Generative AI–Powered Navigation System for Precision Puncture: A Multicenter Randomized Controlled Trial Evaluating Efficacy and Safety

Title	Generative AI–Powered Navigation System for Precision Puncture: A Multicenter Randomized Controlled Trial Evaluating Efficacy and Safety
Version	V 2.0
Date	17 June 2025
Trial Principal Investigator Other investigator(s)	Prof. Bo Du Dr. Huangxuan Zhao Dr. Lei Chen
Subsidising party <if applicable>	Not applicable

Statistical Analysis Plan

Generative AI–Powered Navigation System for Precision Puncture: A Multicenter Randomized Controlled Trial Evaluating Efficacy and Safety

Contents

1 Introduction and rationale 2

2 Objectives and endpoints 3

2.1 Objectives 3

2.2 Endpoints 4

2.2.1 Primary endpoint 4

2.2.2 Second endpoints 4

2.2.3 Other endpoints 4

3 Study Methods 4

3.1 General study design and plan 4

3.2 Inclusion and exclusion criteria 7

3.3 Randomization and binding 7

3.4 Detailed sample size calculation 7

4 Investigational product 9

4.1 Name and description of investigational product(s) 9

4.2 Summary of findings from non-clinical studies 9

4.3 Summary of findings from clinical studies 9

4.4 Summary of known and potential risks and benefits 9

5 Statistical analysis 10

5.1 Withdrawal of individual subjects 10

5.2 Replacement of individual subjects after withdrawal 10

5.3 Follow-up of subjects withdrawn from treatment 10

5.4 Missing data processing 10

5.5 Primary and secondary study parameter(s) 11

5.6 Other study parameter(s) 11

6 Ethical considerations 11

6.1 Regulation statement 11

6.2 Recruitment and consent 11

7 Data retention and management 12

8 Amendments 12
9 References 12

1 Introduction and rationale

Lung puncture biopsy is the gold standard for qualitative diagnosis of pulmonary nodules[1]. There are various techniques for performing lung puncture biopsy, among which cone beam CT (CBCT)-guided percutaneous lung puncture (PLP) has emerged as a novel approach in recent years[2-4]. Compared with conventional CT-guided lung puncture biopsy, CBCT reconstructs a CT-like images while offering the advantages of reduced radiation dose exposure. Additionally, CBCT-guided lung puncture biopsy is performed in real-time under the fluoroscopic guidance using the DSA, which not only enhances procedural efficiency but also minimizes radiation damage to the patient. However, one limitation of CBCT is its relatively poor image quality. This becomes particularly problematic in cases involving small nodules or nodules adjacent to vital organs or blood vessels, as the poor image quality may affect the procedural accuracy, potentially leading to failed puncture or serious complications. Therefore, improving the quality of CBCT images is crucial to enhancing the efficacy of CBCT-guided lung puncture biopsy procedures and reduce complications.

Artificial Intelligence (AI) has been heavily researched in image segmentation and generation and has demonstrated promising results[5-7]. In particular, generative models such as Generative Adversarial Networks (GANs) and diffusion models[8,9] have shown potential in enhancing the quality of images and transforming multimodal images. Previous studies have been conducted to improve CBCT image quality through generative AI models and have demonstrated that these models can effectively improve CBCT image quality, achieving quality comparable to that of conventional CT images[10-12]. These studies demonstrate the great potential of generative AI in improving the quality of CBCT images particularly in guiding lung puncture procedures. However, AI has barely been used in guiding clinical surgical practice. This is mainly due to the

heterogeneity of clinical data used in daily practice, with the datasets used for training AI algorithms often being considered unrepresentative of real-world clinical conditions. Therefore, prospective clinical trials are essential to validate the applicability of AI in clinical applications. As with any medical innovation, clinical trials are indispensable for attributing tangible benefits to the employment of AI in CBCT-guided lung aspiration biopsy procedures.

A generative AI model was developed by retrospectively analyzing data from 3106 participants who underwent CBCT-guided lung puncture biopsy procedures, which the aim of improving the quality of CBCT images. Then, a Generative AI-Powered Puncture Surgery Navigation System for lung punctures (GPS-Lung) system is constructed based on the synthesis CT (sCT) images and intraoperative guidance system. We raised a question: Can GPS-Lung system contribute to the improvement of clinical lung puncture procedures? The application of GPS-Lung system to guide lung puncture procedures can improve the procedural efficacy, reduce radiation exposure, and minimize surgical complications, thereby providing tangible benefits to patients. To validate the clinical applicability of the constructed GPS-Lung system, a follow-up prospective randomized controlled trail (RCT) was conducted to compare the surgical outcomes, radiation dose, and complications rates between conventional CBCT-guided lung puncture procedures and GPS-Lung system guided lung puncture procedures.

2 Objectives and endpoints

2.1 Objectives

- Constructing a Generative AI-Powered Puncture Surgery Navigation System for lung punctures (GPS-Lung) system for lung puncture.
- Using the GPS-Lung system to guide lung puncture surgery, improving puncture surgery efficiency, reducing radiation exposure and minimizing complications.

2.2 Endpoints

2.2.1 Primary endpoint

Puncture times of participants for GPS-Lung system and CBCT-guided lung puncture procedures.

2.2.2 Second endpoints

Radiation dose, duration of surgery and intraoperative and postoperative complications within 7 days of participants for GPS-Lung system and CBCT-guided lung puncture procedures.

2.2.3 Other endpoints

- Algorithmic performance for image quality.
- Doctors' scoring of image quality and lesion visualization.

3 Study Methods

3.1 General study design and plan

- Study configuration and experimental design: This study is a clinical randomized controlled trail.
- Type of comparison: Puncture times, duration of surgery, radiation dose and complications of participants guided by the GPS-Lung system for PLP versus those guided by conventional CBCT are compared.
- Type of control(s): CBCT-guided PLP without the use of the GPS-Lung system.
- Level and method of blinding: This is a double-blind study (Investigators and participants will be unaware of whether they are randomized to GPS-Lung system guided PLP group or CBCT-guided PLP group without aid of AI).

Statistical Analysis Plan

- Method of treatment assignment: Participants will be randomly assigned to treatment or control group using a computer-generated random number sequence, with a 1:1 allocation ratio.
- At what point in time subjects are randomized relative to treatments: Participants will be randomized before PLP initiation.

The proposed clinical workflow is present in **Figure 1**

Statistical Analysis Plan

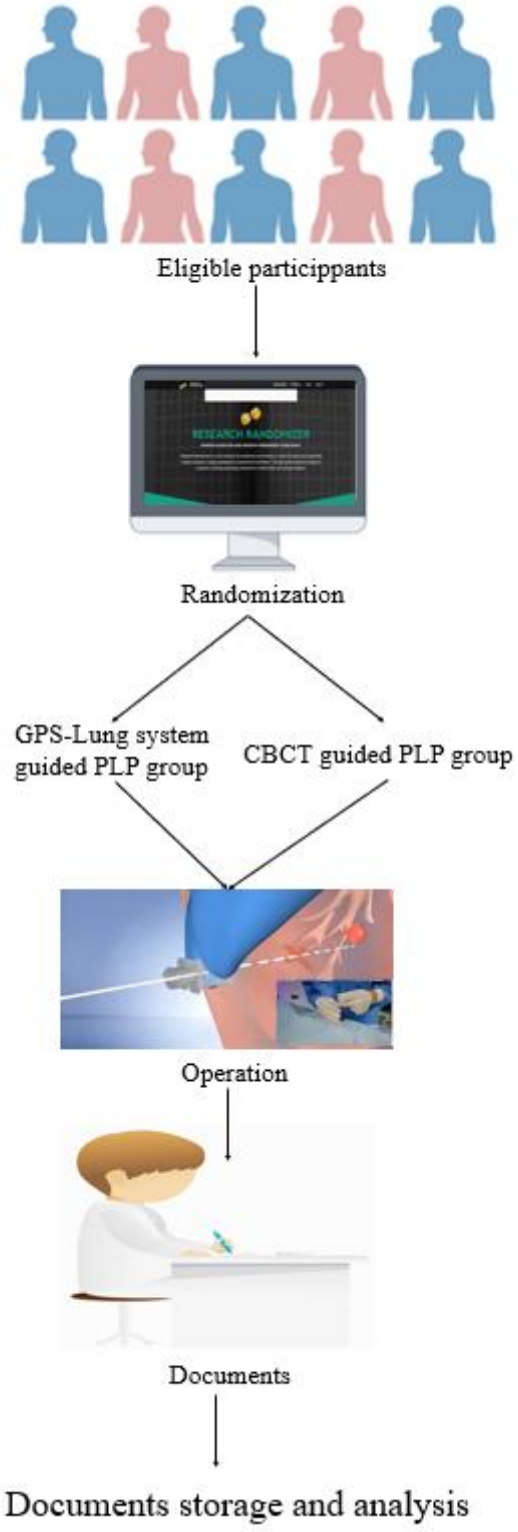


Figure 1. Clinical workflow of this study.

3.2 Inclusion and exclusion criteria

Inclusion criteria

- Participants who require CBCT-guided PLP and meet the clinical indications for the procedure.
- Participants with pre-operation CT images available.
- Participants whose physical condition is suitable for PLP.

Exclusion criteria

- Participants with a history of allergy or serious adverse reactions to iodine contrast media or other related drugs.
- Participants are pregnant or breastfeeding.
- Participants are unwilling or unable to sign informed consent.

3.3 Randomization and binding

Randomization will be performed using a web-based system, employing computer-generated permuted-block sequences with stratification by center. Notably, this centralized, web-based randomization service ensures concealment of trial-group assignments.

A double-blind design will be conducted to ensure that neither participants nor researchers will be aware of group allocation during this study. Participants, researchers, data collectors and outcome assessors will be blinded to group assignment. Statisticians and authors will be also blinded to subgroup allocation during data analysis. Doctors, technicians and trial nurses will be aware of subgroup assignments but will not be involved in data collection or analysis.

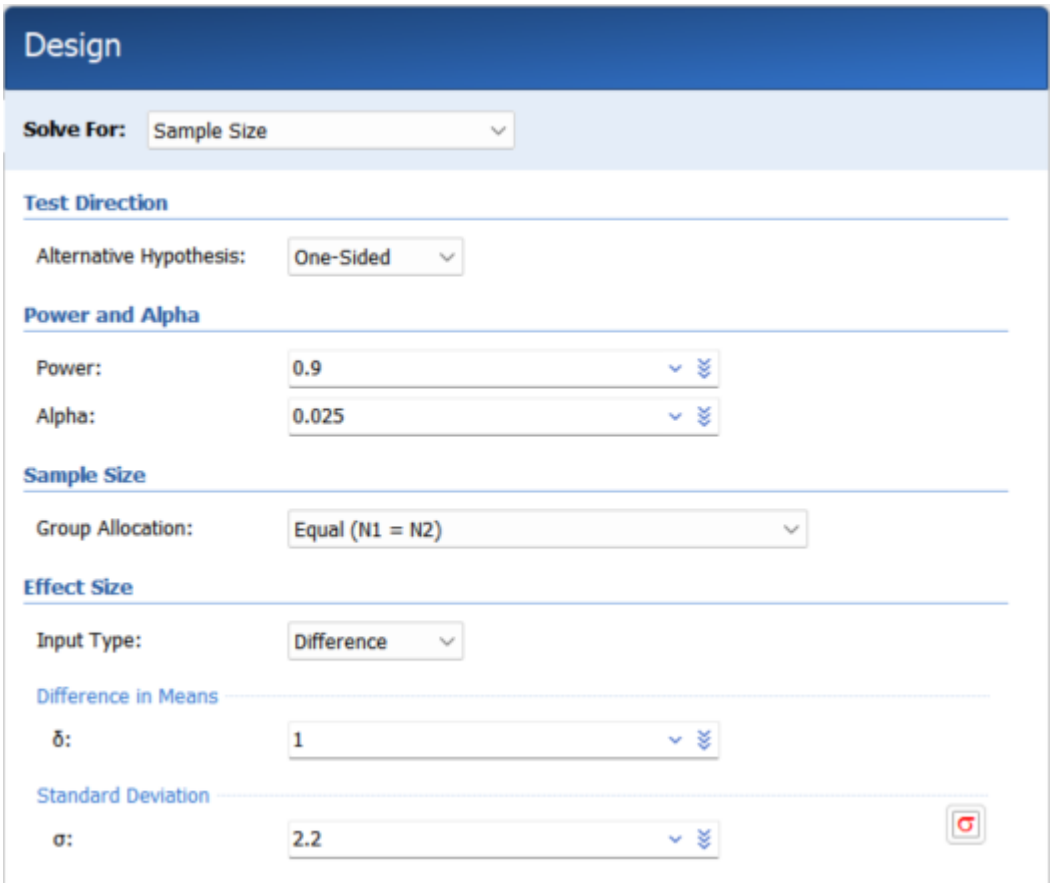
3.4 Detailed sample size calculation

The randomization process was conducted by an independent statistician who was not involved in the recruitment, treatment, or data analysis of the participants. The sample size was calculated for the primary endpoint (The mean number of intraoperative

puncture times). A total of 38 participants received CBCT-guided PLP from retrospective cohort were included from Wuhan Union Hospital. The mean (SD) puncture times of these patients were 2.3±2.2. We used a 2.5% one-sided significance level with 90% efficacy for sample allocation on a 1:1 basis. The sample size was calculated based on the following formula:

$$n = \frac{2(Z_{\alpha/2} + Z_{\beta/2})^2 * \sigma^2}{\Delta^2}$$

n represents the sample size of each group. $Z_{\alpha/2}$ represents the critical value of the significance level (α), and the α=0.025 in the study. $Z_{\beta/2}$ represents the critical value of efficacy (1 - β), and the β=0.9 in the study. σ^2 represents the within-group variance. In the current, σ was substituted by the standard deviation of puncture times from 38 participants. Δ represents the smallest difference between the means of the two groups. In the current, the Δ was considered as 1. Therefore, the sample size to be included in this study for each group is 103. We will included 220 participants (each group 110 participants) (**Figure 2**).



PASS 2021, v21.0.3

Two-Sample T-Tests Assuming Equal Variance

Numeric Results for an Equal-Variance T-Test

$\delta = \mu1 - \mu2$

Hypotheses: $H0: \delta \leq 0$ vs. $H1: \delta > 0$

Target Power	Actual Power	N1	N2	N	δ	σ	Alpha
0.9	0.90089	103	103	206	1	2.2	0.025

Figure 2. Sample size calculation process

Finally, we decided to adopt a conservative approach and include 220 participants in this study. This study will take approximately three months to complete.

4 Investigational product

4.1 Name and description of investigational product(s)

Generative AI-Powered Puncture Surgery Navigation System for lung punctures (GPS-Lung) system.

4.2 Summary of findings from non-clinical studies

Although the algorithm has been developed to improve the quality of the CBCT images, no clinical studies have been published as yet.

4.3 Summary of findings from clinical studies

To date, no clinical studies have been performed, as only retrospective studies have been reported.

4.4 Summary of known and potential risks and benefits

Potential risks: During the procedure, it may not be possible to reconstruct the CBCT image into sCT image, resulting in the inability to use the sCT image for

guided lung puncture procedures. In such cases, CBCT images could still be used to guide the puncture.

Potential benefits: Using GPS-Lung system for lung puncture guidance may reduce the puncture times, lower radiation dose and decrease complications of participants.

5 Statistical analysis

5.1 Withdrawal of individual subjects

Not applicable.

Patient data will be anonymized, and collected as part of standard practice of PLP. Therefore, participants cannot withdraw once data collection begins. Although the data will be anonymized, participants who object to the use of their data for research purposes (i.e., general 'no-object') will be excluded from the study.

5.2 Replacement of individual subjects after withdrawal

Not applicable.

5.3 Follow-up of subjects withdrawn from treatment

If participants refuse to undergo the 7-day postoperative complications follow-up, they will be excluded from this study.

5.4 Missing data processing

Missing data are not expected, as most parameters are obligatory fields in the PLP surgery reports that will be used in this study.

5.5 Primary and secondary study parameter(s)

For primary outcome, secondary outcomes and baseline comparisons between the two groups, appropriate statistical tests (parametric or non-parametric) will be applied to categorical variables (Chi-squared test/Fisher's exact test) or continuous variables (*t*-test, Mann-Whitney U test).

5.6 Other study parameter(s)

Paired *t*-tests will be used to compare the mean scores of overall image quality and the degree of lesion visualization between paired groups. The weighted kappa coefficient (κ), kendall coefficient of concordance (W) or intraclass correlation coefficient (ICC) will be used to assess the inter-reader agreement on image quality and the lesions score of sCT.

6 Ethical considerations

6.1 Regulation statement

This study will be conducted according to Consolidated Standards of Reporting Trials (CONSORT) and Declaration of Helsinki[13,14], and in accordance with the EU GDPR (General Data Protection Regulation).

6.2 Recruitment and consent

Informed consent is required from all participants.

7 Data retention and management

Data will be anonymized to the researchers, who will access it in Picture Archiving and Communication Systems (PACS) to ensure security. Data will be kept for a period of 15 years. For further details, please see the Study Protocol.

8 Amendments

Amendments refer to changes made to the research after an ethical committee approval. Any change will be submitted to the ethical committee for review and approval.

9 References

1 Mazzone, P. J. & Lam, L. Evaluating the Patient With a Pulmonary Nodule: A Review. *Jama* **327**, 264-273 (2022). https://doi.org/10.1001/jama.2021.24287

2 Sharma, A. & Shepard, J. O. Lung Cancer Biopsies. *Radiol Clin North Am* **56**, 377-390 (2018). https://doi.org/10.1016/j.rcl.2018.01.001

3 Reid, C. E. & Wallis, A. G. Percutaneous image-guided lung biopsy. *Br J Hosp Med (Lond)* **78**, C178-c182 (2017). https://doi.org/10.12968/hmed.2017.78.12.C178

4 Verhoeven, R. L. J., Kops, S. E. P., Wijma, I. N., Ter Woerds, D. K. M. & van der Heijden, E. Cone-beam CT in lung biopsy: a clinical practice review on lessons learned and future perspectives. *Annals of translational medicine* **11**, 361 (2023). https://doi.org/10.21037/atm-22-2845

5 Hosny, A., Parmar, C., Quackenbush, J., Schwartz, L. H. & Aerts, H. Artificial intelligence in radiology. *Nature reviews. Cancer* **18**, 500-510 (2018). https://doi.org/10.1038/s41568-018-0016-5

6 van der Velden, B. H. M., Kuijf, H. J., Gilhuijs, K. G. A. & Viergever, M. A. Explainable artificial intelligence (XAI) in deep learning-based medical image analysis. *Medical image analysis* **79**, 102470 (2022). https://doi.org/10.1016/j.media.2022.102470

7 England, J. R. & Cheng, P. M. Artificial Intelligence for Medical Image Analysis: A Guide for Authors and Reviewers. *AJR Am J Roentgenol* **212**, 513-519 (2019). https://doi.org/10.2214/ajr.18.20490

8 Goodfellow, I. *et al.* Generative adversarial networks. *Communications of the ACM* **63**, 139-

144 (2020).

9 Ho, J., Jain, A. & Abbeel, P. Denoising diffusion probabilistic models. *Advances in neural information processing systems* **33**, 6840-6851 (2020).

10 Deng, L., Ji, Y., Huang, S., Yang, X. & Wang, J. Synthetic CT generation from CBCT using double-chain-CycleGAN. *Computers in biology and medicine* **161**, 106889 (2023). https://doi.org/10.1016/j.compbiomed.2023.106889

11 Li, Z. *et al.* Using RegGAN to generate synthetic CT images from CBCT images acquired with different linear accelerators. *BMC Cancer* **23**, 828 (2023). https://doi.org/10.1186/s12885-023-11274-7

12 Chen, X. *et al.* CBCT-based synthetic CT image generation using a diffusion model for CBCT-guided lung radiotherapy. *Medical physics* **51**, 8168-8178 (2024). https://doi.org/10.1002/mp.17328

13 Piaggio, G., Elbourne, D. R., Pocock, S. J., Evans, S. J. & Altman, D. G. Reporting of noninferiority and equivalence randomized trials: extension of the CONSORT 2010 statement. *Jama* **308**, 2594-2604 (2012). https://doi.org/10.1001/jama.2012.87802

14 World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. *Jama* **310**, 2191-2194 (2013). https://doi.org/10.1001/jama.2013.281053