

AI-based large-scale screening of gastric cancer from noncontrast CT imaging

Received: 23 January 2025

Accepted: 19 May 2025

Published online: 24 June 2025

 Check for updates

A list of authors and their affiliations appears at the end of the paper

Early detection through screening is critical for reducing gastric cancer (GC) mortality. However, in most high-prevalence regions, large-scale screening remains challenging due to limited resources, low compliance and suboptimal detection rate of upper endoscopic screening. Therefore, there is an urgent need for more efficient screening protocols. Noncontrast computed tomography (CT), routinely performed for clinical purposes, presents a promising avenue for large-scale designed or opportunistic screening. Here we developed the Gastric Cancer Risk Assessment Procedure with Artificial Intelligence (GRAPE), leveraging noncontrast CT and deep learning to identify GC. Our study comprised three phases. First, we developed GRAPE using a cohort from 2 centers in China (3,470 GC and 3,250 non-GC cases) and validated its performance on an internal validation set (1,298 cases, area under curve = 0.970) and an independent external cohort from 16 centers (18,160 cases, area under curve = 0.927). Subgroup analysis showed that the detection rate of GRAPE increased with advancing T stage but was independent of tumor location. Next, we compared the interpretations of GRAPE with those of radiologists and assessed its potential in assisting diagnostic interpretation. Reader studies demonstrated that GRAPE significantly outperformed radiologists, improving sensitivity by 21.8% and specificity by 14.0%, particularly in early-stage GC. Finally, we evaluated GRAPE in real-world opportunistic screening using 78,593 consecutive noncontrast CT scans from a comprehensive cancer center and 2 independent regional hospitals. GRAPE identified persons at high risk with GC detection rates of 24.5% and 17.7% in 2 regional hospitals, with 23.2% and 26.8% of detected cases in T1/T2 stage. Additionally, GRAPE detected GC cases that radiologists had initially missed, enabling earlier diagnosis of GC during follow-up for other diseases. In conclusion, GRAPE demonstrates strong potential for large-scale GC screening, offering a feasible and effective approach for early detection. ClinicalTrials.gov registration: [NCT06614179](https://clinicaltrials.gov/ct2/show/study/NCT06614179).

Gastric cancer (GC) is the fifth most commonly diagnosed cancer and the fourth leading cause of cancer-related deaths worldwide¹, with nearly 75% of new cases and related deaths occurring in Asian countries, particularly China, Japan and Korea^{2,3}. Effective screening is crucial to reduce mortality rate as studies show that the 5-year survival rate for early-stage, resectable GC (EGC) is 95–99%, compared with less than 30% for advanced stages (AGC)^{4,5}. Endoscopy is the standard diagnostic test for GC because it allows direct visualization of the gastric mucosa and biopsy sampling for histologic evaluation. Japan and Korea have implemented national GC screening programs via endoscopy since 1983 and 1999, respectively, leading to higher early diagnosis rates and reduced mortality, with survival rates significantly higher in Japan (60.3%) and Korea (68.9%) compared with China (35.9%)^{6,7}.

However, in most countries with high GC prevalence rates, large-scale GC screening in the general population via endoscopy is not feasible or cost-effective, leading to a high mortality rate. This is due mainly to the high cost, suboptimal detection rate (the percentage of detected GC cases out of the total number of people screened) and the invasive nature of endoscopy, resulting in low compliance among the population⁸. Screening high-risk people is preferable, as they are more likely to undergo endoscopic screening following positive test results. Currently, serological tests, including *Helicobacter pylori* serology, serum pepsinogen levels and gastrin-17 testing, are the methods used most commonly for identifying people at high risk for GC. However, the detection rate of GC by gastroscopy after serological screening and risk stratification was only 1.25%, showing limited improvement compared with the detection rate of gastroscopy screening in the general population (1.20%)⁹. Therefore, there is an urgent need to develop new noninvasive, low-cost, efficient and reliable screening techniques to identify people at high risk for GC. Such techniques could prioritize endoscopic examinations for high-risk populations, improving cost-effectiveness of large-scale screening programs and ultimately reducing GC-related mortality¹⁰.

Noncontrast computed tomography (CT) is a low-cost and widely used imaging protocol in physical examination centers and hospitals (especially in low-resource regions)^{11,12}. Low-dose noncontrast chest CT has become an established diagnostic paradigm for pulmonary nodule surveillance in clinical practice^{13,14}. Noncontrast abdominal CT is usually indicated for rapid diagnosis of acute conditions, evaluating traumatic injuries and assessing patients who have contraindications to contrast agents, which serves as a critical tool for the initial clinical evaluation of abdominal pathologies. Recent advances in artificial intelligence (AI) have shown promising results in cancer screening via noncontrast CT, notably in the detection of pancreatic cancer¹⁵. Currently, contrast-enhanced CT is used to evaluate GC¹⁶, including invasion depth, lymph node metastasis and distant metastasis. However, identifying the primary GC lesion, particularly assessing the invasion depth, is often influenced by stomach filling and gastrointestinal peristalsis. These limitations have hindered the exploration of the role of CT scans in GC evaluation. Routine CT imaging, although a valuable tool for opportunistic screening¹¹, faces substantial challenges in detecting GC. The value of routine CT imaging in GC detection, especially noncontrast CT, has the potential to be explored further by AI.

In this study, we proposed GRAPE (GC risk assessment procedure with AI), using noncontrast CT and deep learning to identify patients with GC. The study was conducted in three distinct phases. First, we developed GRAPE using a cohort from two centers across China, including 3,470 GC cases and 3,250 non-GC (NGC) cases. Its effectiveness and utility were validated using internal validation cohorts (1,298 cases) and an independent external validation cohort from 16 centers (18,160 cases). Then, we compared the interpretations of GRAPE with those of radiologists and explored its potential in assisting radiologists during interpretation. Finally, we validated the performance of GRAPE in opportunistic screening using two real-world study cohorts comprising 78,593 consecutive participants with

noncontrast CT from one comprehensive cancer center and two independent regional hospitals. We aim to validate the ability of GRAPE in incidental GC detection in opportunistic screening on routine patients included from various scenarios (physical examination, emergency, inpatient and outpatient department) in the first real-world cohort and on patients treated or being followed up for other cancer in the second real-world cohort (Fig. 1).

Results

The GRAPE model

The GRAPE model is a deep-learning framework designed to analyze three-dimensional (3D) noncontrast CT scans for GC detection and segmentation. GRAPE was trained on the internal training cohort, including 3,470 GC cases and 3,250 NGC cases (Extended Data Table 1). It generates two outputs: a pixel-level segmentation mask of the stomach and tumors, and a classification score distinguishing GC patients from NGC. The model follows a two-stage approach (Extended Data Fig. 1a). In the first stage, a segmentation network is used to locate the stomach within the full CT scan, generating a segmentation mask that is then used to crop and isolate the stomach region. This cropped region is fed into the second stage, where a joint classification and segmentation network with dual branches is employed. The segmentation branch detects tumors within the identified stomach region, whereas the classification branch integrates multilevel features to classify the patient as either GC-positive or NGC. Detailed description, the architecture and the interpretability analysis of GRAPE are illustrated in Methods and Extended Data Fig. 1.

Internal and external evaluation

GRAPE achieved an area under the receiver operating characteristic (ROC) curve (AUC) of 0.970 (95% confidence interval (CI) 0.962–0.978), sensitivity of 0.851 (95% CI 0.825–0.877) and specificity of 0.968 (95% CI 0.953–0.980) in the internal validation cohort (Fig. 2a,d), whereas external validation confirmed stable performance with an AUC of 0.927 (95% CI 0.922–0.931), sensitivity 0.817 (95% CI 0.807–0.827), specificity 0.905 (95% CI 0.900–0.910) (Fig. 2c,d and Extended Data Table 2). No statistically significant differences emerged between Zhejiang Cancer Center (ZJCC) and Ningbo Second Hospital (NBSH) cohorts (Fig. 2b and Extended Data Table 2). Multicenter analysis ($n \geq 100$ per center) showed a consistent AUC range of 0.902–0.995 (Extended Data Figs. 2 and 3a and Extended Data Table 2).

GRAPE score distribution analysis revealed clear differentiation between GC and NGC groups (Fig. 2e,f and Extended Data Table 1), with >0.5 scores predominating in GC. T stage progression showed significant score escalation (Fig. 2g,h), whereas tumor location demonstrated no correlation (Fig. 2g–j). Detection rates increased with T stage advancement in both cohorts. Location-specific analysis showed optimal detection for whole-stomach lesions, with similar performance across different locations (Fig. 2k,l). Sensitivity remained consistent across locations within identical T stages (Extended Data Fig. 3b–e). Notably, well-filled stomachs showed 10.72% higher EGC detection versus poor filling (Extended Data Fig. 3f). Additionally, the detection rate of GC was associated with TNM stage but showed no correlation with gender or age (Extended Data Fig. 3g–i).

Reader studies

In a multicenter reader study ($n = 13$ radiologists) interpreting 297 non-contrast CT scans, GRAPE surpassed all readers in diagnostic accuracy (AUC, GRAPE 0.92 versus readers' range 0.76–0.85), demonstrating significant performance advantages in both sensitivity and specificity. When reassessing cases with AI assistance after a ≥ 1 -month washout, radiologists achieved mean improvements of 6.6% sensitivity and 13.3% specificity. Notably, while both senior and junior radiologists showed significant accuracy gains ($P < 0.05$), their augmented performance remained below GRAPE's standalone results (Fig. 3a–c).

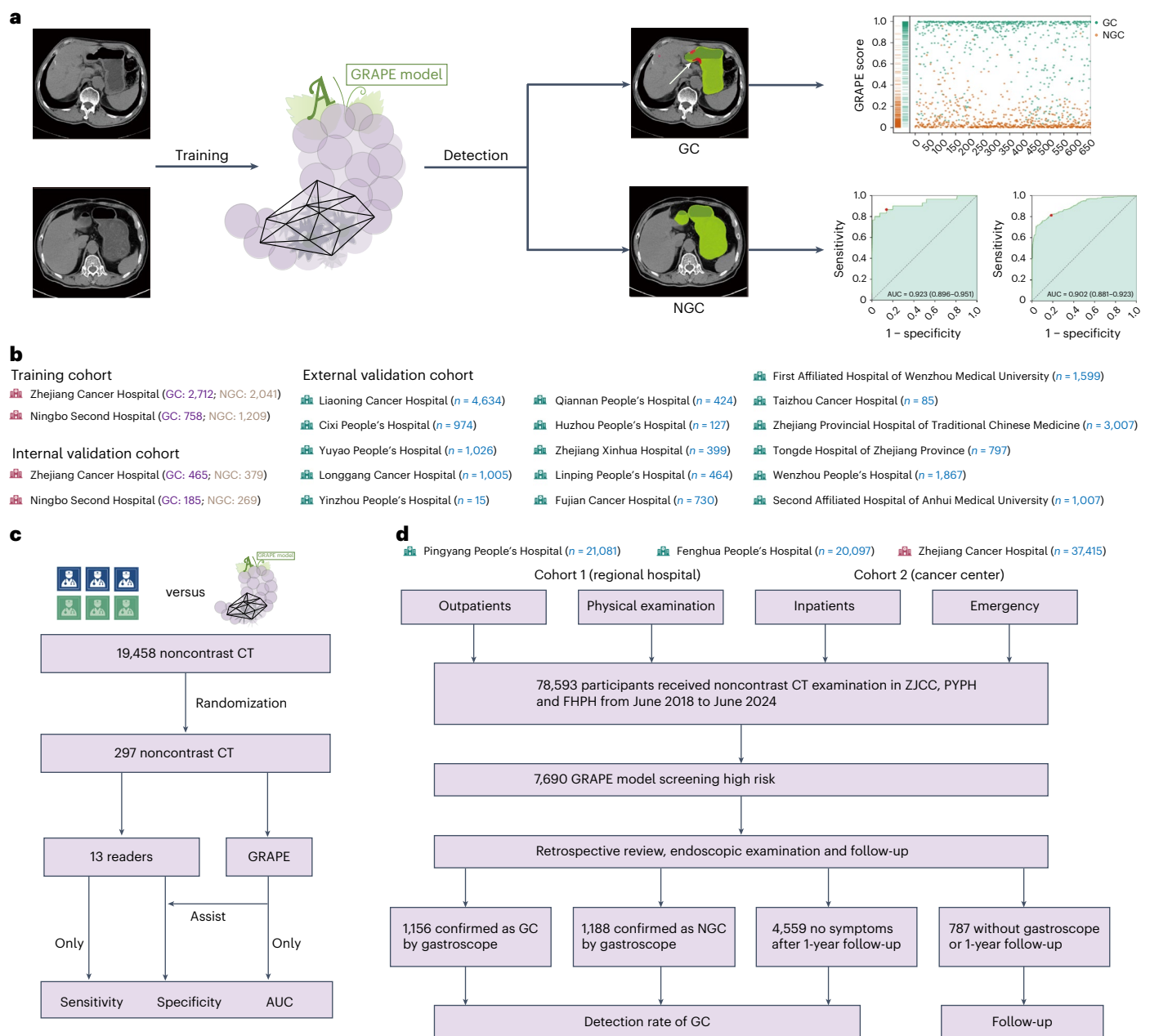


Fig. 1 | Overview of the development, evaluation and clinical translation of GRAPE. a, Model development. GRAPE takes noncontrast CT as input and outputs the probability and the segmentation mask of possible primary gastric lesions. GRAPE was trained with noncontrast CT from gastroscopy-confirmed GC patients. The performance of GRAPE was evaluated via GRAPE scores, ROC

curves and so on. **b**, Overview of the training cohort, internal validation cohort and external validation cohort. **c**, Overview of reader studies. **d**, Real-world study. The performance of GRAPE in realistic hospital opportunistic screening was validated using two real-world study cohorts, including a regional hospital cohort and a cancer center cohort.

We then analyzed the proportion of GC detected by GRAPE, radiologists alone and radiologist with GRAPE assistant in EGC and AGC. GRAPE showed a higher proportion of EGC detected compared with radiologists alone or with GRAPE assistance (Fig. 3d). We also found that GRAPE performed significantly better in the middle-third location compared with radiologists alone or with GRAPE assistance (Fig. 3e), which may result from the difficulty in identifying GC at this location with various degrees of stomach filling. Figure 3f shows that representative CT scans of T1/T2 stage GC that radiologists failed to identify.

Real-world study of hospital opportunistic screening

We further validated the performance of GRAPE in opportunistic screening using two real-world study cohorts comprising 78,593

consecutive participants with noncontrast CT between 2018 and 2024 from two independent regional hospitals and one comprehensive cancer center.

First, we validated the opportunistic screening capacity of GRAPE in a real-world cohort comprising 41,178 routine patients from diverse clinical settings (physical examination, emergency, inpatient and outpatient departments) across regional hospitals. GRAPE identified 11.28% (4,645) as high risk, with distinct proportions across cohorts: 6.2% (1,248 of 20,097) in Fenghua People's Hospital (FHPH) versus 16.1% (3,397 of 21,081) in Pinyang People's Hospital (PYPH) (Fig. 4). Through comprehensive medical record review and follow-up, we stratified GRAPE-identified high-risk patients into four clinical trajectories: confirmed GC via gastroscopy, NGC diagnoses, asymptomatic

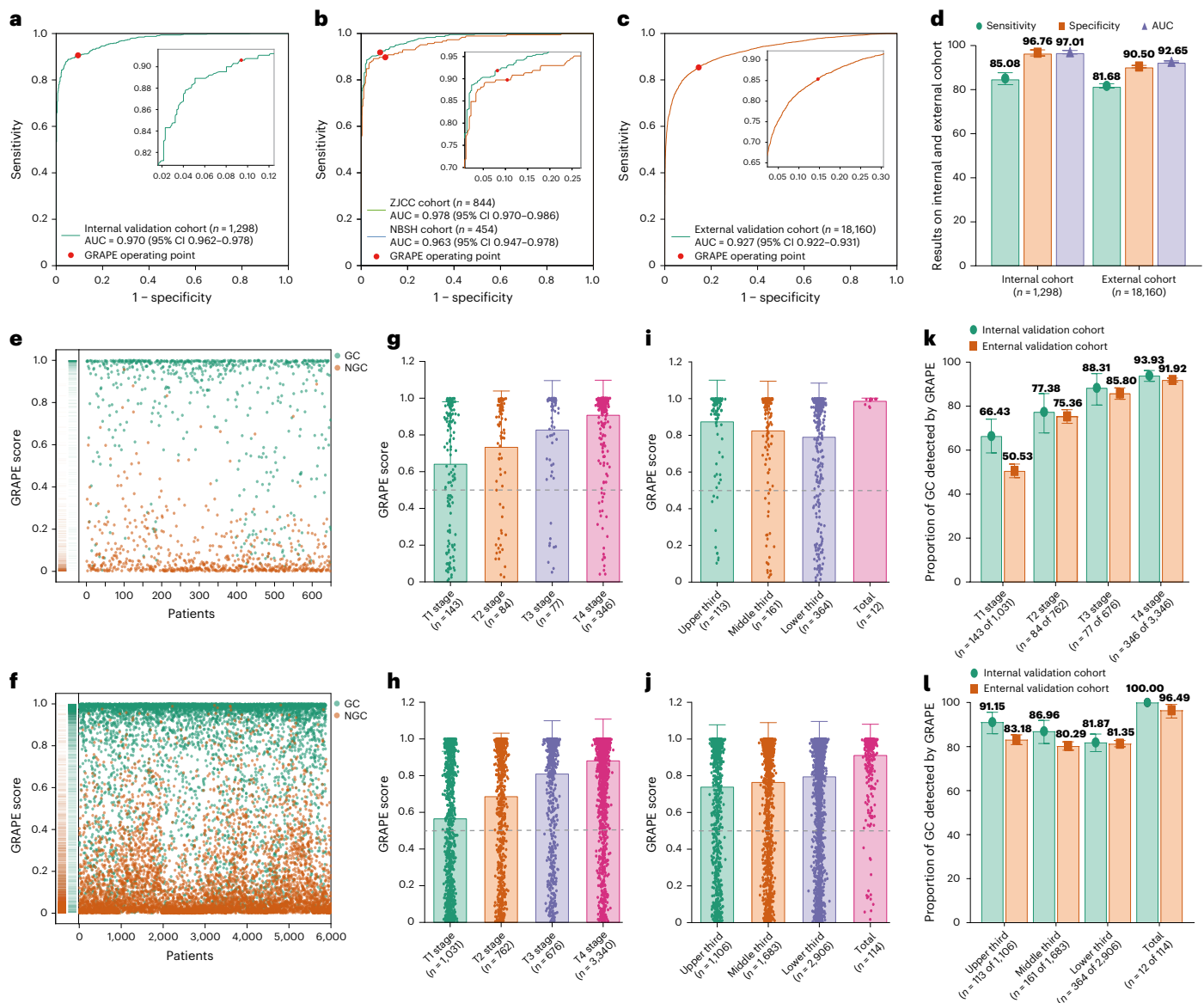


Fig. 2 | GRAPE performance in internal and external validation cohorts.

a, ROC curve of GRAPE in an internal validation cohort. **b**, ROC curve of GRAPE in ZJCC and NBSH in internal validation cohort. **c**, ROC curve of GRAPE in external validation cohort. **d**, Sensitivity, specificity and AUC of GRAPE in internal and external validation cohorts. **e, f**, Distribution of GRAPE scores of GC and NGC in the internal (**e**) and external (**f**) validation cohorts. **g, h**, GRAPE scores in different

T stages in internal (**g**) and external (**h**) validation cohorts. **i, j**, GRAPE scores in different locations in internal (**i**) and external (**j**) validation cohorts. **k**, Proportion of GC detected by GRAPE in different T stages in internal and external validation cohorts. **l**, Proportion of GC detected by GRAPE in different locations of the stomach in internal and external validation cohorts.

patients with 1-year follow-up and those lacking both gastroscopy and follow-up documentation. Finally, among the high-risk population, gastroscopic and follow-up verification revealed that 289 GC cases were confirmed in FHPH, while 556 GC in PYPH (Fig. 4a,b and Extended Data Table 4). The detection rate was 24.49% (289 of 1,180) and 17.65% (556 of 3,151) in FHPH and PYPH, respectively. Furthermore, we retrospectively analyzed the patients with confirmed GC in the ‘high-risk’ population and found that 40.48% (117 of 289) and 38.31% (213 of 556) patients did not have abdominal symptoms in FHPH and PYPH, respectively. The multidisciplinary team (MDT) searched and reviewed the available electronic health record of low-risk patients in the FHPH and PYPH cohorts, including all existing follow-up information, and found 37 and 49 GCs patients, respectively, confirmed by pathology. Thus, the estimated sensitivity and specificity are 0.887 (95% CI 0.850–0.920) and 0.951 (95% CI 0.949–0.954) for the FHPH

cohort, and 0.919 (95% CI 0.899–0.939) and 0.861 (95% CI 0.857–0.866) for the PYPH cohort, respectively.

Next, we aimed to validate the ability of GRAPE in GC detection on a comprehensive cancer center, where most patients were diagnosed with or suspected of having tumors. We consecutively collected the abdominal noncontrast CT between January 2022 and June 2024 from ZJCC. GRAPE identified 8.1% (3,045) as high risk. Comprehensive medical review stratified that 311 GC cases were confirmed (Fig. 5a and Extended Data Table 4), including 34.41% (107 of 311) patients without obvious abdominal symptoms. The overall detection rate was 12.1%. Notably, one patient was predicted high risk by GRAPE on an abdominal noncontrast CT scan in June 2024. MDT review showed that the patient underwent gastroscopy, and was diagnosed as GC in August 2024. Surgical pathology confirmed the GC was pT2N0M0 (AJCC Stage IB) (Fig. 5b).

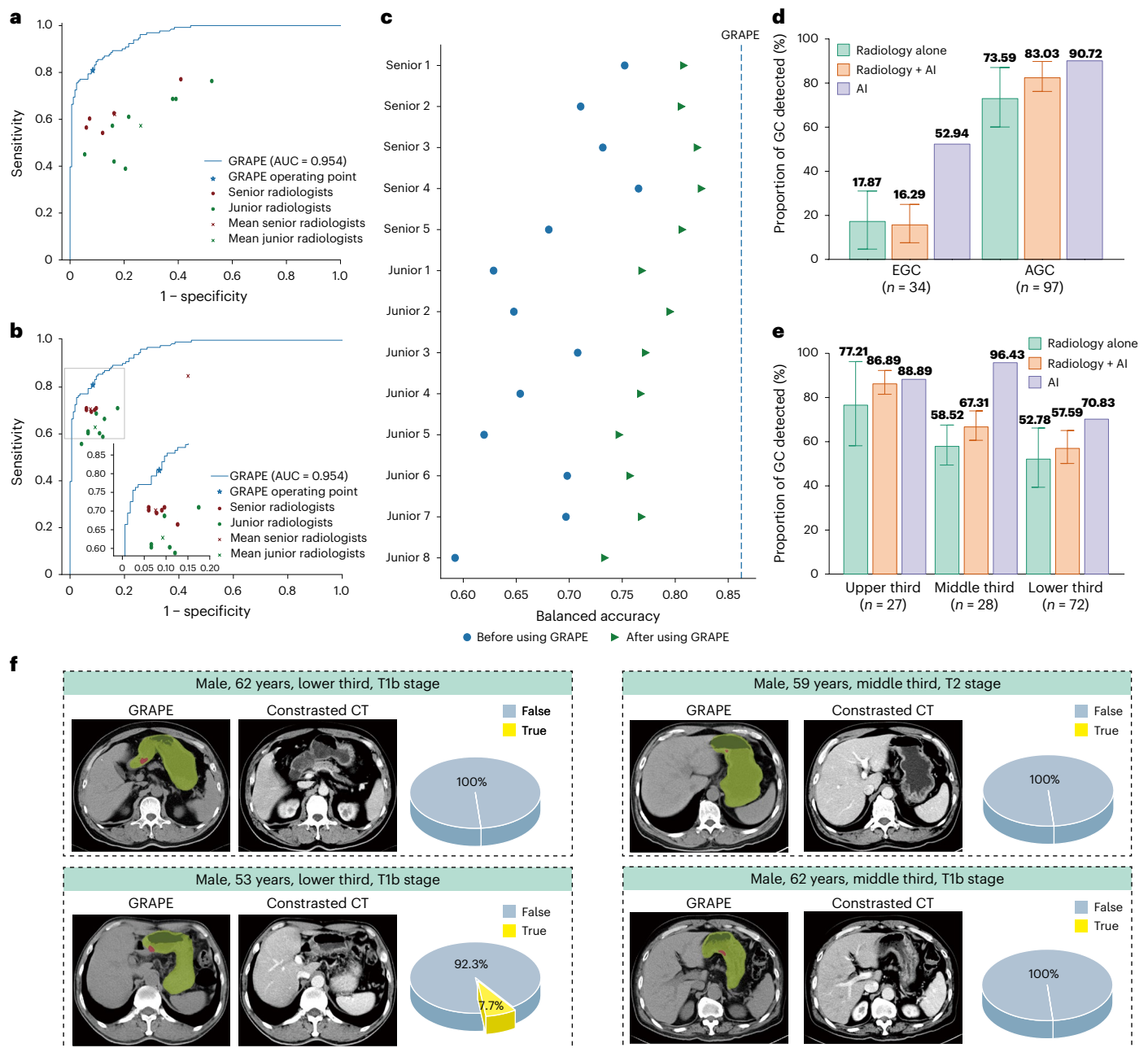


Fig. 3 | Reader study. **a**, Comparison between GRAPE and 13 readers with 13 radiologists with different levels of expertise on GC. **b**, Performance in GC discrimination of the same set of radiologists with the assistance of GRAPE on noncontrast CT. **c**, Balanced accuracy improvement in radiologists with different levels of expertise for GC discrimination. **d**, Detection rate of EGC and AGC by

radiologists alone, radiologists with the assistance of GRAPE and GRAPE alone. **e**, Detection rate of GC in different locations by radiologists alone, radiologists with the assistance of GRAPE and GRAPE alone. **f**, Examples of T1 and T2 GC discrimination by GRAPE, which were missed by readers.

We validated the ability of GRAPE in early diagnosis. Among 311 GC cases in the second real-world cohort (ZJCC), 26 patients were under follow-up, being reviewed or under treatment for other cancer, whereas the other 285 patients visited the hospital for the first time before diagnoses with GC. Among the 26 patients, 11 had prediagnostic abdominal noncontrast CT scans taken during the 6 months before GC diagnosis. GRAPE was tested on 11 prediagnostic scans and predicted GC in 63.64% (7 of 11) patients (Fig. 5c). A patient underwent two abdominal CT examinations at 14 and 6 months before diagnosis of GC due to pulmonary nodules, with no notable abnormalities of stomach reported in 2023. Later, this patient was diagnosed with poorly differentiated GC in T4aN1M0 stage via gastroscopy after more than 4 months

of abdominal discomfort in April 2024. We evaluated the noncontrast CT before and at the time of diagnosis and the results showed that the GRAPE indicated GC in noncontrast CT 6 months before diagnoses. Based on the GRAPE prediction and retrospective review of the image, the MDT suspected the patient was in stage T2, which was detected successfully by GRAPE 6 months in advance (Fig. 5d).

Discussion

CT provides a comprehensive and objective evaluation of anatomical structures, irrespective of clinical indications, establishing its significance as an effective cancer detection tool¹⁷. The emergence of AI methods has enabled the automatic segmentation and processing of

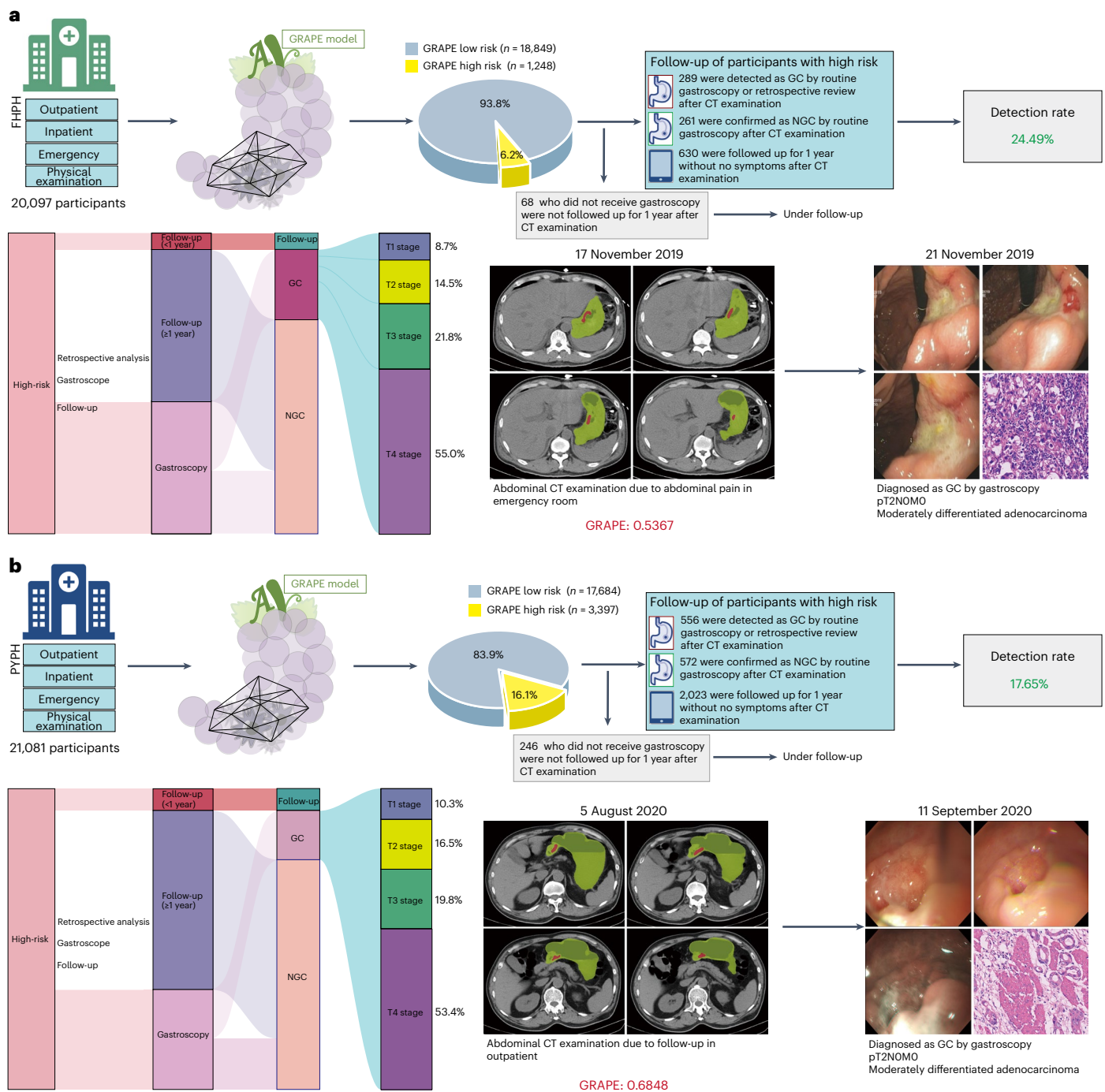


Fig. 4 | Performance of GRAPE in realistic hospital opportunistic screening in regional hospitals validated using a real-world study cohort. a, Overview of GRAPE's performance in FHPH cohort and a case study. **b,** Overview of GRAPE's performance in PYPH cohort and a case study.

underutilized data through deep learning, facilitating fast, objective algorithms for accurate lesion identification^{18,19}. In this study, we present GRAPE: an AI model based on routine noncontrast CT scans for GC detection. The effectiveness of this model was further validated in an independent, external multicenter study. GRAPE demonstrated superior detection performance for GC compared with previous models based on clinical information and serological diagnostics (AUC 0.757–0.79)^{20–22}, and the performance was comparable to liquid biopsy^{23,24}, including circular RNAs²⁵, microRNAs^{26,27}, cell-free DNA²⁸ and metabolites²³ in blood samples, which reported AUC values between 0.83 and 0.94. The subgroup analysis demonstrated GRAPE sensitivity of approximately 50% for EGC, and more than 90% for GC at T3 and

T4 stages. Although GRAPE demonstrated superior performance in advanced GC detection, it may allow GC to be detected at an earlier stage, which can also improve the overall survival of GC patients.

To validate the value of GRAPE in realistic hospital opportunistic screening, we conducted two real-world studies comprising 78,593 consecutive participants. GRAPE achieved high detection rates, significantly surpassing the 0.9% detection rate of all GC from questionnaire-based methods²⁹. GRAPE also showed good performance on prediagnosis scans performed in the 6 months before GC diagnosis. Despite its advantage, GRAPE is not intended to replace endoscopic evaluation. We believe it provides a valuable alternative for symptomatic patients hesitant to undergo initial endoscopic screening.

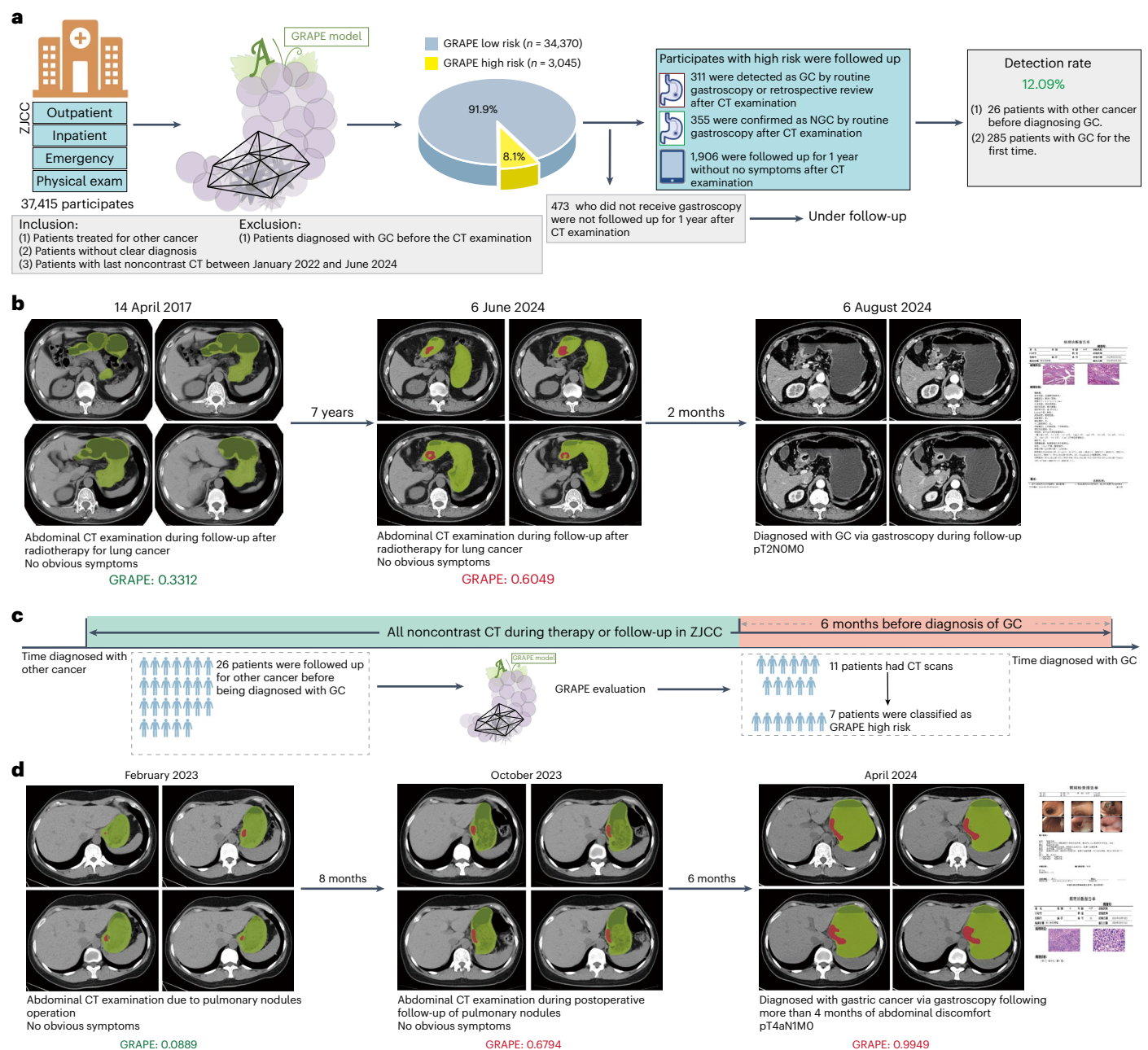


Fig. 5 | Performance of GRAPE in realistic hospital opportunistic screening in cancer center validated using a real-world study cohort. a, Overview of GRAPE's performance in the ZJCC cohort. **b**, Illustration of GC patient diagnosed after detection using GRAPE. This patient was being followed up for lung cancer treatment in June 2024 but was detected by GRAPE. MDT review showed that the patient underwent gastroscopy and was diagnosed with moderately differentiated to well-differentiated GC in August 2024. **c**, Data collection process of prediagnosis CT scans. Among 26 patients under follow-up for other cancers, 11 had prediagnostic CT scans taken in the 6 months before GC diagnosis. GRAPE suggested GC in 63.64% (7 of 11) patients in the 6 months before their diagnosis.

d, A patient underwent 2 abdominal CT examinations at 14 and 6 months before diagnosis of GC due to pulmonary nodules, with no notable abnormalities of stomach reported in 2023. Later, this patient was diagnosed with poorly differentiated GC in T4aN1M0 stage via gastroscopy after more than 4 months of abdominal discomfort in April 2024. We evaluated the noncontrast CT before and at the time of diagnosis, and the results showed that the GRAPE indicated GC in noncontrast CT 6 months before diagnoses. Based on the GRAPE prediction and retrospective review of the image, the MDT suspected the patient was in stage T2, which was detected successfully by GRAPE 6 months in advance.

Currently, no other new GC screening methods have been implemented in large-scale screening cohorts, although similar studies have been conducted for other malignancies. The National Lung Screening Trial in the United States reported a lung cancer detection rate of 2.4–5.2% using low-dose helical CT³⁰. The detection rate of ovarian and fallopian tube cancers was 5.5% using transvaginal ultrasound³¹. In liquid screening, the detection rate of nasopharyngeal carcinoma was 11.0% using Epstein–Barr virus DNA levels in plasma samples³², and the

detection rate of colorectal cancer was 3.2% via cell-free DNA detection in plasma in the theoretical ECLIPSE study population³³. As a highly accurate high-risk identification tool, GRAPE has the potential to enable large-scale screening programs by boosting the detection rate for the secondary endoscopy examinations, thus reducing GC mortality.

The GRAPE model adopts a simple yet effective architecture to the detection of GC on noncontrast CT scans by integrating both classification and segmentation into a single deep-learning framework.

Traditional methods often fall short by focusing solely on segmentation, which limits the ability to provide patient-level probability assessments, or by relying exclusively on classification approaches applied directly to organ regions of interest (ROIs), thereby reducing the interpretability of predictions. In contrast, our joint model leverages the advantages of both strategies, allowing us to benefit from detailed local textural pattern recognition of tumors while maintaining a comprehensive understanding of the stomach's overall shape and structure. By adapting the nnUNet backbone—a proven architecture known for its robust and high-performance visual feature extraction capabilities—our model achieved enhanced capacity to handle the complex visual characteristics present in medical images specific to GC. GRAPE's overall design not only ensures efficient feature learning but also offers improved generalization across diverse datasets. Moreover, in terms of interpretability, GRAPE combined segmentation and classification pipeline, providing clinicians with detailed images where tumor segmentation masks are explicitly delineated while simultaneously offering decisive classification outputs. We also visually analyzed interpretability through Grad-CAM and found the coarse heatmap corresponded well with the tumor region. Although noncontrast CT is not a routine modality for GC diagnosis, the performance of radiologists can be improved significantly via the assistance of GRAPE. This improvement could be attributed to GRAPE's tumor segmentation output, which is easy for radiologists to interpret.

Despite GRAPE's promising results, further research is required to address key aspects of its applicability. First, a large prospective screening cohort is necessary to validate GRAPE's effectiveness in GC screening more robustly. We are implementing large-scale prospective validation of GRAPE's performance through a nationwide GC screening program. Due to scenario difference, for example, high-risk versus opportunistic screening, the optimal cutoff value is expected to be selected to balance the sensitivity and specificity during clinical implementation. Second, enhancing the sensitivity of the model for early-stage GC remains a priority. We will amplify the training dataset of GRAPE with more EGC, thereby improving model sensitivity. We will also utilize more information from endoscopic and pathology reports as additional supervision to train the GRAPE model to further improve sensitivity, for example, the description of shape, texture, size and degree of invasion of the tumor. As mentioned before, patients with better gastric filling had higher sensitivity in early GC detection. As a result, we would recommend distension of the stomach before noncontrast CT imaging in our prospective trial. Third, although the prevalence of non-GC tumors such as gastrointestinal stromal tumors and gut-associated lymphoid tissue tumors are relatively low, GRAPE could be expanded to detect and diagnose these rare tumors, making it a more comprehensive screening tool. Finally, to address generalization challenges, future direction would be the incorporation of training data from more centers. From the technical aspect, the test-time adaptation³⁴ technique is a recent machine-learning approach for better model generalizability and has the potential to boost external validation performance for GRAPE. Overall, GRAPE represents a new approach for mass GC screening, demonstrating high sensitivity, specificity and detection rates, thus enhancing both the cost-effectiveness and compliance of GC screening efforts.

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41591-025-03785-6>.

References

- Smyth, E. C., Nilsson, M., Grabsch, H. I., van Grieken, N. C. & Lordick, F. Gastric cancer. *Lancet* **396**, 635–648 (2020).
- Arnold, M. et al. Is gastric cancer becoming a rare disease? A global assessment of predicted incidence trends to 2035. *Gut* **69**, 823–829 (2020).
- Lin, J.-L. et al. Global incidence and mortality trends of gastric cancer and predicted mortality of gastric cancer by 2035. *BMC Public Health* **24**, 1763 (2024).
- Xia, J. Y. & Aadam, A. A. Advances in screening and detection of gastric cancer. *J. Surg. Oncol.* **125**, 1104–1109 (2022).
- Jiang, Y. et al. Predicting peritoneal recurrence and disease-free survival from CT images in gastric cancer with multitask deep learning: a retrospective study. *Lancet Digit Health* **4**, e340–e350 (2022).
- Sun, D. et al. The effect of nationwide organized cancer screening programs on gastric cancer mortality: a synthetic control study. *Gastroenterology* **166**, 503–514 (2024).
- Allemani, C. et al. Global surveillance of trends in cancer survival 2000–14 (CONCORD-3): analysis of individual records for 37 513 025 patients diagnosed with one of 18 cancers from 322 population-based registries in 71 countries. *Lancet* **391**, 1023–1075 (2018).
- Xia, R. et al. Estimated cost-effectiveness of endoscopic screening for upper gastrointestinal tract cancer in high-risk areas in China. *JAMA Netw. Open* **4**, e2121403 (2021).
- Hu, Y. et al. Performance evaluation of four prediction models for risk stratification in gastric cancer screening among a high-risk population in China. *Gastric Cancer* **24**, 1194–1202 (2021).
- Wang, Z. et al. Nationwide gastric cancer prevention in China, 2021–2035: a decision analysis on effect, affordability and cost-effectiveness optimisation. *Gut* **71**, 2391–2400 (2022).
- Pickhardt, P. J. et al. Opportunistic screening: radiology scientific expert panel. *Radiology* **307**, e222044 (2023).
- Pickhardt, P. J. et al. Automated CT biomarkers for opportunistic prediction of future cardiovascular events and mortality in an asymptomatic screening population: a retrospective cohort study. *Lancet Digit Health* **2**, e192–e200 (2020).
- McWilliams, A. et al. Probability of cancer in pulmonary nodules detected on first screening CT. *N. Engl. J. Med.* **369**, 910–919 (2013).
- Prosper, A. E., Kammer, M. N., Maldonado, F., Aberle, D. R. & Hsu, W. Expanding role of advanced image analysis in CT-detected indeterminate pulmonary nodules and early lung cancer characterization. *Radiology* **309**, e222904 (2023).
- Cao, K. et al. Large-scale pancreatic cancer detection via non-contrast CT and deep learning. *Nat. Med.* **29**, 3033–3043 (2023).
- Jiang, Y. et al. Radiographical assessment of tumour stroma and treatment outcomes using deep learning: a retrospective, multicohort study. *Lancet Digit Health* **3**, e371–e382 (2021).
- Gillies, R. J. & Schabath, M. B. Radiomics improves cancer screening and early detection. *Cancer Epidemiol. Biomark. Prev.* **29**, 2556–2567 (2020).
- Wang, Z., Liu, Y. & Niu, X. Application of artificial intelligence for improving early detection and prediction of therapeutic outcomes for gastric cancer in the era of precision oncology. *Semin Cancer Biol.* **93**, 83–96 (2023).
- Moor, M. et al. Foundation models for generalist medical artificial intelligence. *Nature* **616**, 259–265 (2023).
- Wang, Z., Mo, T.-M., Tian, L. & Chen, J.-Q. Gastrin-17 combined with CEA, CA12-5 and CA19-9 improves the sensitivity for the diagnosis of gastric cancer. *Int. J. Gen. Med.* **14**, 8087–8095 (2021).
- Sorscher, S. *Helicobacter pylori* and gastric cancer screening. *J. Clin. Oncol.* **42**, 3162–3163 (2024).
- Cai, Q. et al. Development and validation of a prediction rule for estimating gastric cancer risk in the Chinese high-risk population: a nationwide multicentre study. *Gut* **68**, 1576–1587 (2019).
- Xu, Z. et al. Efficient plasma metabolic fingerprinting as a novel tool for diagnosis and prognosis of gastric cancer: a large-scale, multicentre study. *Gut* **72**, 2051–2067 (2023).

24. Yu, P. et al. Genomic and immune microenvironment features influencing chemoimmunotherapy response in gastric cancer with peritoneal metastasis: a retrospective cohort study. *Int J. Surg.* **110**, 3504–3517 (2024).
 25. Roy, S. et al. Diagnostic efficacy of circular RNAs as noninvasive, liquid biopsy biomarkers for early detection of gastric cancer. *Mol. Cancer* **21**, 42 (2022).
 26. So, J. B. Y. et al. Development and validation of a serum microRNA biomarker panel for detecting gastric cancer in a high-risk population. *Gut* **70**, 829–837 (2021).
 27. Mok, J. W., Oh, Y. H., Magge, D. & Padmanabhan, S. Racial disparities of gastric cancer in the USA: an overview of epidemiology, global screening guidelines, and targeted screening in a heterogeneous population. *Gastric Cancer* **27**, 426–438 (2024).
 28. Yu, P. et al. Multi-dimensional cell-free DNA-based liquid biopsy for sensitive early detection of gastric cancer. *Genome Med* **16**, 79 (2024).
 29. Zhu, X. et al. Development, validation, and evaluation of a risk assessment tool for personalized screening of gastric cancer in Chinese populations. *BMC Med* **21**, 159 (2023).
 30. Aberle, D. R. et al. Results of the two incidence screenings in the National Lung Screening Trial. *N. Engl. J. Med.* **369**, 920–931 (2013).
 31. Menon, U. et al. Sensitivity and specificity of multimodal and ultrasound screening for ovarian cancer, and stage distribution of detected cancers: results of the prevalence screen of the UK Collaborative Trial of Ovarian Cancer Screening (UKCTOCS). *Lancet Oncol.* **10**, 327–340 (2009).
 32. Chan, K. C. A. et al. Analysis of plasma Epstein–Barr virus DNA to screen for nasopharyngeal cancer. *N. Engl. J. Med.* **377**, 513–522 (2017).
 33. Chung, D. C. et al. A cell-free DNA blood-based test for colorectal cancer screening. *N. Engl. J. Med.* **390**, 973–983 (2024).
 34. Zhang, M., Levine, S. & Finn, C. Memo: test time robustness via adaptation and augmentation. *Adv. Neural Inf. Process. Syst.* **35**, 38629–38642 (2022).
- Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.
- Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.
- © The Author(s) 2025

Can Hu^{1,2,3,31}, **Yingda Xia**^{4,5,31}, **Zhilin Zheng**^{5,6,31}, **Mengxuan Cao**^{7,31}, **Guoliang Zheng**^{8,31}, **Shangqi Chen**^{9,31}, **Jiancheng Sun**^{10,31}, **Wujie Chen**⁷, **Qi Zheng**⁹, **Siwei Pan**¹, **Yanqiang Zhang**¹, **Jiahui Chen**¹⁰, **Pengfei Yu**¹, **Jingli Xu**¹, **Jianwei Xu**^{5,6}, **Zhongwei Qiu**^{5,6}, **Tiancheng Lin**^{5,6}, **Boxiang Yun**^{5,6}, **Jiawen Yao**^{5,6}, **Wenchao Guo**^{5,6}, **Chen Gao**¹¹, **Xianghui Kong**¹², **Keda Chen**¹², **Zhengle Wen**¹³, **Guanxia Zhu**¹⁴, **Jinfang Qiao**¹⁵, **Yibo Pan**¹⁶, **Huan Li**¹⁷, **Xijun Gong**¹⁷, **Zaisheng Ye**¹⁸, **Weiun Ao**¹⁹, **Lei Zhang**⁷, **Xing Yan**⁷, **Yahan Tong**⁷, **Xinxin Yang**²⁰, **Xiaozhong Zheng**²¹, **Shufeng Fan**²¹, **Jielu Cao**²¹, **Cheng Yan**²², **Kangjie Xie**², **Shengjie Zhang**², **Yao Wang**²³, **Lin Zheng**²⁴, **Yingjie Wu**²⁵, **Zufeng Ge**²⁶, **Xiyuan Tian**²⁶, **Xin Zhang**²⁷, **Yan Wang**²⁸, **Ruolan Zhang**¹, **Yizhou Wei**¹, **Weiwei Zhu**¹, **Jianfeng Zhang**^{5,6}, **Hanjun Qiu**^{10,29}, **Miaoguang Su**³⁰, **Lei Shi**⁷, **Zhiyuan Xu**^{1,2,3}, **Ling Zhang**^{4,5} & **Xiangdong Cheng**^{1,2,3}

¹Department of Gastric Surgery, Zhejiang Cancer Hospital, Hangzhou Institute of Medicine (HIM), Chinese Academy of Sciences, Hangzhou, China.

²Zhejiang Provincial Research Center for Upper Gastrointestinal Tract Cancer, Zhejiang Cancer Hospital, Hangzhou, China. ³Key Laboratory of Prevention, Diagnosis and Therapy of Upper Gastrointestinal Cancer of Zhejiang Province, Hangzhou, China. ⁴DAMO Academy, Alibaba Group, Washington, DC, USA. ⁵Hupan Laboratory, Hangzhou, China. ⁶DAMO Academy, Alibaba Group, Hangzhou, China. ⁷Department of Radiology, Zhejiang Cancer Hospital, Hangzhou Institute of Medicine (HIM), Chinese Academy of Sciences, Hangzhou, China. ⁸Department of Gastric Surgery, Liaoning Cancer Hospital, Shenyang, China. ⁹Department of Gastrointestinal surgery, Ningbo Second Hospital, Ningbo, China. ¹⁰Department of Gastrointestinal surgery, First Affiliated Hospital of Wenzhou Medical University, Wenzhou, China. ¹¹Department of Radiology, First Affiliated Hospital of Zhejiang Chinese Medical University, Hangzhou, China. ¹²Department of Urology, Wenzhou People's Hospital, Wenzhou, China. ¹³Department of Urology, Pingyang People's Hospital, Wenzhou, China. ¹⁴Department of Radiology, Longgang People's Hospital, Wenzhou, China. ¹⁵Department of Radiology, Yuyao People's Hospital, Ningbo, China. ¹⁶Department of Radiology, Cixi People's Hospital, Ningbo, China. ¹⁷Department of Radiology, Second Affiliated Hospital of Anhui Medical University, Hefei, China. ¹⁸Department of Gastric Surgery, Fujian Cancer Hospital, Fuzhou, China. ¹⁹Department of Radiology, Tongde Hospital of Zhejiang Province, Hangzhou, China. ²⁰Department of Hematology, First People's Hospital of Linping District, Hangzhou, China. ²¹Department of Radiology, Second Affiliated Hospital of Zhejiang Chinese Medical University, Hangzhou, China. ²²Department of Gastrointestinal Surgery, The People's Hospital of Qiannan, Duyun, China. ²³Department of Gastrointestinal Surgery, Huzhou People's Hospital, Huzhou, China. ²⁴Department of Radiation Oncology, Taizhou Cancer Hospital, Wenling, China. ²⁵Department of Gastrointestinal Surgery, Yinzhou People's Hospital, Ningbo, China. ²⁶Department of Radiology, Fenghua People's Hospital, Ningbo, China. ²⁷Department of Nuclear Medicine, Shengjing Hospital of China Medical University, Shenyang, China. ²⁸School of Communication and Electronic Engineering, East China Normal University, Shanghai, China. ²⁹Department of Orthopedics, Fenghua People's Hospital, Ningbo, China. ³⁰Department of Radiology, Pingyang People's Hospital, Wenzhou, China. ³¹These authors contributed equally: Can Hu, Yingda Xia, Zhilin Zheng, Mengxuan Cao, Guoliang Zheng, Shangqi Chen, Jiancheng Sun.

✉ e-mail: 289971282@qq.com; 39265825@qq.com; shilei@zjcc.org.cn; zuzy@zjcc.org.cn; ling.z@alibaba-inc.com; chengxd@zjcc.org.cn

Methods

Trial design and population

This study employed a phase design and included three independent cohorts (Fig. 1). The study was approved by the centralized institutional review board (IRB) covering 20 participating centers (IRB-2024-279, Clinicaltrials.gov registration [NCT06614179](https://clinicaltrials.gov/ct2/show/study/NCT06614179)). The training cohort was enrolled between September 2006 and June 2024 across 2 centers in China, comprising 3,470 participants with GC and 3,250 participants with NGC, aged 18–99 years. The internal validation cohort was enrolled between December 2006 and April 2024 across 2 centers in China, comprising 650 participants with GC and 648 participants with NGC, aged 20–94 years. Moreover, between January 2011 and August 2024, we conducted an independent external validation cohort, enrolled 18,160 participants from 16 centers who underwent gastroscopic examination, aged 18–80 years. The inclusion criteria for the GC group required participants to have a diagnosis of GC confirmed by gastroscopic pathological biopsy and to have undergone a noncontrast CT examination before receiving any treatment at the time of diagnosis. Participants were excluded if they underwent the noncontrast CT examination after receiving treatment. For the NGC group, inclusion criteria required participants to be confirmed as not having GC by gastroscopy and to have undergone a noncontrast CT examination within 6 months of the gastroscopy. Alternatively, NGC participants were included if they underwent a noncontrast CT examination and were confirmed without GC based on a 1-year follow-up. We conducted a comprehensive review of the medical records for all enrolled patients, including age, sex, T stage, location and so on. Full details of the internal and external cohorts, including patients' clinical information and CT acquisition parameters, are shown in Extended Data Table 1.

Finally, we validated the performance of GRAPE in realistic hospital opportunistic screening using 2 real-world study cohorts, including a cohort comprising 41,178 consecutive participants with noncontrast CT between 2018 and 2024 from 2 regional hospitals and a cohort comprising 37,415 participants with last noncontrast CT between 2022 and 2024 from a cancer center. In total, 45.19% (9,083 of 20,097) patients came from the outpatient and emergency departments and physical examination center, while 54.80% (11,014 of 20,097) patients came from department of inpatient in FHPH; 41.56% (8,761 of 21,081) patients came from the outpatient and emergency departments and physical examination center, whereas 58.44% (12,320 of 21,081) patients came from the inpatient department in PYPH. In ZJCC, we additionally excluded 4,573 patients diagnosed with GC before routine CT examination. In total, 30.66% (11,472 of 37,415) patients came from the outpatient and emergency departments and physical examination center, while 39.34% (25,943 of 37,415) patients came from the inpatient department.

Image acquisition and processing

CT images in the training cohort, internal validation cohort and external validation cohort were collected before gastroscopy. Stomach segmentation was obtained via a semi-supervised self-training approach using a combination of publicly available dataset³⁵ with manually annotated masks and our internal training set. Tumor segmentation was performed meticulously by two proficient radiologists using ITK-SNAP software (v.3.8, available at <http://www.itksnap.org>)³⁶. As GC can be distinguished readily from normal gastric tissues in portal venous-phase CT images, we delineated two-dimensional sections containing tumors to effectively outline tumor boundaries in venous-phase CT images. In cases where marked discrepancies in the delineation of the ROI arose between the two radiologists, resolution was achieved through constructive discussion, resulting in a consensus. To obtain tumor annotations on noncontrast images, we employed the DEEDS³⁷ registration algorithm to align venous-phase images with noncontrast images. The resulting deformation fields were then applied to the

annotations on the venous phase, thereby generating the transformed tumor annotations for the noncontrast phase. Finally, the transferred tumor annotations on the noncontrast phase were again delineated and confirmed via the same procedure as on the venous phase.

AI model: GRAPE

GRAPE model was developed to analyze 3D noncontrast CT scans aimed at the detection and segmentation of GC. The model yields two outputs: a pixel-level segmentation mask outlining the stomach and potential tumors, and a classification score that categorizes patients into those with GC and normal controls. The model employs a two-stage strategy. The first stage involves the localization of the stomach region within the entire CT scan. This stage is implemented using a segmentation network, nnUNet³⁸, with the configuration for 3D full resolution. The output stomach segmentation mask is used to generate a 3D bounding box of the stomach region, which is cropped and served as input for the second stage. This cropping enhances the efficiency and simplification of the subsequent stage by focusing the tumor segmentation specifically on the area of interest. In the second stage, a joint classification and segmentation network with dual branches is employed. The segmentation branch focuses on segmenting tumors in this identified stomach region. Meanwhile, the classification branch, which integrates multilevel feature maps from the segmentation branch, is designed to classify the patients as either having GC or as normal controls. Whereas noncontrast CT has inherent resolution limitations (~0.5–1 mm), GRAPE detects early GCs not through direct lesion visualization but by integrating subtle 3D morphometric patterns (for example, local wall thickening, mucosal heterogeneity) and contextual radiological features (for example, perigastric fat stranding, lymph node microcalcifications).

To address variations in CT imaging parameters across several centers, we adopted a robust preprocessing pipeline as follows. First, all scans were resampled to the median voxel spacing of the training dataset, specifically 0.717 mm × 0.717 mm × 5.0 mm, to harmonize differences in slice thickness and spatial resolution. Next, all intensity values were normalized using the mean and s.d. calculated from the training dataset to minimize interscanner intensity discrepancies. During training, fixed-size 3D patches with a size of 256 × 256 × 28 were extracted randomly from the resampled volumes, and random Gaussian noise was injected before feeding into the deep network to increase model robustness. At test time, predictions were generated using a sliding window strategy across the entire resampled volume, ensuring consistent performance regardless of original scan dimensions.

The training procedure involved a fivefold cross-validation approach by dividing the training set into five folds. Five models were trained individually using four folds, with validation on the one remaining fold, following the standard nnUNet cross-validation protocol. In the testing phase, ensemble learning was applied, averaging classification probabilities across five models to determine the classification score, and using pixel-level voting to determine segmentation output. The class with the higher probability between the two categories was chosen as the final classification for each test case. In addition, different status of stomach filling is challenging for the recognition of GC. GRAPE inherently addresses stomach filling variability through comprehensive training on a wide range of gastric volumes (fasting, <100 cm³ to over-distended, >800 cm³).

An advantage of the GRAPE model is its inherent interpretability. Given that it produces both segmentation and classification outputs, the segmentation output provides a pixel-level visual map that aids in understanding and confirming the classification results. To further enhance interpretability, we visualized the heatmap of the convolutional feature map in the classification branch using Grad-CAM (Gradient-weighted Class Activation Mapping)³⁹, which is a widely applicable technique to identify the regions contribute most importantly to the classification.

Evaluation metrics

Our main goal is a binary classification task to determine whether a patient has GC. A GRAPE score greater than 0.5 is considered as the ‘high risk’ category for calculating AUC, sensitivity, specificity and accuracy. Furthermore, we access the proportion of GC detection stratified by T stage, TNM stage, lesion location and stomach filling during CT examination.

Reader studies

The aim of the reader study was to assess the difference in performance between GRAPE and radiologists in detecting GC on noncontrast CT. A total of 13 radiologists were enrolled in this study, comprising 5 senior radiologists and 8 junior radiologists. Details of the radiologists is shown in Extended Data Table 3. The study comprised two sessions. In the first session, GRAPE’s performance was compared with that of radiologists with varying levels of expertise in GC imaging. The second session evaluated GRAPE’s potential to assist radiologists where we provided the radiologists with GRAPE’s prediction in addition to the noncontrast image. A washout period of at least 1 month was maintained between the 2 sessions for each radiologist.

Statistical analysis

The performance of the GC and NGC classification was evaluated using the AUC, sensitivity, specificity, positive predictive value and balanced accuracy. Confidence intervals were calculated based on 1,000 bootstrap replications of the corresponding data. The significance comparisons of sensitivity, specificity and balanced accuracy were conducted using permutation tests to calculate two-sided *P* values with 10,000 permutations. The threshold to determine statistical significance is $P < 0.05$. Data analysis was conducted in Python using the numpy (v.1.26.2), scipy (v.1.11.4) and scikit-learn (v.1.3.2) packages.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The datasets analyzed in this study are not publicly available due to restrictions imposed by the respective IRBs. Researchers may request access to the anonymized data and supporting clinical documentation by contacting corresponding author X.C. (chengxd@zjcc.org.cn). Access will be granted subject to IRB approval, a signed data-use agreement and will be for noncommercial academic purposes only. Requests will be processed within 6 weeks.

Code availability

The code used for the implementation of GRAPE has dependencies on internal tooling and infrastructure and is under patent protection (CN116188392; the other application number is not currently in the public domain), and thus is not able to be publicly released. All experiments and implementation details are described in sufficient detail in Methods to support replication with nonproprietary libraries. Several principal components of GRAPE are available in open-source repositories: PyTorch (<https://pytorch.org/v2.2.0>) and nnU-NetV1 (<https://github.com/MIC-DKFZ/nnUNet/tree/nnunetv1>).

References

35. Gibson, E. et al. Automatic multi-organ segmentation on abdominal CT with dense V-networks. *IEEE Trans. Med. Imaging* **37**, 1822–1834 (2018).
36. Yushkevich, P. A., Yang, G. & Gerig, G. ITK-SNAP: an interactive tool for semi-automatic segmentation of multi-modality biomedical images. *Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.* **2016**, 3342–3345 (2016).

37. Heinrich, M. P., Jenkinson, M., Brady, M. & Schnabel, J. A. MRF-based deformable registration and ventilation estimation of lung CT. *IEEE Trans. Med. Imaging* **32**, 1239–1248 (2013).
38. Isensee, F., Jaeger, P. F., Kohl, S. A. A., Petersen, J. & Maier-Hein, K. H. nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation. *Nat. Methods* **18**, 203–211 (2021).
39. Zhang, H. & Ogasawara, K. Grad-CAM-based explainable artificial intelligence related to medical text processing. *Bioengineering* **10**, 1070 (2023).

Acknowledgements

We thank the International Scientific and Technological Cooperation Base for tumor molecular diagnosis and intelligent screening. X.C. was supported by The National Key Research and Development Program of China (2021YFA0910100) and Healthy Zhejiang One Million People Cohort (K-20230085). C.H. was supported by National Natural Science Foundation of China (82304946), postdoctoral Innovative Talent Support Program (BX2023375), China Postdoctoral Science Foundation (2023M733563) and Natural Science Foundation of Zhejiang Province (LMS25H160006). Z.X. was supported by National Natural Science Foundation of China (82473489), Medical Science and Technology Project of Zhejiang Province (WKJ-ZJ-2323) and Natural Science Foundation of Zhejiang Province (LBD24H290001); S.P. was supported by National Natural Science Foundation of China (82403546).

Author contributions

X.C., Ling Zhang, C.H., Y.X. and Z.X. conceived of and designed the study. C.H., M.C., G. Zheng, S.C., J.S., W.C., Q.Z., S.P., Y.Z., J. Chen, P.Y., Jingli Xu, C.G., X.K., K.C., Z.W., G. Zhu, J.Q., Y.P., H.L., X.G., Z.Y., W.A., Lei Zhang, X. Yan, Y.T., X. Yang, X. Zheng, S.F., J. Cao, C.Y., K.X., Yao Wang, L. Zheng, Y. Wu, Z.G., X.T., X. Zhang, Yan Wang, R.Z., Y. Wei, W.Z., H.Q., M.S., S.Z. and L.S. carried out data acquisition. C.H., Y.X., Z.Z., M.C., Jianwei Xu, Z.Q., T.L., B.Y., J.Y. and W.G. carried out the data preprocessing. Y.X. and Z.Z. developed the AI model. C.H., Y.X., Z.Z., M.C., Jianwei Xu, Z.Q., T.L., B.Y., W.G., L.S., Z.X., Ling Zhang, J.Z. and X.C. analyzed and interpreted the data. X.C., Z.X. and C.H. carried out the clinical deployment. C.H., Y.X., Z.Z., M.C., Jianwei Xu, Z.Q., T.L., B.Y. and W.G. carried out the statistical analysis. C.H., Y.X. and Z.Z. wrote and revised the paper.

Competing interests

Alibaba Group has filed for patent protection (CN116188392; the other application number is not currently in the public domain) for the work related to the methods of detection of GC in noncontrast CT. Y.X., Z.Z., Jianwei Xu, Z.Q., T.L., B.Y., J.Y., W.G., J.Z. and Ling Zhang are employees of Alibaba Group and own Alibaba stock as part of the standard compensation package. The other authors declare no competing interests.

Additional information

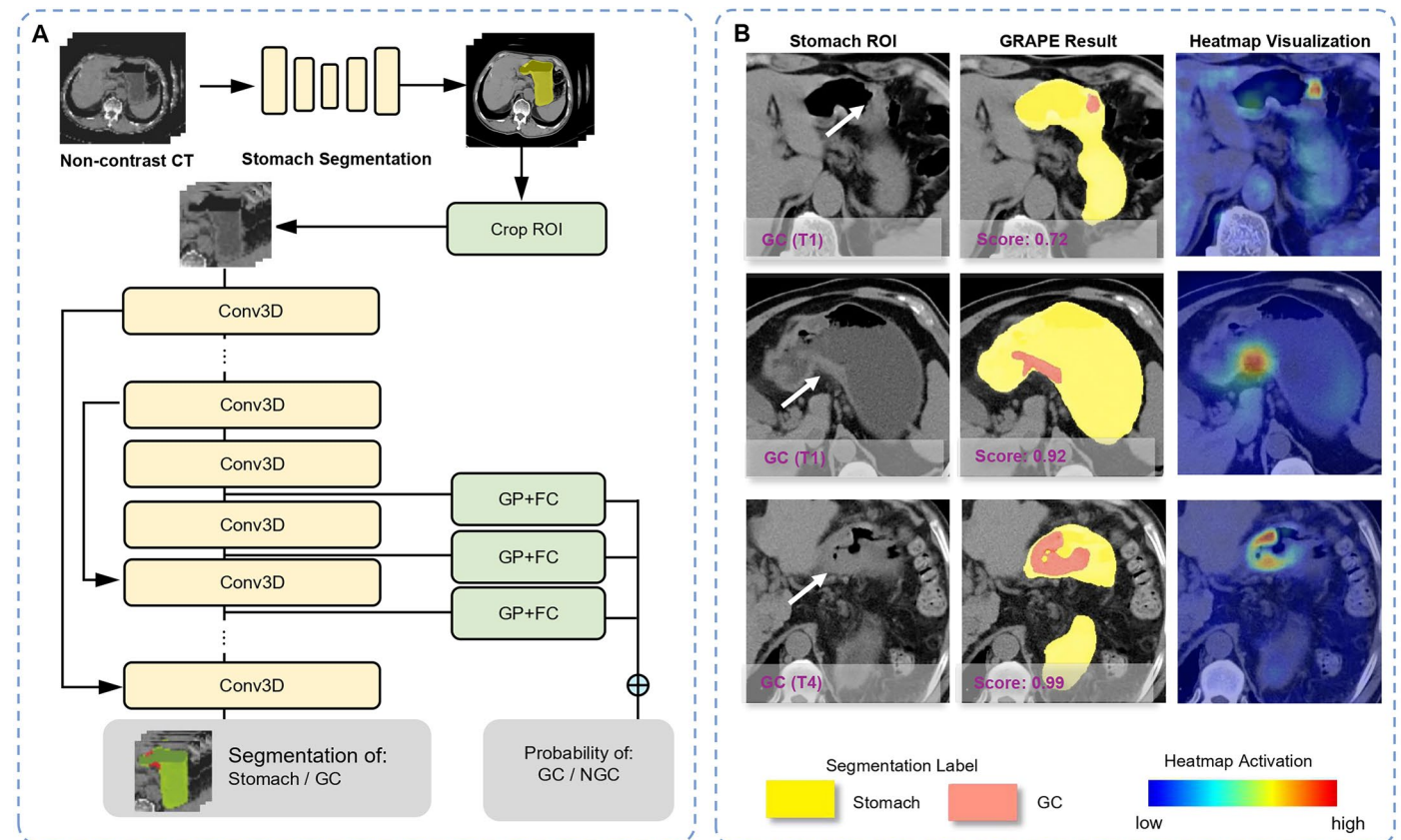
Extended data is available for this paper at <https://doi.org/10.1038/s41591-025-03785-6>.

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41591-025-03785-6>.

Correspondence and requests for materials should be addressed to Hanjun Qiu, Miaoguang Su, Lei Shi, Zhiyuan Xu, Ling Zhang or Xiangdong Cheng.

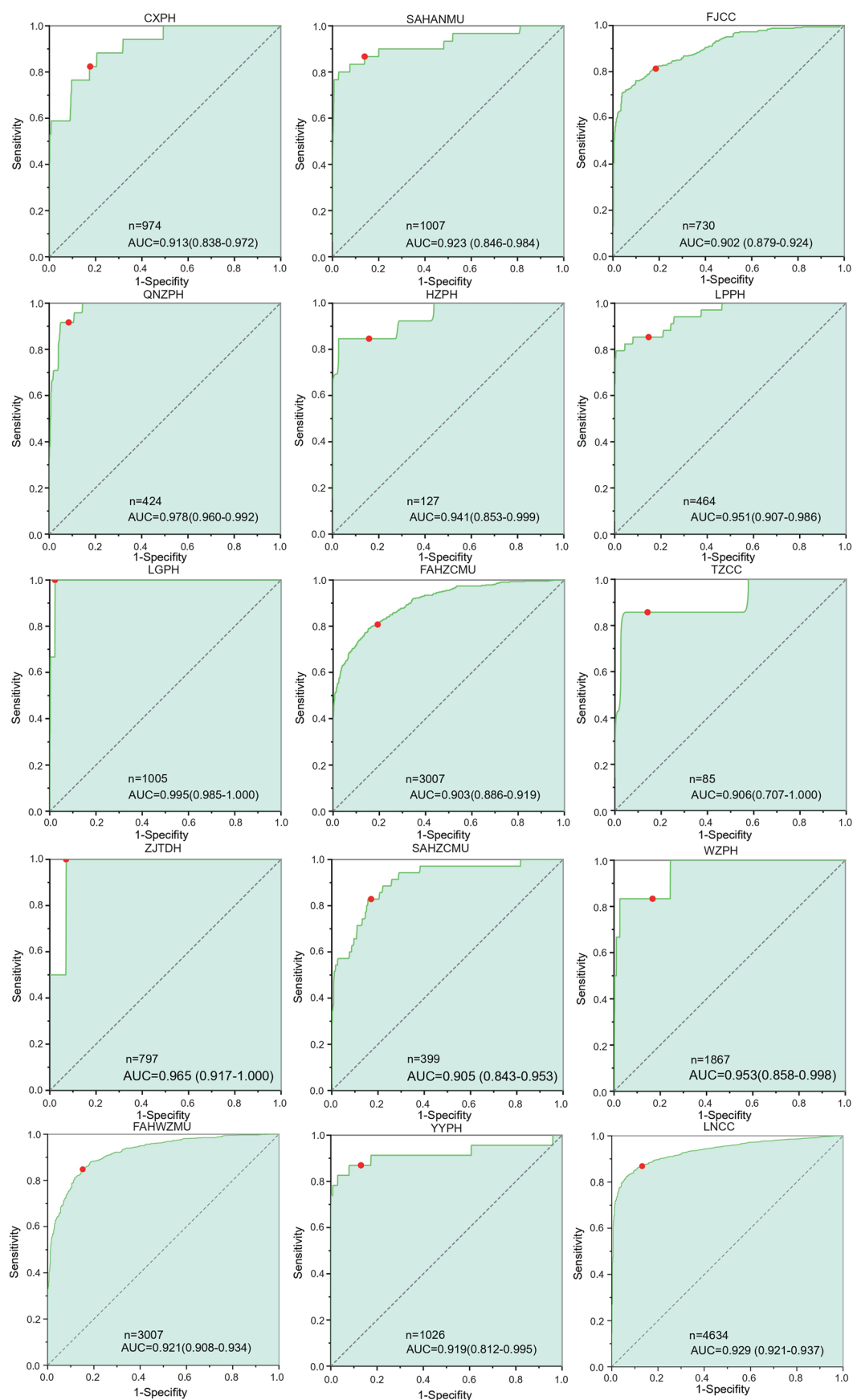
Peer review information *Nature Medicine* thanks Florian Lordick and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Primary Handling Editor: Lorenzo Righetto, in collaboration with the *Nature Medicine* team.

Reprints and permissions information is available at www.nature.com/reprints.

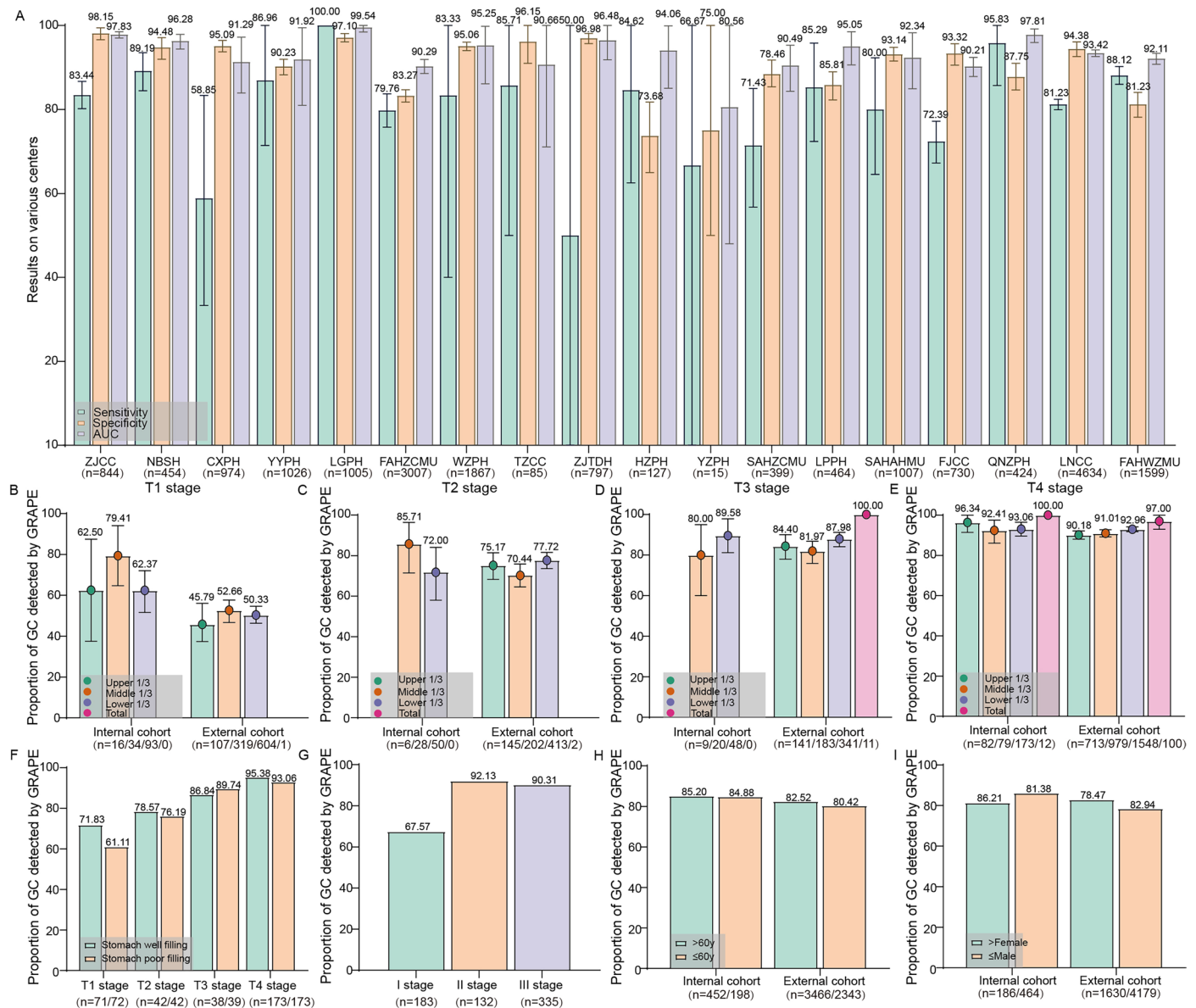


Extended Data Fig. 1 | The GRAPE model and its interpretability analysis.
a. Model workflow and architecture. The GRAPE model takes the input of a non-contrast CT scan and first segment the stomach with a U-Net to obtain the ROI of the stomach region. It then processes the ROI region with a joint segmentation and classification network which extracts the multi-level feature

of a U-Net backbone and perform classification after global pooling (GP) and fully connected layers (FC). **b.** Examples of interpretability analysis and three GC cases. The GRAPE model outputs the localization of the detected GC and aligns well with its heatmap visualization via the Grad-CAM approach.



Extended Data Fig. 2 | ROC curves of individual centers. The ROC curves show the performance of GRAPE in centers with more than 100 participants in the external validation cohort.



Extended Data Fig. 3 | The performance of GRAPE in the internal and external cohorts. a. The sensitivity, specificity and AUC of GRAPE in centers with more than 100 participants in the external validation cohort. **b–e.** The sensitivity of GRAPE in GC detection stratified by T stages and tumor locations. **f.** The sensitivity of GRAPE in GC detection in different status of stomach filling.

g. The sensitivity of GRAPE in GC detection stratified by TNM stage. **h.** The sensitivity of GRAPE in GC detection in patients younger or older than 60. **i.** The sensitivity of GRAPE in GC detection stratified by sex. Subgroups with less than 10 samples were omitted.

Extended Data Table 1 | Data characteristics of the GC and non-GC participants in the training cohort, internal validation cohort and external validation cohort

	Training cohort (n=6720)	Internal validation cohort (n=1298)	External validation cohort (n=18160)	Training cohort		Internal validation cohort		External validation cohort															
				ZJCC (n = 4753)	NBSH (n = 1967)	ZJCC (n = 844)	NBSH (n = 454)	CXPH (n = 974)	YYPH (n = 1026)	LQPH (n = 1005)	FAHZCMU (n = 3007)	WZPH (n = 1867)	TZCC (n = 85)	ZJTDH (n = 797)	HZPH (n = 127)	YZPH (n = 15)	SAHZCMU (n = 399)	LPPI (n = 464)	SAHANMU (n = 1007)	FJCC (n = 730)	QNZPH (n = 424)	LNCC (n = 4634)	FAHWZMU (n = 1599)
GC	3470	650	5809	2712	758	465	185	17	23	5	425	6	7	2	13	3	35	34	30	326	24	3958	901
Non-GC	3250	648	12351	2041	1209	379	269	957	1003	1000	2582	1861	78	795	114	12	364	430	977	404	400	676	698
Age (IQR)	61 (55-71)	61 (55-72)	59 (33-75)	61 (55-69)	62 (56-71)	61 (55-69)	62 (56-72)	56 (45-68)	60 (51-71)	46 (33-59)	55 (45-66)	49 (37-60)	61 (54-69)	50 (33-66)	67 (60-75)	65 (65-70)	57 (47-67)	54 (41-66)	54 (46-65)	59 (53-67)	50 (49-71)	62 (56-69)	64 (57-72)
Sex (Female/Male)	2245/4475	450/848	7123/11037	1629/3124	616/1351	306/538	144/310	466/508	506/520	386/619	1378/1629	782/1085	27/58	345/452	49/78	5/10	186/213	223/241	441/566	257/473	179/245	1397/3237	496/1103
T stage																							
T1	657	143	1031	539	118	129	14	2	1	0	127	0	0	0	0	0	10	0	2	62	0	577	250
T2	500	84	762	341	159	53	31	8	3	0	47	1	2	1	2	1	9	12	5	65	2	471	133
T3	476	77	676	266	210	32	45	1	4	1	77	2	3	0	4	1	2	5	4	20	7	441	104
T4	1783	346	3340	1512	271	251	95	6	15	4	174	3	2	1	7	1	14	17	19	179	15	2469	414
Location																							
Upper 1/3	759	113	1106	634	125	91	22	1	1	1	81	1	3	0	4	0	7	10	9	110	1	734	143
Middle 1/3	777	161	1683	560	217	104	57	5	6	1	88	1	0	1	4	1	10	8	11	99	8	1233	207
Lower 1/3	1872	364	2906	1471	401	261	103	11	16	3	242	4	4	1	5	2	18	14	9	114	15	1927	521
Total stomach	62	12	114	47	15	9	3	0	0	0	14	0	0	0	0	0	0	2	1	3	0	64	30
Slice Thickness mean (mm)	5.4	5.3	4.8	5.4	5.0	5.3	5.0	5.0	5.0	5.0	5.1	5.0	5.0	5.0	5.0	5.0	5.0	2.6	5.0	5.0	3.0	4.4	5.0
Slice Thickness Q1	5.0	5	5.0	5.0	5.0	5.0	5.0	5.0	5.0	5.0	5.0	5.0	5.0	5.0	5.0	5.0	5.0	2.5	5.0	5.0	3.0	3.0	5.0
Slice Thickness Q3	5.0	5	5.0	6.0	5.0	5.0	5.0	5.0	5.0	5.0	5.0	5.0	5.0	5.0	5.0	5.0	5.0	2.5	5.0	5.0	3.0	5.0	5.0
Slice Thickness IPR	0.0	0.0	0.0	1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	2.0	0.0
CTDvol mean (mGy)	11.3	10.9	12.7	15.4	9.2	15.6	9.3	-	11.0	-	8.5	12.8	16.3	13.9	7.6	-	-	8.0	8.6	16.0	13.0	14.2	12.9
CTDvol Q1	7.3	7.0	8.0	11.4	6.0	11.8	5.8	-	10.8	-	7.1	8.7	15.5	13.2	6.8	-	-	7.0	6.6	11.2	4.2	10.7	9.9
CTDvol Q3	15.7	14.4	15.7	18.6	11.3	18.8	11.5	-	10.8	-	9.4	15.9	17.6	13.9	8.1	-	-	8.7	10.3	21.5	18.3	17.0	15.2
CTDvol IPR	8.3	7.4	7.7	7.2	5.3	7.0	5.7	-	0.0	-	2.3	7.2	2.1	0.7	1.3	-	-	1.7	3.7	10.3	14.0	6.3	5.4
KVP mean	118	117	119	120	106	120	107	120	120	120	120	115	120	120	120	120	119	120	120	122	109	120	120
KVP Q1	120	120	120	120	90	120	90	120	120	120	120	100	120	120	120	120	120	120	120	120	100	120	120
KVP Q3	120	120	120	120	120	120	120	120	120	120	120	120	120	120	120	120	120	120	120	120	120	120	120
KVP IPR	0	0	0	0	30	0	30	0	0	0	0	20	0	0	0	0	0	0	0	0	20	0	0
Exposure mean (mAs)	150	149	127	149	154	148	151	23	161	33	98	191	214	104	113	140	19	259	128	64	205	170	80
Exposure Q1	98	102	65	77	123	85	117	17	160	28	66	124	200	35	101	125	19	215	99	14	145	136	26
Exposure Q3	206	206	169	210	190	208	194	28	160	38	126	235	250	169	120	157	19	305	153	100	250	194	121
Exposure IPR	108	104	104	133	67	123	77	11	0	10	60	111	50	134	20	32	0	90	54	86	105	58	95
ExposureTime mean (ms)	549	536	652	548	557	533	546	792	504	805	423	603	6767	2591	500	500	607	769	509	500	501	611	814
ExposureTime Q1	500	500	500	500	500	500	500	800	500	800	350	500	4783	600	500	503	600	800	500	500	500	500	500
ExposureTime Q3	600	600	750	600	625	600	625	800	500	800	500	752	10538	6431	500	503	600	800	500	500	500	750	600
ExposureTime IPR	100	100	250	100	125	100	125	0	0	0	150	252	5756	5831	0	0	0	0	0	0	0	250	100
Tube Current mean (mA)	295	298	267	300	268	305	273	235	274	323	208	258	370	345	136	280	252	336	178	310	306	311	260
Tube Current Q1	228	232	190	233	213	236	223	170	272	276	166	182	354	331	122	248	249	280	119	189	183	223	185
Tube Current Q3	343	341	336	349	313	349	313	287	272	375	217	290	425	350	145	313	249	378	208	422	400	375	321
Tube Current IPR	115	109	146	116	100	113	90	117	0	99	51	108	71	19	23	65	0	98	89	234	217	152	136
z_axis extent mean (mm)	363	370	443	341	466	340	465	406	334	485	376	412	340	452	441	455	422	455	331	518	520	541	435
z_axis extent Q1	250	250	400	240	445	240	445	310	230	455	366	395	265	435	415	433	410	438	275	435	426	462	420
z_axis extent Q3	470	474	490	460	495	457	490	470	450	510	435	455	360	480	463	465	440	478	345	665	621	675	445
z_axis extent IPR	220	224	90	220	50	217	45	160	220	55	69	60	95	45	48	33	30	40	70	230	195	213	25
Spacing mean (mm)	0.720	0.718	0.733	0.722	0.713	0.720	0.713	0.711	0.679	0.775	0.658	0.724	0.730	0.735	0.687	0.686	0.707	0.751	0.741	0.777	0.746	0.803	0.704
Spacing Q1	0.672	0.673	0.683	0.674	0.663	0.676	0.669	0.684	0.635	0.703	0.625	0.684	0.684	0.684	0.631	0.664	0.684	0.697	0.684	0.727	0.705	0.742	0.699
Spacing Q3	0.768	0.763	0.782	0.770	0.756	0.764	0.762	0.734	0.717	0.820	0.683	0.768	0.775	0.774	0.736	0.710	0.730	0.793	0.793	0.820	0.797	0.855	0.703
Spacing IPR	0.096	0.090	0.099	0.096	0.093	0.088	0.093	0.051	0.082	0.117	0.058	0.084	0.092	0.090	0.105	0.047	0.047	0.096	0.109	0.094	0.092	0.113	0.004

ZJCC: Zhejiang Cancer Hospital; NBSH: Ningbo Second Hospital; CXPH: Cixi People Hospital; YYPH: Yuyao People Hospital; LQPH: Longgang People Hospital; FAHZCMU: First Affiliated Hospital of Zhejiang Chinese Medical University; WZPH: Wenzhou People Hospital; TZCC: Taizhou Cancer Hospital; ZJTDE: Zhejiang Tongde Hospital; HZPH: Huzhou People Hospital; YZPH: Yinzhou People Hospital; SAHZCMU: Second Affiliated Hospital of Zhejiang Chinese Medical University; LPPI: Lingshan People Hospital; SAHANMU: Second Affiliated Hospital of Anhui Medical University; FJCC: Fujian Cancer Hospital; QNZPH: Quannanzhou People Hospital; LNCC: Liaoning Cancer Hospital; FAHWZMU: First Affiliated Hospital of Wenzhou Medical University

Extended Data Table 2 | The performance of GRAPE models in internal validation cohort and external validation

Center	Sensitivity	95% CI of Sensitivity	Specificity	95% CI of Specificity	Accuracy	95% CI of Accuracy	AUC	95% CI of AUC
Internal cohort (n=1298)	0.851	0.825-0.877	0.968	0.953-0.980	0.909	0.893-0.924	0.970	0.962-0.978
ZJCC (n = 844)	0.834	0.802-0.867	0.982	0.967-0.995	0.901	0.880-0.921	0.978	0.970-0.985
NBSH (n = 454)	0.892	0.844-0.935	0.948	0.920-0.973	0.925	0.989-0.949	0.963	0.943-0.979
External cohort (n=18160)	0.817	0.807-0.827	0.905	0.900-0.910	0.877	0.872-0.882	0.927	0.922-0.931
CXPH (n = 974)	0.588	0.333-0.812	0.951	0.936-0.965	0.945	0.930-0.958	0.913	0.838-0.972
YYPH (n = 1026)	0.870	0.706-1.000	0.902	0.883-0.921	0.902	0.883-0.919	0.919	0.812-0.995
LGPH (n = 1005)	1.000	1.000-1.000	0.971	0.961-0.981	0.971	0.960-0.981	0.995	0.985-1.000
FAHZCMU (n = 3007)	0.798	0.759-0.838	0.833	0.817-0.848	0.827	0.814-0.841	0.903	0.886-0.919
WZPH (n = 1867)	0.833	0.500-1.000	0.951	0.942-0.961	0.950	0.940-0.960	0.953	0.858-0.998
TZCC (n = 85)	0.857	0.500-1.000	0.962	0.917-1.000	0.953	0.906-0.988	0.906	0.707-1.000
ZJTDH (n = 797)	0.500	0.000-1.000	0.970	0.957-0.981	0.969	0.956-0.980	0.965	0.917-1.000
HZPH (n = 127)	0.846	0.636-1.000	0.737	0.658-0.819	0.748	0.669-0.819	0.941	0.853-0.999
YZPH (n = 15)	0.667	0.000-1.000	0.750	0.500-1.000	0.733	0.533-0.933	0.806	0.500-1.000
SAHZCMU (n = 399)	0.714	0.556-0.862	0.885	0.851-0.917	0.870	0.837-0.902	0.905	0.843-0.953
LPPH (n = 464)	0.852	0.719-0.967	0.858	0.825-0.893	0.858	0.825-0.888	0.951	0.907-0.986
SAHANMU (n =1007)	0.800	0.656-0.935	0.931	0.916-0.946	0.928	0.912-0.943	0.923	0.846-0.984
FJCC (n = 730)	0.724	0.675-0.771	0.933	0.907-0.957	0.840	0.814-0.866	0.902	0.879-0.924
QNZPH (n = 424)	0.958	0.862-1.000	0.878	0.844-0.910	0.882	0.851-0.913	0.978	0.960-0.992
LNCC (n = 4634)	0.798	0.796-0.810	0.944	0.927-0.961	0.819	0.808-0.830	0.929	0.921-0.937
FAHWZMU (n = 1599)	0.881	0.860-0.902	0.812	0.782-0.840	0.851	0.834-0.868	0.921	0.908-0.934

ZJCC: Zhejiang Cancer Hospital; NBSH: Ningbo Sencond Hospital; CXPH: Cixi People Hospital; YYPH: Yuyao People Hospital; LGPH: Longgang People Hospital; FAHZCMU: First Affiliated Hospital of Zhejiang Chinese Medical University; WZPH: Wenzhou People Hospital; TZCC: Taizhou Cancer Hospital; ZJTDH: Zhejiang Tongdei Hospital; HZPH: Huzhou People Hospital; YZPH: Yinzhou People Hospital; SAHZCMU: Second Affiliated Hospital of Zhejiang Chinese Medical University; LPPH: Linping People Hospital; SAHANMU: Second Affiliated Hospital of Anhui Medical University; FJCC: Fujian Cancer Hospital; QNZPH: Qiannanzhou People Hospital; LNCC: Liaoning Cancer Hospital; FAHWZMU: First Affiliated Hospital of Wenzhou Medical University

Extended Data Table 3 | Reader experience

Reader ID	Experience year	CT read per year (approximately)	GC CT read pear year (approximately)	Seniority
1	18	20,000	1,100	Senior
2	8	15,000	900	Senior
3	13	18,000	1,000	Senior
4	8	12,000	600	Senior
5	9	11,000	600	Senior
6	5	11,000	500	Junior
7	5	10,000	600	Junior
8	5	9,000	400	Junior
9	5	12,000	400	Junior
10	4	10,000	500	Junior
11	4	11,000	500	Junior
12	3	9,000	500	Junior
13	2	8,000	100	Junior

Readers were stratified into senior vs. junior radiologists endorsed by the National Health Commission's Physician Title Evaluation Standards (2021 Edition) and the Chinese Medical Association Radiology Branch certification guidelines. Senior status required either: (1) attainment of associate chief physician or higher academic rank; or (2) accumulated ≥8 years of independent diagnostic reporting experience, as verified through institutional credentialing databases.

Extended Data Table 4 | Data characteristics of the GRAPE-high risk group in the real world study

	FHPH (n = 1248)	PYPH (n = 3397)	ZJCC (n = 3045)
GC confirmed by gastroscopy	289	556	311
NGC confirmed by gastroscopy	261	572	355
NGC confirmed after more than one year of follow-up	630	2023	1906
Patients without gastroscopy or less than one year of follow-up	68	246	473
T stage			
T1	25	57	12
T2	42	92	49
T3	63	110	74
T4	159	297	180
Location			
Upper 1/3	58	142	121
Middle 1/3	87	127	66
Lower 1/3	128	267	114
Total stomach	16	20	14

FHPH: Fenghua People Hospital; PYPH: Pingyang People Hospital; ZJCC: Zhejiang Cancer Hospital;

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our [Editorial Policies](#) and the [Editorial Policy Checklist](#).

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

- | | |
|-------------------------------------|--|
| n/a | Confirmed |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> The statistical test(s) used AND whether they are one- or two-sided
<i>Only common tests should be described solely by name; describe more complex techniques in the Methods section.</i> |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> A description of all covariates tested |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals) |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> For null hypothesis testing, the test statistic (e.g. F , t , r) with confidence intervals, effect sizes, degrees of freedom and P value noted
<i>Give P values as exact values whenever suitable.</i> |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Estimates of effect sizes (e.g. Cohen's d , Pearson's r), indicating how they were calculated |

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

Software and code

Policy information about [availability of computer code](#)

- | | |
|-----------------|--|
| Data collection | Dicom files were handled with the open source libraries Pydicom(https://pydicom.github.io/ , version 2.2.2), SimpleITK(https://simpleitk.org/ , version 2.0.2), and NiBabel(https://nipy.org/nibabel/ , version 3.2.1). |
| Data analysis | The code used for the implementation of GRAPE has dependencies on internal tooling and infrastructure, is under patent protection (CN116188392; the other application number is not currently in the public domain) and thus is not able to be publicly released. However, all experiments and implementation details are described in sufficient detail in the Methods section to support replication with public libraries. Several major components of our work are available in open-source repositories: PyTorch (https://pytorch.org/ , v2.2.0) and nnU-NetV1 (https://github.com/MIC-DKFZ/nnUNet/tree/nnunetv1). Data analysis was conducted in Python using the numpy (version 1.26.2), scipy (version 1.11.4), and scikit-learn (version 1.3.2) packages. ITK-SNAP software (version 3.8, available at http://www.itksnap.org) |

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio [guidelines for submitting code & software](#) for further information.

Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our [policy](#)

The datasets analyzed in this study are not publicly available due to restrictions imposed by the respective institutional review boards (IRBs). Researchers may request access to the anonymized data and supporting clinical documentation by contacting the corresponding author (Xiangdong Cheng; chengxd@zjcc.org.cn). Access will be granted subject to IRB approval, a signed data use agreement, and for non-commercial academic purposes only. Requests will be processed within 6 weeks.

Research involving human participants, their data, or biological material

Policy information about studies with [human participants or human data](#). See also policy information about [sex, gender \(identity/presentation\), and sexual orientation](#) and [race, ethnicity and racism](#).

Reporting on sex and gender

We are using only retrospective data collected through clinical practice. Sex was assigned based on the government-issued ID. The datasets used in the internal training and test cohorts, and the external multi-center test cohorts have sex distributions reported in the paper. Sex-based analysis was not reported because sex was unrelated to model implementation or deployment. Self-identification gender was not collected from patients.

Reporting on race, ethnicity, or other socially relevant groupings

We are using retrospective data collected through clinical practice. Race, ethnicity, and other socially relevant groupings were not collected from the patients and were unrelated to model implementation or deployment.

Population characteristics

Full characteristics were displayed in Extended Data Table 1. This study employed a phase design and included three independent cohorts. The study was approved by the centralized institutional review board (IRB) covering 20 participating centers (IRB-2024-279, Clinicaltrials.gov number, NCT06614179). The training cohort was enrolled between September 2006 and June 2024 across 2 centers in China, comprising 3470 participants with GC and 3250 participants with NGC, aged 18 to 99 years. The internal validation cohort was enrolled between December 2006 and April 2024 across 2 centers in China, comprising 650 participants with GC and 648 participants with NGC, aged 20 to 94 years. Moreover, between January 2011 and August 2024, we conducted an independent external validation cohort, enrolled 18,160 participants from 16 centers, aged 18 to 80. Finally, we validated the performance of GRAPE in realistic hospital opportunistic screening using two real-world study cohorts, including a cohort comprising 41,178 consecutive participants with non-contrast CT between 2018 and 2024 from two basic hospitals and a cohort comprising 37,415 participants with last non-contrast CT between 2021 and 2024 from a cancer center.

Recruitment

The training cohort was enrolled between September 2006 and June 2024 across 2 centers in China, comprising 3470 participants with GC and 3250 participants with NGC, aged 18 to 99 years. The internal validation cohort was enrolled between December 2006 and April 2024 across 2 centers in China, comprising 650 participants with GC and 648 participants with NGC, aged 20 to 94 years. Moreover, between January 2011 and August 2024, we conducted an independent external validation cohort, enrolled 18,160 participants from 16 centers aged 18 to 80. The inclusion criteria for the GC group required participants to have a diagnosis of GC confirmed by gastroscopic pathological biopsy and to have undergone a non-contrast CT examination before receiving any treatment at the time of diagnosis. Participants were excluded if they underwent the non-contrast CT examination after receiving treatment. For the NGC group, inclusion criteria required participants to be confirmed as not having GC by gastroscopy and to have undergone a non-contrast CT examination within six months of the gastroscopy. Alternatively, NGC participants were included if they underwent a non-contrast CT examination and were confirmed without GC based on a one-year follow-up. We conducted a comprehensive review of the medical records for all enrolled patients, including age, gender, T stage, location and etc.

We validated the performance of GRAPE in realistic hospital opportunistic screening using two real-world study cohorts, including a cohort comprising 41,178 consecutive participants with non-contrast CT between 2018 and 2024 from two basic hospitals and a cohort comprising 37,415 participants with last non-contrast CT between 2021 and 2024 from a cancer center. All participants with confirmed GC at the time of non-contrast CT have been excluded in the cohort of cancer center.

Ethics oversight

The study was approved by the centralized institutional review board (IRB) covering 20 participating centers (IRB-2024-279, Clinicaltrials.gov number, NCT06614179).

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

- ☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	<p>The internal cohort included 8,018 participants. We randomly allocated approximately 80% of participants (GC: Non-GC at approximately 1:1) to the training cohort and 20% to the internal validation cohort. No formal sample size calculation was performed; instead, the sample size was determined based on data availability and the need to ensure a balanced representation of GC and NGC cases for model training and evaluation. The training cohort consisted of 3,470 GC and 3,250 NGC participants, aged 18 to 99 years, enrolled between September 2006 and June 2024 across two centers in China. The internal validation cohort comprised 650 GC and 648 NGC participants, aged 20 to 94 years, enrolled between December 2006 and April 2024 across the same centers. These sample sizes were deemed sufficient for model development and initial validation based on prior studies and empirical assessments of convergence and generalization.</p> <p>For external validation, we utilized an independent dataset collected between January 2011 and August 2024, consisting of 18,160 participants from 16 centers, aged 18 to 80 years. The sample size was determined based on the availability of real-world data to assess the model's generalizability across multiple institutions.</p> <p>To further evaluate real-world performance, we validated GRAPE in two hospital-based opportunistic screening cohorts. The first comprised 41,178 consecutive participants who underwent non-contrast CT between 2018 and 2024 at two basic hospitals. The second included 37,415 participants who had their most recent non-contrast CT between 2021 and 2024 at a cancer center. These large sample sizes were chosen to ensure robust assessment in clinical settings, covering a diverse patient population and reflecting real-world screening conditions.</p>
Data exclusions	<p>For the internal and external cohorts, the inclusion criteria for the GC group required participants to have a diagnosis of GC confirmed by gastroscopic pathological biopsy and to have undergone a non-contrast CT examination before receiving any treatment at the time of diagnosis. Participants were excluded if they underwent the non-contrast CT examination after receiving treatment. For the NGC group, inclusion criteria required participants to be confirmed as not having GC by gastroscopy and to have undergone a non-contrast CT examination within six months of the gastroscopy. Alternatively, NGC participants were included if they underwent a non-contrast CT examination and were confirmed without GC based on a one-year follow-up. We conducted a comprehensive review of the medical records for all enrolled patients, including age, gender, T stage, location and etc.</p> <p>In real word study of cancer center cohort:We consecutively collected the abdominal non-contrast CT between January 2022 and June 2024 from ZJCC. For patients with multiple abdominal non-contrast CT examinations, only the last performed scan was included. In this cohort, we additionally excluded 4,573 patients diagnosed with GC prior to routine CT examination, to simulate the prospective utility for the opportunistic screening purpose.</p>
Replication	<p>First, we developed GRAPE using a cohort from two centers in China, including 3,470 GC cases and 3,250 non-GC (NGC) cases. Its effectiveness and utility were validated using internal validation cohort (1,298 cases) with the AUC of 0.970 and an independent external validation cohort from 16 centers (18,160 cases) with the AUC of 0.927. Subgroup analysis showed that the proportion of GC detected by GRAPE increased with advancing T stage but showed no significant association with tumor location. Then, we compared GRAPE's interpretations with those of radiologists and explored its potential in assisting radiologists during diagnostic interpretation. Reader studies showed that GRAPE significantly outperformed the average reader performance by 21.8% in sensitivity and 14.0% in specificity, especially in early-stage GC. Finally, we validated the performance of GRAPE in realistic hospital opportunistic screening using a real-world study cohort of 78,593 consecutive participants with non-contrast CT scans in one comprehensive cancer center and two independent regional hospitals. The real-world study showed that the detection rate of GC in GRAPE-predicted high-risk group were 24.5% and 17.7% in two basic hospitals, including 23.2% and 26.8% T1/2 stage GC. In addition, GRAPE can also detect GC that easily missed by radiologists in advance in follow up due to other disease in cancer center.</p>
Randomization	<p>For the dataset in the internal training cohort and the internal test cohort, patients were randomly assigned into training and test splits. In the internal training cohort, patients were randomly assigned to training and validation in the process of the cross-validation.</p>
Blinding	<p>The internal validation cohort, the external validation cohort and the real -world clinical test cohort were not used for the development of GRAPE. In the reader studies, readers were binded to results and other clinical information, except for patient age and sex. Readers were also binded to the data collection, exact ratio of the positive patients, and blinded to other readers. Readers were blinded to the ground-truth labels and their performance after the study. In real-word study, researchers were blinded to results of consecutive participants with non-contrast CT scans before the interpretation by GRAPE.</p>

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> Antibodies
<input checked="" type="checkbox"/>	<input type="checkbox"/> Eukaryotic cell lines
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology and archaeology
<input checked="" type="checkbox"/>	<input type="checkbox"/> Animals and other organisms
<input type="checkbox"/>	<input checked="" type="checkbox"/> Clinical data
<input checked="" type="checkbox"/>	<input type="checkbox"/> Dual use research of concern
<input checked="" type="checkbox"/>	<input type="checkbox"/> Plants

Methods

n/a	Involved in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> ChIP-seq
<input checked="" type="checkbox"/>	<input type="checkbox"/> Flow cytometry
<input checked="" type="checkbox"/>	<input type="checkbox"/> MRI-based neuroimaging

Clinical data

Policy information about [clinical studies](#)

All manuscripts should comply with the ICMJE [guidelines for publication of clinical research](#) and a completed [CONSORT checklist](#) must be included with all submissions.

Clinical trial registration ClinicalTrials.gov registration: NCT06614179

Study protocol <https://clinicaltrials.gov/study/NCT06614179?cond=NCT06614179&rank=1>

Data collection The training cohort was enrolled between September 2006 and June 2024 across 2 centers in China, comprising 3470 participants with GC and 3250 participants with NGC, aged 18 to 99 years. The internal validation cohort was enrolled between December 2006 and April 2024 across 2 centers in China, comprising 650 participants with GC and 648 participants with NGC, aged 20 to 94 years. Moreover, between January 2011 and August 2024, we conducted an independent external validation cohort, enrolled 18,160 participants from 16 centers, aged 18 to 80. The inclusion criteria for the GC group required participants to have a diagnosis of GC confirmed by gastroscopic pathological biopsy and to have undergone a non-contrast CT examination before receiving any treatment at the time of diagnosis. Participants were excluded if they underwent the non-contrast CT examination after receiving treatment. For the NGC group, inclusion criteria required participants to be confirmed as not having GC by gastroscopy and to have undergone a non-contrast CT examination within six months of the gastroscopy. Alternatively, NGC participants were included if they underwent a non-contrast CT examination and were confirmed without GC based on a one-year follow-up. We conducted a comprehensive review of the medical records for all enrolled patients, including age, gender, T stage, location and etc. We validated the performance of GRAPE in realistic hospital opportunistic screening using two real-world study cohorts, including a cohort comprising 41,178 consecutive participants with non-contrast CT between 2018 and 2024 from two regional hospitals and a cohort comprising 37,415 participants with last non-contrast CT between 2021 and 2024 from a comprehensive cancer center. All participants with confirmed GC at the time of non-contrast CT have been excluded in the cohort of cancer center.

Outcomes The AUCs, sensitivity and specificity of the AI models. The detection rate of GC in AI model-predicted high-risk group

Plants

Seed stocks *Report on the source of all seed stocks or other plant material used. If applicable, state the seed stock centre and catalogue number. If plant specimens were collected from the field, describe the collection location, date and sampling procedures.*

Novel plant genotypes *Describe the methods by which all novel plant genotypes were produced. This includes those generated by transgenic approaches, gene editing, chemical/radiation-based mutagenesis and hybridization. For transgenic lines, describe the transformation method, the number of independent lines analyzed and the generation upon which experiments were performed. For gene-edited lines, describe the editor used, the endogenous sequence targeted for editing, the targeting guide RNA sequence (if applicable) and how the editor was applied.*

Authentication *Describe any authentication procedures for each seed stock used or novel genotype generated. Describe any experiments used to assess the effect of a mutation and, where applicable, how potential secondary effects (e.g. second site T-DNA insertions, mosaicism, off-target gene editing) were examined.*