Future Technology, Xi'an Jiaotong University, Xi'an 710100, Shaanxi Province, China. [2]Institute of Medical Engineering, Department of Biophysics, School of Basic Medical Sciences, Xi'an Jiaotong University, Xi'an 710061, Shaanxi Province, China.

Corresponding author: Chunsheng Wu.

Acknowledgements: This work was supported by National Key Research and Development Program of China, the Ministry of Science and Technology, People's Republic of China (2023YFC2411902).

Declaration of conflict of interest: None.

Received March 12, 2025; Accepted June 18, 2025; Published September 30, 2025

Highlights

- This review explores advancements in leadless pacemaker technology, focusing on optimized wireless communication, energy-efficient strategies, and artificial intelligence-enhanced clinical applications.
- As a minimally invasive innovation, these devices enhance patient outcomes through adaptive algorithms and secure data transmission. Key developments include load modulation to maintain signal integrity and intelligent remote monitoring for real-time diagnostics.
- The review also addresses cybersecurity challenges and underscores the transformative potential of integrated intelligent systems in revolutionizing cardiac therapeutics.

Abstract

Leadless pacemakers have emerged as a mainstream clinical solution, and their communication capabilities, crucial for reliable pacing and device monitoring, continue to evolve. This review systematically examines the fundamental principles of leadless pacemaker communication systems, current design requirements, existing challenges, and future development trends. We outline the bidirectional communication mechanism between leadless pacemakers and external programmers through wireless technologies, focusing on radio-frequency field communication coupled with load modulation techniques to optimize energy efficiency and transmission reliability. Additionally, we analyze the role of artificial intelligence in adaptive communication protocols and explore the clinical potential of remote monitoring and control systems. This comprehensive analysis aims to serve as a reference for the development of communication architectures for leadless pacemakers.

Keywords: Pacemaker, leadless, wireless communication, radio frequency field, programming, remote control

Introduction

Leadless pacemakers represent a transformative advancement in cardiac rhythm management, offering a minimally invasive alternative to traditional lead-based systems by eliminating the need for transvenous leads and subcutaneous pockets. These devices mitigate lead-related complications such as infection, lead fracture, and dislodgement, while also improving patient comfort and cosmetic outcomes.

As the technology evolves, leadless pacemakers have expanded from single-chamber ventricular pacing to dual-chamber and even multi-chamber systems, enabling more physiological pacing modes.

However, the absence of physical leads necessitates robust wireless communication systems to ensure reliable bidirectional data transmission between the implanted device and exter-

Address correspondence to: Chunsheng Wu, Institute of Medical Engineering, Department of Biophysics, School of Basic Medical Sciences, Xi'an Jiaotong University, 76 Yanta West Road, Yanta Campus, Xi'an 710061, Shaanxi Province, China. E-mail: wuchunsheng@xjtu.edu.cn.

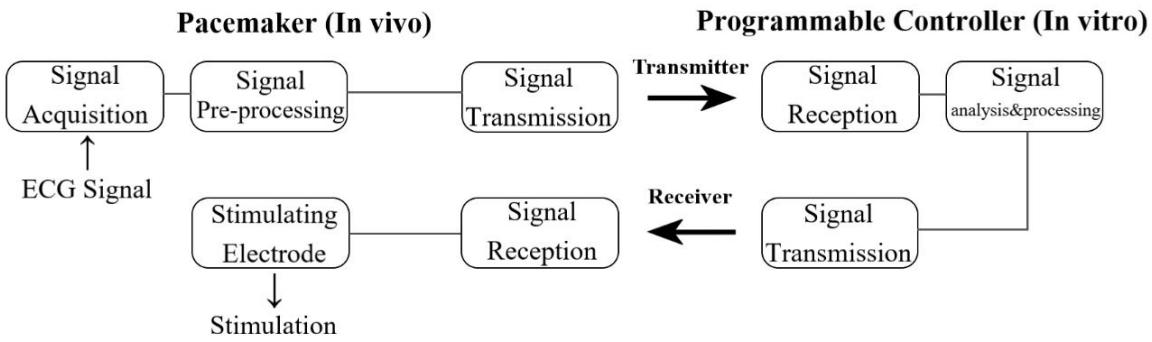


Figure 1. Schematic diagram of communication structure of the pacemaker in vivo and in vitro. ECG, electrocardiogram.

nal programmers, as well as among multiple devices within the heart.

The communication system is integral to the functionality of leadless pacemakers, enabling not only real-time telemetry and parameter adjustment but also advanced features such as adaptive pacing and remote monitoring. Despite significant progress, challenges remain in achieving energy-efficient, high-fidelity communication through biological tissues while minimizing power consumption and electromagnetic interference. This review provides a comprehensive overview of current state of leadless pacemaker communication technologies, including methods based on radio frequency (RF), load modulation, energy harvesting, and the emerging role of artificial intelligence in enhancing adaptive communication and cybersecurity. We also discuss clinical applications, ongoing challenges, and future directions aimed at optimizing the performance and safety of these innovative devices.

Communication with a leadless pacemaker

Basic communication functions of leadless pacemakers

For patients with arrhythmia, heart block, or cardiac insufficiency, a permanent cardiac pacemaker is commonly implanted to deliver electrical stimulation pulses at the required frequency and intensity, restoring sinus rhythm [1]. Pacemakers can be classified into leaded and leadless types, based on the presence or absence of leads. Leadless pacemakers are smaller and differ from leaded pacemakers in terms of size and implantation site. Without leads traversing blood vessels or cardiac chambers, leadless pacemakers offer simpler implantation, reduced trauma, and a lower risk of associated complications. Additionally, the absence of subcutaneous implantation device, further reduces the risk of infection. Leadless pacemakers also offer advantages such as lighter weight, smaller size, and minimal interference with daily activities [2]. Currently, leadless pacemakers are the mainstream choice among implantable cardiac pacemakers in clinical practice.

The primary communication requirement of leadless pacemakers pertains to bidirectional signal transmission between the implantable device and the external programming unit. The implantable component of a leadless cardiac pacing system includes a stimulation circuit, electrode, communication system, battery, and protective housing. For pacemakers with adaptive functions, a sensing circuit may also be included to capture specific signals such as acceleration, pressure, and pH. The communication system includes signal acquisition, processing, and conversion modules, along with signal transmitter and receiver modules [3]. It functions by converting physiological signals, such as electrocardiogram (ECG), into digital signals via analog-to-digital conversion and transmitting them to the external programming device. The receiver module accepts command signals from the external programming system, decodes them through digital-to-analog conversion, and adjusts the pacing pulse amplitude and frequency accordingly. This facilitates real-time, bidirectional wireless transmission of both data and commands [4]. **Figure 1** illustrates the interaction between the intracavity and extracavity structures of the pacemaker communication system.

Wireless communication methods for leadless pacemakers

Since the leadless pacemaker is directly implanted in the heart chamber, 6-10 cm beneath the body surface, and its position changes with cardiac motion, wireless communication is the only viable method for achieving two-way communication between the pacemaker and the

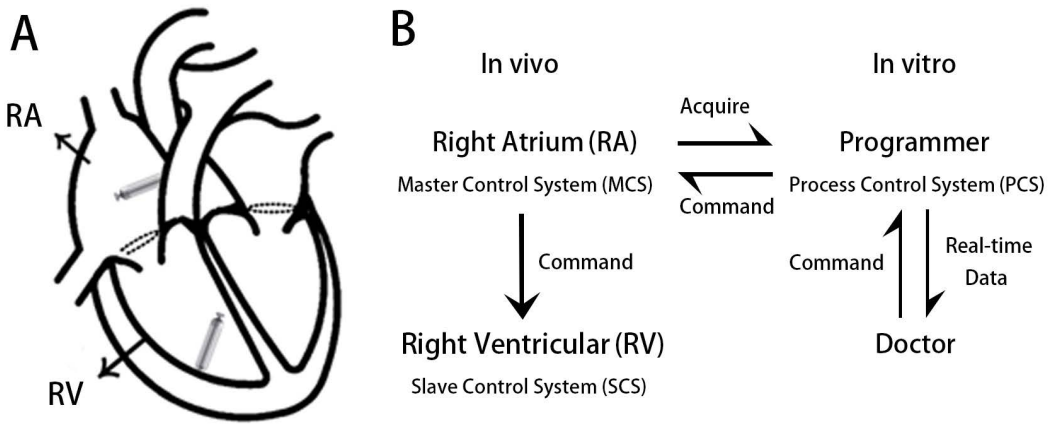


Figure 2. Configuration and intracardiac communication of a dual-chamber leadless pacemaker. (A) Schematic diagram of the implantation position of the dual-chamber leadless pacemaker in the heart. The two pacemakers are typically located in the right atrium and right ventricle respectively; (B) The intracavitary and extracavitary communication mode diagram of the dual-chamber pacemaker.

external programmer. Common wireless communication methods include Bluetooth, Wi-Fi, 5G, and RF field communication [5].

Currently, most pacemakers used in clinical practice rely on RF electromagnetic waves for communication between pacemakers and program controllers [6]. Research and development by universities and companies primarily focus on improving RF-based communication between pacemakers and external devices.

This method requires both the pacemaker and the program controller to have coil structures for signal transmission and reception. When the pacemaker sends a signal to the program controller, effective coupling between the coils of the pacemaker and the program controller is necessary to ensure efficient and accurate signal transmission.

The microprocessor first converts the signal through analog-to-digital coding, then modulates it using a modulation circuit. The signal is then transmitted as a high-frequency electromagnetic wave capable of penetrating the pacemaker housing and human tissues. The modulated signal is amplified by the power amplification circuit and transmitted via the transmitting coil. The receiver circuit in the program controller receives the signal, separates the high-frequency carrier signal using a demodulation circuit, and decodes it to obtain the original ECG signal and other data collected by the pacemaker. The process by which the program controller sends control signals to the pacemaker follows a similar mechanism.

For dual-chamber or multi-chamber pacing systems, communication is required not only between the internal pacemaker and the external programming device but also between pacemakers in different cardiac chambers to ensure synchronized pacing between the atrium and the ventricle [7]. In a dual-chamber pacemaker system, for example, the pacemaker implanted in the atrium initiates the pacing pulse, and the pacemaker implanted in the ventricle follows suit. Thus, only one-way communication from the atrium to the ventricle is needed [8]. While the distance between the two pacemakers is minimal, strict requirements for pacing time intervals, communication transmission rates, and the safety and accuracy of communication still make RF field communication the preferred method.

In this system, the atrial pacemaker serves as the master pacemaker, communicating bidirectionally with the external programming device while simultaneously sending control signals to the ventricular pacemaker. The ventricular pacemaker, as a slave pacemaker, receives control signals from the atrial pacemaker, enters a period of atrioventricular delay, and then delivers pacing pulses to control ventricular pacing. **Figure 2** depicts the communication relationship between the components of the dual-chamber pacing system. A similar communication mode applies to multi-chamber pacemakers.

Design requirements for communication systems in leadless pacemakers

The communication system of a leadless pacemaker includes modules for signal acquisition, processing, and conversion, as well as signal transmitter and receiver modules. These components convert ECG signals into digital data,

which is transmitted to the external device. This communication is bidirectional: the pacemaker transmits information about its working state, parameters, and collected electrophysiological signals to the external programming device for analysis and storage for future research [9]. In turn, the external device can send control commands to adjust pacing parameters, such as pacing frequency, threshold, maximum pulse stimulation amplitude, and refractory period to optimize the pacemaker's performance [10-12].

For a leadless pacemaker to function effectively, its communication system must fulfill several hardware and software requirements:

• Accurate and Secure Two-Way Communication: The pacemaker must reliably receive and interpret signals from the external programming device to avoid incorrect responses that could harm the user's health.

• Basic Communication Functions: These include collecting the intracavity ECG, heart rate, and real-time supply voltage of the pacemaker.

• Minimized Resource Consumption: The pacemaker's communication system should be designed to minimize energy use while ensuring normal operation, aiming to keep the average operating current at microampere levels or lower. Alternatively, external devices should provide power for the pacemaker's communication system.

• Adaptability to Multi-Parameter Adjustment: The communication system should facilitate easy and fast adjustment of pacing parameters across multiple levels to avoid complications caused by mismatches between pacemaker settings and the user's physiological needs.

• Support for Optimization and Additional Functions: The system should support functions such as achieving a transmission rate of 20 kb/s or higher and maintaining normal operation under MRI conditions.

Challenges in leadless pacemaker communication

Ideally, the communication function of a pacemaker should be instant, controllable, and resource-efficient. However, due to the current technological limitations, achieving these requirements simultaneously remains difficult. Based on a comprehensive analysis of leadless pacemaker products, such as Medtronic Micra and St. Yoda Nanostim, approved by the US Food and Drug Administration or undergoing clinical trials, several communication challenges need to be ddressed [13-16].

• Excessive Resource Consumption by the Communication Module: While leadless pacemakers are designed to be small to fit within the heart chamber, the size of the communication module often exceeds that of the pacemaker itself, reducing available space for the battery and pacing circuits. This reduces the pacemaker's battery life. Additionally, the communication function contributes significantly to the internal power consumption of the pacemaker, as it requires signal amplification, processing, and bidirectional transmission. If not carefully managed, this can lower battery life.

• Limited Communication Range and Interference: The communication signal must penetrate both the pacemaker's metal housing and the human chest during transmission. This results in significant signal attenuation, meaning the external programming device must be positioned close to the chest to maintain normal signal transmission. Otherwise, the pacemaker may fail to transmit signals, or severe signal distortion may occur.

• Communication Efficiency Challenges: The pacemaker must perform signal acquisition, processing, transmission, analysis, and reception within a short time frame—typically less than one second—since the R-R interval between heartbeats is around 1 second, and can shorten to about 0.6 seconds during intense exercise or with certain medications. If the pacemaker's communication system fails to meet this speed requirement, pacing may be delayed or even induce arrhythmias. Achieving such high communication speeds poses significant challenges for both the hardware and algorithms. The need for interchamber communication, which has higher efficiency demands, further complicates this issue. Currently, leadless pacemakers struggle to meet these requirements due to algorithm limitations, hardware constraints, and space limitations, particularly for interchamber communication, where effective solutions are lacking.

Leadless pacemaker communication technology

Factors influencing communication efficiency and accuracy

Like other medical devices with communication functions, evaluating the communication performance of leadless pacemakers requires a comprehensive analysis of both internal system

factors and external influences on communication efficiency and accuracy.

In wireless communication systems, the signal transmission rate is a key factor affecting communication efficiency. The transmission rate between the sending and receiving devices is determined by Shannon's formula (Equation (1)), where C represents the maximum transmission rate (in bits per second), W represents the channel bandwidth (in Hertz), S is the signal power, and N is the noise power (both in watts). As shown in Equation (1), the transmission rate is mainly determined by the Channel Bandwidth and the Signal-to-Noise Ratio (SNR). Channel bandwidth is typically determined by network equipment (e.g., switches, routers, hubs), topology, data type, the number of devices, and other factors.

For leadless pacemakers, increasing the bandwidth requires advanced network hardware, which would place additional strain on the pacemaker's power supply [17]. Therefore, increasing the transmission rate by expanding channel bandwidth is generally not considered feasible in leadless pacemakers. Instead, enhancing the signal transmission rate by improving the SNR is a more feasible strategy in pacemaker design [18]. To improve the SNR, two approaches can be employed:

Designing appropriate filters based on the frequency of the effective signal to exclude interference at the acquisition end.

Amplifying the effective signal to facilitate transmission without significantly increasing the resource consumption of the transmission circuit [19].

$$C = W * log_2(1 + S/N) \quad (1)$$

Communication accuracy refers to the reliability of the communication system, commonly measured by the Bit Error Rate and Mean Time Between Failures. The bit error rate is influenced by factors such as channel noise, transmission distance, data rate, and the transmission power and receiving sensitivity. Since reducing the communication distance between the internal communication module of the pacemaker and the external programming device is challenging, and the data rate is constrained by the hardware at the time of implantation, efforts to minimize Bit Error Rate focus on two areas:

• Limiting the interference from external noise, particularly electromagnetic radiation and signal interference from other electronic devices.

• Increasing the transmission power and improving the receiving sensitivity of the communication module, while ensuring user safety and maintaining efficient resource consumption.

In addition to the communication system itself, the environment in which the system operates also affects communication efficiency. Wireless signals inevitably experience loss as they penetrate obstacles, and the material of the obstacles significantly influences signal attenuation [20]. Additionally, nearby equipment such as switching power supplies, high-voltage grids, welding machines, or other high-frequency devices can negatively impact the pacemaker's communication function [21]. For example, patients with traditional pacemakers cannot undergo MRI scans unless their pacemaker is specifically designed for MRI compatibility. Even with MRI-compatible pacemakers, the device must be adjusted to the "ventricle paced, none sensed, none response" mode or another pacing-only mode during MRI operation [22]. Failure to do so may result in severe electromagnetic interference that disrupts the pacemaker's sensing and communication functions, posing significant risks to the patient's health and safety.

Energy-efficient solutions for communication

Several solutions have been proposed to address energy consumption issues in leadless pacemakers, including advanced battery technologies and wireless power transfer (WPT) techniques.

The core issue limiting the lifetime and communication functionality of leadless pacemakers is their small size, which limits battery capacity. Once implanted in the heart chamber, the pacemaker is difficult to retrieve for recharging or replacement, and attempting to do so could cause significant harm to the patient. As a result, the communication power and transmission efficiency of the pacemaker are often constrained, which can lead to complications or misdiagnosis during follow-up due to communication issues [23].

A direct solution to this issue is to use batteries with higher energy storage capacities. Biofuel cells and micro-nuclear batteries are promising alternatives. However, biofuel cells have limited energy storage and lifespan, while nuclear power batteries, although efficient and clean, pose uncontrollable risks, such as potential radiation leakage and disposal issues after patient

death. Moreover, both types of batteries are still too bulky for compact, implantable medical devices like leadless pacemakers.

WPT technology, which allows wireless energy transfer from an external source to the pacemaker, offers a more sustainable solution by reducing dependence on the pacemaker's internal power supply [24]. Near-field WPT technology enables non-invasive, contact-free energy delivery from an external source to the pacemaker, extending its battery life and reducing the need for additional surgeries.

Two main types of near-field WPT exist: inductive coupling and magnetic resonance coupling. Inductive coupling uses magnetic field induction, where two resonating coils transfer energy. Magnetic resonance coupling, on the other hand, employs oscillating electromagnetic fields to transmit energy between the coils, offering higher efficiency, larger coupling distances, and greater resistance to interference. When applying WPT to leadless pacemakers, it is necessary to consider the limited space and geometry of the coils within the pacemaker, as well as the uncertainty of pacemaker positioning after implantation [25]. This requires careful selection of operating frequencies and optimization of the inductor's quality factor in the design of the transmission system.

The combination of WPT and RF technology has become one of the most effective methods for overcoming the limited battery capacity in communication systems [26]. Load modulation enables both signal transmission and energy delivery by sending carrier signals from the transmitting unit to the receiving unit. In leadless pacemakers, load modulation works by sending a carrier signal from the external programming device to the implanted device. The load change in the receiving unit is reflected as a voltage deviation at the transmitting unit, and the timing of this deviation serves as a time reference.

The advantage of this approach is that the receiver unit can communicate with the external device without additional power consumption, while also establishing synchronization between leadless pacemakers. In dual-chamber leadless pacemakers, load modulation facilitates communication between the atrial and ventricular modules. If the atrium generates a depolarized ECG due to contraction, a voltage deviation signal is sent after a delay of T_o, and after T_α, it is received and recognized by the envelope detection circuit in the ventricle. The ventricular module then depolarizes accordingly. The time interval between atrial and ventricular depolarization meets the interval T_v shown in Equation (2), where τ represents the error time due to signal recognition.

$$T_v = T_o + T_a + \tau \tag{2}$$

Another way to reduce the communication power consumption is through the use of mixed-signal Application-Specific Integrated Circuits, which enable low-power signal transmission. These Application-Specific Integrated Circuits can implement pulse width modulation to facilitate signal transmission. Besides the channel for collecting intracardiac ECG, the pacemaker includes channels for extracting the phase and amplitude of intracardiac impedance, a pulse generator for myocardial stimulation, and a transmitter for signal transmission.

Additionally, digital/analog converters and low-voltage module bias generators are used for acquiring bioelectric signals, while high-voltagebias generators are used for stimulation [27]. Application-Specific Integrated Circuits convert biological signals into pulse width modulation signals, which are transmitted to the transmitter through the switching circuit without requiring additional analog-to-digital conversion or digital modulation [28]. This approach reduces the power consumption of the pacemaker, with the bulk of signal processing occurring in the external device, while the implanted pacemaker provides minimal conversion power.

Enhancing signal transmission efficiency

As previously mentioned, improving the SNR is the most effective strategy to enhance signal transmission efficiency in leadless pacemakers. This requires the pacemakers to have better filtering and amplification capabilities. A feasible solution involves using an antenna implanted within the pacemaker to collect RF energy from outside the body, which is then used for pacemaker pulse delivery and communication. Deeply implantable conformal antennas based on this design have shown good impedance matching characteristics in High Frequency Structure Simulator and Tissue Simulation Liquid experiments. These antennas exhibit radiation patterns similar to planar microstrip antennas in free space and can be integrated with commercial pacemakers.

Compared to previous compact rectifier antennas, the size of the antenna increases by only about 4 mm when integrated with the pacemaker, and it can still be implanted via cathe-

Table 1. Comparison of characteristics of different types of adaptive sensors

Sensor Types	Sensing speed	Perceive proportionality	Perceived specificity	Perceptual sensitivity
Body movement	High	Low	Low	Low
Q-T interval	Low	Medium	High	Medium
Arterial diastolic pressure	Lower	Higher	Medium	Low
Ventilation per minute	Medium	High	Medium	Low
Combination sensor	Higher	High	High	Medium

ter. Additionally, using the emerging Wide Band Numerical Model has further improved the antenna's design, offering more flexible resonant frequencies, lower return loss in the same frequency band, and a further reduction in the antenna's size [29]. The integrated antenna diameter is only about 7 mm.

Tissue Simulation Liquid experiments reveal that blood is a highly attenuating medium, and the pacemaker's communication effectiveness decreases as the blood content in the cardiac chamber increases. The Human Genome Organisation (HUGO) simulation experiments also showed a 3.1 dB difference in received power during systole and diastole, and a 1.3 dB difference in external communication received power. By continuously monitoring and analyzing the received power, it is possible to determine whether the heart is in systole or diastole without additional energy expenditure for heart state detection. This analysis helps the signal transmitting module select the optimal communication time, minimizing mismatch effects, enhancing coupling between the transmitting (T_x) and receiving (R_x) antennas, and improving communication efficiency while reducing communication frequency. The received power normalization formula at the R_x end is as follows.

$$|S_{21}|_{norm} = |S_{21}|/\sqrt{1-|S_T|^2}\sqrt{1-|S_R|^2} \quad (3)$$

In Equations (3), $|S_{21}|_{norm}$ represents the normalized standard received power. $|S_{21}|$ represents the received power before normalization, and S_T and S_R represent the return loss of T_x and R_x, respectively.

Adaptive communication systems in pacemakers

Adaptive pacemakers, which have been in development since the 1980s, can adjust pacing parameters based on physiological signals, unlike traditional pacemakers that require manual adjustment after implantation. These pacemakers use sensors or algorithms to create a closed-loop system, calculating the difference between the desired pacing rate and the actual pacing rate, and automatically adjusting the pacing rate and amplitude. Adaptive functionality has become a core trend in leadless pacemaker development, aiming to address the mismatch between pacemaker performance and user needs.

Currently, most adaptive pacemakers are frequency adaptive, using body movement or physiological parameters such as ventricular blood impedance, arterial diastolic pressure, or minute ventilation for adjustment [30]. **Table 1** lists the characteristics of adaptive pacemakers based on different sensor types.

The basic principle and communication mode of adaptive pacemakers are similar to those of other pacemakers. However, adaptive pacemakers require enhanced anti-interference capabilities for their algorithms and automatic programming due to frequent pacing parameter adjustments. Multi-sensing pacemakers, which combine multiple sensors, offer stronger resistance to interference from the chest and external sources. By stacking, fusing, or cross-checking sensor data with appropriate algorithms, the pacemaker can determine the optimal pacing mode for the user, making the pacing more suitable for the user's physiological needs.

Compared to traditional pacemakers that require manual programming, adaptive pacemakers can self-learn, eliminating the need to set pacing rate slopes, upper/lower limits, and sensor thresholds [31]. This reduces the pacemaker's resource usage and minimizes communication-related issues, especially in children and the elderly with complex physiological conditions [32].

New trends in pacemaker communication

Role of artificial intelligence in pacemaker communication

Currently, pacemaker parameter adjustments are primarily based on manual analysis of user data, which is time-consuming, labor-intensive, and not real-time. Even pacemakers with adaptive functions rely on pre-set algorithms, adjusting pacing parameters within a limited range based on the user's physiological state. However, as patients with permanent pacemakers age or develop acute or chronic conditions,

their physiological parameters often change in ways that are not autonomously detected and adjusted by the pacemaker [33, 34]. In special conditions such as intense exercise or large amplitude movements after prolonged rest, pacemaker algorithms may fail to make necessary adjustments, leaving the pacemaker unable to meet the user's heart requirements.

AI can address this issue by enabling pacemakers to autonomously update their algorithms in response to changes in the user's physiological state, optimizing the structure and algorithm of the pacemaker. For example, a deep learning-based method using multi-dimensional features can detect ECG signals under compressed sampling conditions. A multi-channel convolutional neural network can identify pacing and non-pacing heart rates, avoiding missed detection of pacemaker spikes, even at low sampling rates. AI also reduces errors in manual pacemaker control and improves pacemaker algorithm performance. Using tools like Matlab and Simulink, deep reinforcement learning can be applied to pacemaker algorithms to achieve high-precision continuous control [35].

One significant challenge in pacemaker communication is low SNR, especially when the pacemaker battery voltage is low, making it difficult to detect pacing spikes. T_o address this, an improved ECG signal processing algorithm based on convolutional networks was proposed, with its performance evaluated using Leave-One-Out Cross-Validation to test robustness [36]. The robustness of the pacemaker can be improved using the Mitchell and Schaeffer model, commonly used in cardiac electrophysiology. This model helps derive the recovery curve of Action Potential Duration under single or multiple stimuli. Equations (4) and (5) show the improved Mitchell and Schaeffer model, where J_s is the externally applied voltage, v_m is the transmembrane potential, h is the gate variable of the incoming current, and T_i, T_p, T_o, T_c are the four time constants affecting the four phases of the transmembrane potential.

Considering the possible malfunction of the pacemaker in the communication process, it has been proposed that software agents can continuously verify the behavior of implantable pacemakers based on Colored Pett-net and Hierarchical Fuzzy Colored Pett-net [37].

The autonomy and intelligence of AI software can be used to help pacemakers avoid behavior delay or decision-making errors caused by faults, especially in checking the operation rules of the pacemaker and enhancing the communication decision-making with more than 90% efficiency improvement.

$$\partial v_m/\partial t = h v_m(v_m - v_g)(1 - v_m)/T_{in} - (1 - h)v_m/T_o + J_s \quad (4)$$

$$\partial h/\partial t = \begin{cases} (1-h)/T_p & v_m \le v_g \\ -h/T_c & v_m > v_g \end{cases} \quad (5)$$

Remote monitoring (RM) and control technologies for pacemakers

Postoperative follow-up for implantable pacemakers is crucial to prevent malfunctions or complications that could harm the patient's health. However, data shows that approximately a quarter of pacemaker users fail to follow up as scheduled [38]. RM functions allow pacemakers to send real-time data to medical professionals, overcoming the limitation of face-to-face follow-ups and improving patient prognosis [39]. RM enables continuous monitoring of changes in pacing threshold, pacemaker duration, and remaining battery life. This allows for early detection of when pacemaker replacement is needed and assists in diagnosing changes in the user's physiological state, identifying complications or maladaptive symptoms.

For hospitals, RM facilitates the collection and aggregation of patient data via online channels such as the Internet, significantly reducing the difficulty and time required for postoperative follow-up [40]. Additionally, big data analysis models can be introduced to provide personalized diagnosis and treatment recommendations for patients. RM also helps reduce care disparities across patients, facilitating cross-referencing of pacing statuses to better address individual differences. For patients, RM serves as a real-time monitoring and alert system, with studies showing that the survival rate is positively correlated with RM usage. It can also reduce unnecessary pacing shocks and improve the overall comfort of the pacemaker experience [41].

RM has shown significant value in pacemaker fault detection. By constructing a Petri net model of the leadless pacemaker, targeted checkpoints can be integrated to create a fault diagnosis network. This allows pacemaker data to be sent as check codes, enabling medical staff to assess the pacemaker's operational status and the patient's health remotely. RM-based follow-up has demonstrated a better ability to identify rare pacemaker failures, particularly overwork situations where the pacemaker operates beyond its normal capacity. These abnormalities, such as atrial/ventricular episodes exceeding daily limits, excessive pacemaker

mode switching, or high heart rates during switching periods, can be observed and recorded over extended time frames, providing crucial data for diagnosis.

For frail populations, such as the elderly, RM offers a valuable solution. Since these patients often have difficulty attending regular follow-ups due to mobility issues, RM enables healthcare providers to remotely monitor patients, significantly reducing the cost and burden of postoperative care and making it more convenient and accessible for both patients and their families.

However, there are ongoing concerns about the validity and reliability of RM. The primary issue is the potential risk of information leakage through remote communication and vulnerabilities in the communication network, which could be exploited by hackers [42]. Such attacks could render the pacemaker inoperative or cause disruptions in communication, such as sensing failures or oversensing, leading to serious clinical complications. Additionally, hackers may access sensitive personal data, resulting in privacy breaches. To mitigate these risks, medical device manufacturers must regularly assess the security of pacemaker communication systems, identifying and addressing software vulnerabilities promptly. Moreover, cooperation between manufacturers, hospitals, and governments is essential to strengthen medical network systems and enhance penalties for unauthorized access to medical device communications.

A potential solution to protect pacemaker communications involves using specialized communication firmware to establish a secure, direct link between the pacemaker and the manufacturer's server. However, this firmware must be regularly updated, which could temporarily disrupt communication with the pacemaker. Although the likelihood of update failures and parameter loss is minimal, these risks should not be overlooked. Furthermore, adding additional hardware to improve security can increase the device's resource consumption and failure risk.

Conclusion and future directions

Leadless pacemakers are one of the most widely adopted and promising technologies in active implantable medical devices. As a critical component for ensuring proper pacing, the communication system of leadless pacemakers has seen substantial advancements in both hardware design and algorithm development, with significant clinical applications. As clinical usage grows, the design requirements for pacemaker communication systems have become more defined, highlighting both achievements and challenges.

Current research and development efforts focus primarily focus on optimizing energy utilization and improving signal transmission efficiency. WPT technology supports low-power signal transmission, while load modulation enables the replacement of internal batteries by utilizing externally transmitted energy, addressing the limitations of traditional batteries' low energy storage efficiency and short lifespan. These advances reduce the physiological risks and surgical burden associated with frequent pacemaker replacements. To further enhance communication efficiency, software optimization strategies, such as improved signal filtering and amplification algorithms, are widely adopted, alongside hardware enhancements in antenna design to improve signal transmission reliability.

Adaptive pacing technologies, which incorporate automatic programming and anti-interference communication mechanisms, enhance the pacemaker's sensing speed, compatibility, specificity, and sensitivity. These innovations minimize the need for manual programming and reduce the dependency on regular follow-ups.

AI-driven cardiac pacing algorithms are emerging as a key research direction, enabling pacemakers to autonomously adjust pacing modes and parameters according to patients' physiological condition. This improves the precision of medical assistance and enhances the overall user experience. With the growing integration of telemedicine, pacemaker programming and follow-up are shifting toward remote management, and RM technology is playing a crucial role in this evolution. As AI and RM technologies become increasingly embedded in leadless pacemakers, addressing their reliability and cybersecurity remains vital. This includes ensuring the protection of patients' personal data in the digital era and safeguarding against unauthorized access or potential malicious hardware attacks.

Author contributions: Yundi Zhao and Chunsheng Wu conceived and designed the study. Liping Du and Wei Chen contributed significantly to literature analysis and manuscript preparation. Yundi Zhao performed the literature review and drafted the manuscript. Ping Guo and Chunsheng Wu assisted with literature analysis and provided constructive discussions.

References

[1] Auffret V, Boulmier D, Didier R, et al. Clinical effects of permanent pacemaker implantation after transcatheter aortic valve implantation: Insights from the nationwide FRANCE-TAVI registry. Arch Cardiovasc Dis 2024;117(3):213-223.

[2] Knops RE, Reddy VY, Ip JE, et al. A Dual-Chamber Leadless Pacemaker. N. Engl J Med 2023;388(25):2360-2370.

[3] Steinwender C, Blessberger H, Kiblböck D, et al. [Micra™ leadless pacemaker : Clinical experience and perspectives]. Herzschrittmacherther Elektrophysiol 2018;29(4):334-339.

[4] Catalan-Matamoros D, Lopez-Villegas A, Tore-Lappegard K, et al. Correction: Patients' experiences of remote communication after pacemaker implant: The NORDLAND study. PLoS One 2019;14(7):e0219584.

[5] Clerckx B, Zhang R, Schober R, et al. Fundamentals of Wireless Information and Power Transfer: From RF Energy Harvester Models to Signal and System Designs. IEEE J Sel Area Comm 2019;37(1):4-33.

[6] Graesslin I, Vernickel P, Den Boef J H, et al. RF transmitter with digital feedback for MRI. Patent Cooperation Treaty. W02008135872-A1. 2008/4/3.

[7] Zhang HJ, Zhou WX, Hou WB, et al. A sensory pacing method for a multi-chamber leadless pacemaker system. China. CN109966643-A. 2019/3/29.

[8] Toon LT, Paisey J, Roberts PR. The dual chamber leadless pacemaker: a game changer for bradycardia management? Expert Rev Cardiovasc Ther 2024;22(7):285-288.

[9] Noormohammadi R, Khaleghi A, Bergsland J, et al. Conductive Backscatter Communication for Dual-Chamber Leadless Pacemakers. IEEE Trans Microwave Theory Tech 2022;70(4):2442-2450.

[10] Knops RE, Lloyd MS, Roberts PR, et al. A Modular Communicative Leadless Pacing-Defibrillator System. N Engl J Med 2024;391(15):1402-1412.

[11] Tong F, Sun Z. Strategies for Safe Implantation and Effective Performance of Single-Chamber and Dual-Chamber Leadless Pacemakers. J Clin Med 2023;12(7):2454.

[12] Tsitlik Joshua F, Levin H, Halperin H, et al. MRI interference-filtered monitor, stimulator and pacemaker. Patent Cooperation Treaty. W09221286-A1. 1992/5/26.

[13] Kazmi M, Rashid S, Markovic N, et al. Micra™ Leadless Intracardiac Pacemaker Implantation: A Safer Option During the Coronavirus Disease 2019 Pandemic. J Innov Cardiac Rhythm Manage 2021;12(1):4368-4370.

[14] Oosterwerff EFJ, Salavati A, Lenssen M, et al. Experience with malfunctioning leadless pacemakers: Troubleshooting and management during medium-term follow-up. Heart Rhythm 2022;19(6):894-900.

[15] Brown C, Ryan MP, Chikermane SG, et al. Incremental costs of new permanent pacemaker implantation (PPMI) after transcatheter aortic valve replacement (TAVR). Cardiovasc Revasc Med 2024;65:101-103.

[16] Mararenko A, Udongwo N, Pannu V, et al. Intracardiac leadless versus transvenous permanent pacemaker implantation: Impact on clinical outcomes and healthcare utilization. J Cardiol 2023;82(5):378-387.

[17] Alghamdi A, Alshammari A, Chang L, et al. Miniaturized Implantable Antenna with Ultra-Wide Bandwidth Characteristics for Leadless Pacemakers. 2024 18th European Conference on Antennas and Propagation (EuCAP) 2024;1-5.

[18] Bikki P, Dhiraj Y, Nivas Kumar RVS. Design and Implementation of a Sense Amplifier for Low-Power Cardiac Pacemaker. J Circuit Syst Comp 2022;32(09):2350148.

[19] Mr.Girinath N, Dr.Ganesh Babu C, Mr. Dinesh Kumar JR, et al. A Novel Low Noise Instrumentation Amplifier for Bio-Medical applications. IOP Conference Series: Materials Science and Engineering 2021;1084(1):012068.

[20] Seo DW, Lee JH, Lee H. Integration of resonant coil and antenna for wireless power transfer and data telemetry. 2014 International Symposium on Antennas and Propagation Conference Proceedings 2014;617-618.

[21] Psenakova Z, Smondrk M, Barabas J, et al. Computational Analysis of a Multi-Layered Skin and Cardiac Pacemaker Model Based on Neural Network Approach. Sensors (Basel, Switzerland) 2022;22(17):6359.

[22] Russo RJ, Costa HS, Silva PD, et al. Assessing the Risks Associated with MRI in Patients with a Pacemaker or Defibrillator. N Engl J Med 2017;376(8):755-764.

[23] Tjong FVY, Knops RE, Udo EO, et al. Leadless pacemaker versus transvenous single-chamber pacemaker therapy: A propensity score-matched analysis. Heart Rhythm 2018;15(9):1387-1393.

[24] Perera TDP, Jayakody DNK, Sharma SK, et al. Simultaneous Wireless Information and Power Transfer (SWIPT): Recent Advances and Future Challenges. IEEE Commun Surv Tutor 2018;20(1):264-302.

[25] Soares IV, Nikolayev D. Stochastic Analysis of WPT Efficiency due to Location Uncertainty of mm-Sized Deep-Implanted Pacemakers. 2022 Wireless Power Week (WPW) 2022;284-288.

[26] Falco PEd, Hallberg W, Barton T. Load Modulated RF Amplifiers for Wireless Communications. 2020 IEEE Texas Symposium on Wireless and Microwave Circuits and Systems (WMCS) 2020;1-6.

[27] Kim H, Kim S, Helleputte NV, et al. A Configurable and Low-Power Mixed Signal SoC for Portable ECG Monitoring Applications. IEEE Trans Biomed Circuits Syst 2014;8(2):257-267.

[28] Yang X, Yang S, Abbasi QH, et al. Sparsity-Inspired Nonparametric Probability Characterization for Radio Propagation in Body Area Networks. IEEE J Biomed Health Inf 2015;19(3):858-865.

[29] Asif SM, Iftikhar A, Braaten BD, et al. A Wide-Band Tissue Numerical Model for Deeply Implantable Antennas for RF-Powered Leadless Pacemakers. IEEE Access 2019;7:31031-31042.

[30] Burnam Michael, Gang Eli. An intelligently, continuously and physiologically controlled pacemaker and method of operation of the same. Patent Cooperation Treaty. W02020210060-A1. 2020/3/27.

[31] Schuster P, Faerestrand S, Ohm OJ, et al. Proportionality of rate response to metabolic workload provided by a rate adaptive pacemaker with automatic rate profile optimization. Europace 2005;7(1):54-59.

[32] Aktürk İ F, Erol MK. [Bradyarrhythmias and pacemaker indications in elderly patients]. Turk Kardiyol Dern Ars 2017;45(Suppl 5):71-74.

[33] Dawood M, Elsharkawy E, Abdel-Hay MA, et al. Comparative study between effects of single and dual chamber pacemaker on stroke volume, cardiac output and strain using 3D echocardiography, pulsed Doppler method and global longitudinal strain (GLS). Eur Heart J 2020;41(Supplement_2):ehaa946.0716.

[34] Sikora K, Wawryniuk A, Łuczyk RJ, et al. The Occurrence of Stress, Illness Acceptance and the Quality of Life of Patients after Pacemaker Implantation. Int J Environ Res Public Health 2022;19(21):14133.

[35] Datta A, Kolwadkar B, Ingale VV. Artificial Cardiac Pacemaker Control design using Deep Reinforcement learning: A Continuous Control Approach. 2021 5th International Conference on Trends in Electronics and Informatics (ICOEI) 2021;1031-1038.

[36] Rohr M, Huang Z, Umutcan Uguz D, et al. Limitations of Pacemaker Spike Detection in Capacitive ECGs via Deep Learning. Curr Dir Biomed Eng 2023;9(1):182-185.

[37] Majma N, Babamir SM. Model-Based Monitoring and Adaptation of Pacemaker Behavior Using Hierarchical Fuzzy Colored Petri-Nets. IEEE Trans Syst Man Cybern, Syst 2020;50(9):3344-3357.

[38] Burkart-Kuettner D, Nuernberg M, Zweiker D. Empfehlungen zur Nachsorge von konventionellen Herzschrittmacher-Systemen // Recommendations for aftercare of conventional cardiac pacemaker systems. J fur Kardiol 2023;30(9-10):209-216.

[39] Jacobs AK, Anderson JL, Halperin JL. The evolution and future of ACC/AHA clinical practice guidelines: a 30-year journey: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines. J Am Coll Cardiol 2014;64(13):1373-1384.

[40] Folino AF, Breda R, Calzavara P, et al. Remote follow-up of pacemakers in a selected population of debilitated elderly patients. Europace 2013;15(3):382-387.

[41] Varma N, Pavri BB, Stambler B, et al. Same-day discovery of implantable cardioverter defibrillator dysfunction in the TRUST remote monitoring trial: influence of contrasting messaging systems. Europace 2013;15(5):697-703.

[42] Alexander B, Haseeb S, Baranchuk A. Are implanted electronic devices hackable? Trends Cardiovasc Med 2019;29(8):476-480.