SlideChat: A Large Vision-Language Assistant for Whole-Slide Pathology Image Understanding

Supplementary Material

6. SlideInstruction and SlideBench

6.1. Data Source

In this section, we present the sources of the constructed SlideInstruction and SlideBench, which are derived from ten TCGA datasets as well as the BCNB challenge dataset. The Tab. 4 provides a detailed overview of the specific number of WSIs.

Table 4. Datasets statistics

Dataset	WSIs	Report	Organ	Purpose
TCGA-BRCA	1068	✓	Breast	Train, Test
TCGA-LGG	783	✓	Brain	Train, Test
TCGA-GBM	513	✓	Brain	Train, Test
TCGA-LUAD	506	✓	Lung	Train, Test
TCGA-LUSC	474	✓	Lung	Train, Test
TCGA-HNSC	464	✓	Head and Neck	Train, Test
TCGA-BLCA	424	✓	Bladder	Train, Test
TCGA-COAD	419	✓	Colon	Train, Test
TCGA-READ	157	✓	Rectum	Train, Test
TCGA-SKCM	107	✓	Skin	Train, Test
BCNB	1058	✗	Breast	Test

6.2. Data Statistics

We have compiled statistics on the number of VQA instances for each category within SlideBench VQA (TCGA) in Tab. 5. Each subcategory contains over 500 VQA instances, ensuring a robust representation across all areas, which supports comprehensive model evaluation and facilitates in-depth performance analysis. We provide an overview of the sample sizes and detailed original label information for the seven classification tasks within the BCNB dataset in Tab. 6.

6.3. Curation Scope and Prompt

In this section, we illustrate the various dimensions of VQAs in SlideInstruction and SlideBench, ensuring comprehensive coverage of diverse pathological scenarios. This includes 3 broad categories and 13 narrow categories. Below are the contents for each category, which help to delineate their scope and meaning, thereby enabling GPT to extract high-quality question-answer pairs more effectively.

Table 5. The number of VQA corresponding to each category in SlideBench-VQA (TCGA).

Broad Category	Narrow Catgory	Number
Microscopy	Tissue Architecture and Arrangement	696
	Tumor Characteristics	562
	Cytomorphological Characteristics	601
	Histopathological Changes	633
Diagnosis	Disease Detection	581
	Disease Classification	532
	Staging	671
	Grading	601
	Differential Diagnosis	586
Clinical	Treatment Guidance	597
	Biomarker Analysis	502
	Risk Factors	591
	Prognostic Assessment	674

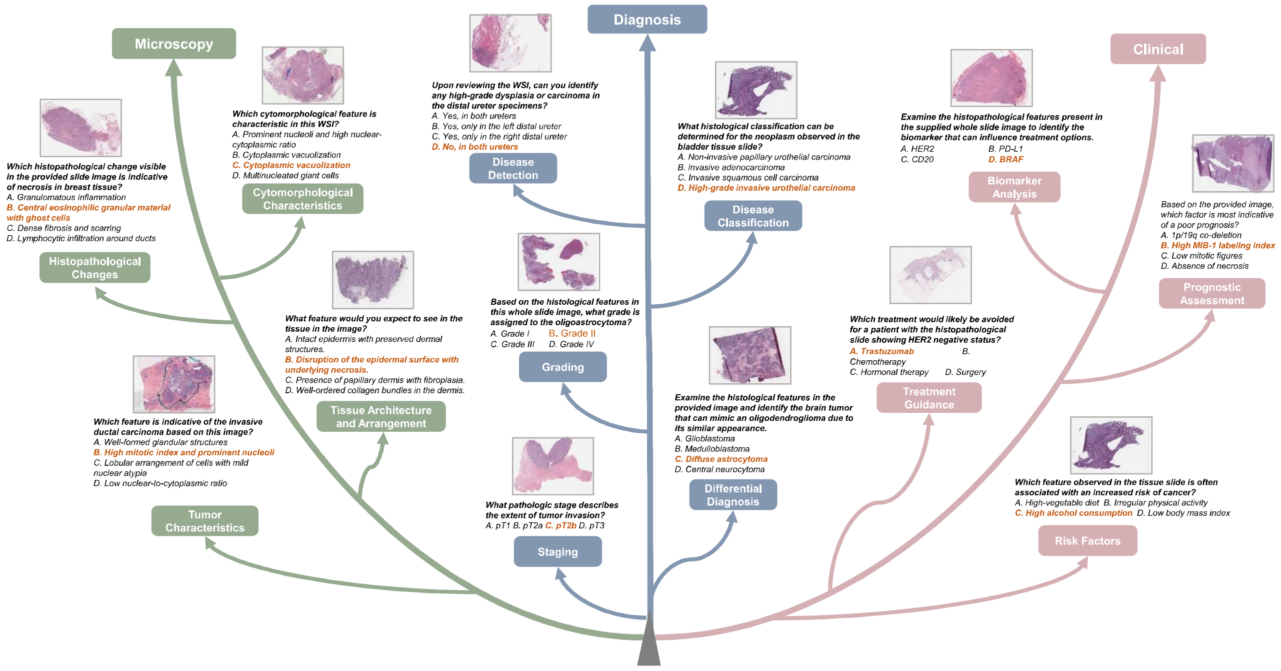
Table 6. The number and options of VQA corresponding to each task in SlideBench-VQA (BCNB).

Task	Number	Option
ER Status	1058	Postive / Negative
HR Status	1058	Postive / Negative
HER2 Status	1058	Postive / Negative
HER2 Expression	1058	0 / 1+ / 2+ / 3+
Histological Grading	926	1 / 2 / 3
Molecular Subtype	1058	Luminal A / Luminal B / HER2(+) / Triple negative
Tumor Type	1058	Invasive ductal carcinoma / Invasive lobular carcinoma / Other Type

6.3.1. Scope

Microscopy This category involves assessing the ability to generate microscopy descriptions of pathology images, focusing on clinically relevant features:

- Tissue Architecture and Arrangement: Questions in this category should evaluate the understanding of overall tissue structure and spatial organization within a histological section.
- Cytomorphological Characteristics: These questions should focus on the detailed description of individual cell morphology, including nuclear and cytoplasmic features.
- Tumor Characteristics: Questions under this category should assess the ability to identify and describe features specific to tumors, such as tumor differentiation, inva-



Which histopathological change visible in the provided slide image is indicative of necrosis in breast tissue?
A. Granulomatous inflammation
B. Central eosinophilic granular material with ghost cells
C. Dense fibrosis and scarring
D. Lymphocytic infiltration around ducts

Which cytomorphological feature is characteristic in this WSI?
A. Prominent nucleoli and high nuclear-cytoplasmic ratio
B. Cytoplasmic vacuolization
C. Cytoplasmic vacuolization
D. Multinucleated giant cells

What feature would you expect to see in the tissue in the image?
A. Intact epidermis with preserved dermal structures.
B. Disruption of the epidermal surface with underlying necrosis.
C. Presence of papillary dermis with fibroplasia.
D. Well-ordered collagen bundles in the dermis.

Which feature is indicative of the invasive ductal carcinoma based on this image?
A. Well-formed glandular structures
B. High mitotic index and prominent nucleoli
C. Lobular arrangement of cells with mild nuclear atypia
D. Low nuclear-to-cytoplasmic ratio

Upon reviewing the WSI, can you identify any high-grade dysplasia or carcinoma in the distal ureter specimens?
A. Yes, in both ureters
B. Yes, only in the left distal ureter
C. Yes, only in the right distal ureter
D. No, in both ureters

What histological classification can be determined for the neoplasm observed in the bladder tissue slide?
A. Non-invasive papillary urothelial carcinoma
B. Invasive adenocarcinoma
C. Invasive squamous cell carcinoma
D. High-grade invasive urothelial carcinoma

Based on the histological features in this whole slide image, what grade is assigned to the oligoastrocytoma?
A. Grade I B. Grade II
C. Grade III D. Grade IV

Examine the histological features in the provided image and identify the brain tumor that can mimic an oligodendroglioma due to its similar appearance.
A. Glioblastoma
B. Medulloblastoma
C. Diffuse astrocytoma
D. Central neurocytoma

What pathologic stage describes the extent of tumor invasion?
A. pT1 B. pT2a **C. pT2b** D. pT3

Examine the histopathological features present in the supplied whole slide image to identify the biomarker that can influence treatment options.
A. HER2 B. PD-L1
C. CD20 **D. BRAF**

Based on the provided image, which factor is most indicative of a poor prognosis?
A. 1p/19q co-deletion
B. High MIB-1 labeling index
C. Low mitotic figures
D. Absence of necrosis

Which treatment would likely be avoided for a patient with the histopathological slide showing HER2 negative status?
A. Trastuzumab B. Chemotherapy
C. Hormonal therapy D. Surgery

Which feature observed in the tissue slide is often associated with an increased risk of cancer?
A. High-vegetable diet B. Irregular physical activity
C. High alcohol consumption D. Low body mass index

Figure 6. Examples of generated structural VQAs in pathology across Microscopy, Diagnosis, and Clinical scenarios.

sion, and specific patterns associated with different types of tumors.

- Histopathological Changes: This category should include questions that evaluate the recognition and description of pathological changes in tissue, such as necrosis, inflammation, fibrosis, and other alterations that indicate disease processes.

Diagnosis This category tests the ability of models to suggest a reasonable diagnosis based on histological images and relevant clinical context:

- Disease Detection: Questions in this category should evaluate the model's ability to identify the presence or absence of a disease based on histological features and clinical information.
- Disease Classification: These questions should focus on distinguishing between different types or subtypes of diseases, assessing the model's capability to classify conditions accurately based on morphological and histopathological criteria.
- Grading: Questions under this category should assess the model's ability to determine the grade of a disease, particularly tumors, based on the degree of differentiation and cellular atypia observed in histological images.
- Staging: This category should include questions that evaluate the ability to assign a stage to a disease, particularly in oncology, by assessing the extent of disease spread and involvement of surrounding tissues or organs.
- Differential Diagnosis: Questions should test the model's ability to provide a differential diagnosis, distinguishing between multiple potential conditions that may present with similar histological and clinical features.

Clinical This category tests the ability of models to retrieve and apply clinically relevant background knowledge about diseases:

- Treatment Guidance: Questions in this category should assess the model's ability to recommend appropriate treatment options based on the disease in question, considering factors such as disease stage, patient demographics, and any specific clinical guidelines.
- Prognostic Assessment: These questions should focus on evaluating the model's ability to predict the likely course and outcome of a disease, including survival rates, potential complications, and long-term outcomes based on clinical and pathological data.
- Risk Factors: Questions under this category should test the model's knowledge of risk factors associated with specific diseases, including genetic, environmental, and lifestyle factors that may influence disease development or progression.
- Biomarker Analysis: This category should include questions that evaluate the ability to identify and interpret biomarkers relevant to the diagnosis, prognosis, or treat-

ment of diseases, emphasizing their role in personalized medicine and targeted therapy.

6.3.2. Designed Prompts

Report Cleaning Prompt.

Report Cleaning Prompt. The prompt used to clean up the report from the original TCGA report is represented in Tab. 7. This process effectively eliminates extraneous noise from the report, thereby establishing a more solid foundation for caption and QA pairs generation.

Table 7. Prompts for report clean and caption generation.

> **[Report Clean Prompt]** This is the content from the pathology report. Please remove some redundant irrelevant information from the original report, such as technical details of pathology department procedures, Symbols unrelated to the pathological report, specimen handling and processing information, redundant administrative or legal statements, and some repeated information. Show me the cleaned report content.
>
> **[Caption Generation Prompt]** Based on the above pathological report content, generate a detailed paragraph that summarizes the essential pathological findings. The paragraph should include key information such as the diagnosis, tumor characteristics, margin status, lymph node involvement, and other relevant pathological findings. The summary should not mention the source being a report and should exclude any specific sizes or measurements. The paragraph should be written in a clear and cohesive manner, covering all important points without unnecessary details.

Caption Generation Prompt. The prompt used for caption generation from the refined report is detailed in Tab. 7, ensuring that the generated caption effectively captures essential summarized information in report.

Question-Answers Generation Prompt. The prompt used to extract QA from reports mainly consist of 4 parts (*i.e.*, <Cleaned Report>+ System Prompt + Objective Prompt + General Prompt), and the detailed content of each part is illustrated in Tab. 8

Table 8. Question-Answers generation prompts, including system prompt, general prompt and objective prompt.

> **[System Prompt]** You are an AI assistant proficient in digital pathology. You will receive a pathology report for whole slide images.
>
> **[General Prompt]** Based on the above pathological report content, your task is to use the provided information, create 2 multi-choice questions amd 2 short-answer questions for each narrow category. The design question should be able to be answered based on the content of the image. Design medical questions very carefully and only ask questions when you are sure of the answer. Answers should be specific and avoid ambiguity. When generating questions, it is necessary to indicate their broad category and narrow category. For multi-choice questions, you should (1) “question type” is “multi-choice questions”. (2) Provide the options and answer and reasoning. Provide four answer choices (A, B, C, and D), ensuring that one choice is correct and the other three are plausible but incorrect. (3) Aim to include one answer that is incorrect but very similar to the correct one to increase the difficulty level. For short-answer questions: (1) “question type” is “short-answer questions”. (2) Generating questions with different content from multiple-choice questions. For all questions: (1) Do not mention that the information source is report in “question”, “anwser”. (2) Return JSON format in “question type”: xxx, “question”: xxx, “options”: [], “answer”: xxx, “broad category”: xxx, “narrow category”: xxx for each question. The “options” section is empty for short-answer questions.
>
> **[Objective Prompt]** Definition of Broad Category and its corresponding Narrow Categories. “ The required broad category is Microscopy, which involves assessing the ability to generate microscopy descriptions of pathology images, focusing on clinically relevant features. For the narrow category: Tissue Architecture and Arrangement: Questions should evaluate the understanding of overall tissue structure and spatial organization within a histological section.”

Label Transformation Prompt. The prompt for transforming BCNB dataset is illustrated in Tab. 9. We employ GPT to transform individual labels into a question-answer format based on the task type and corresponding classification labels, facilitating the testing of MLLM. For instance, in the context of a tumor type classification task, <Task>represents “Tumor Type”, while <label 1>, <label 2>, and <label 3>are “Invasive ductal carcinoma”, “Invasive lobular carcinoma”, and “Other Type”, respectively, enabling the generation of relevant QA pairs.

6.4. Multimodal Dataset Comparsion

Recently, several multimodal pathology datasets have been introduced for pathology applications. However, these datasets are often constrained in both scope and scale, as they primarily focus on either patch-level analysis or limited

Table 9. Prompt for Converting Labels into QA Pairs

> **[Label Transformation Prompt]** Please create prompts for pathology image classification tasks concerning <Task>, transforming traditional labels into a multi-choice question-and-answer format. The original labels include <label 1>, <label 2>, ...

available data. In contrast, our proposed SlideInstruction and SlideBench, provided as open-source resources, significantly expand the dataset size while enhancing its versatility, as shown in Tab. 10.

Table 10. Comparisons of our datasets with other pathology datasets.

Dataset	Level	Data Type	Curation Type	Scope	Number	Availability
PathChat [21]	Patch	Patch and Q/A pairs	Human+GPT	-	257,004	✗
Quilt-Instruct [23]	Patch	Patch and Q/A pairs	GPT	-	107,131	✓
WSI-VQA [8]	Slide	WSI and Q/A pairs	GPT	-	8,672	✓
PathText [7]	Slide	WSI-Caption pairs	GPT	-	9,009	✓
HistGen [13]	Slide	WSI-Reports pairs	GPT	-	7,753	✓
Prov-Path [31]	Slide	WSI-Reports pairs	GPT	-	17,383	✗
CR-PathNarratives [34]	Slide	WSIs with annotations	Human	-	174	✗
PathAlign [3]	Slide	WSI-Reports pairs	Human	-	354,089	✗
Our SlideInstruction	Slide	WSI and Q/A pairs	GPT	13	179,935	✓
Our SlideBench	Slide	WSI and Q/A pairs	Human+GPT	13	15,835	✓

7. Experiment

7.1. Computational Cost Analysis

To evaluate the computational cost of our model architecture, we measured both the inference time and GPU memory consumption throughout the entire pipeline. This pipeline includes the patch-level encoder, slide-level encoder, multimodal projector module, and large language model, all executed on an A100 GPU. After extracting the local and global features of WSIs, the average response time was within 1 second, and GPU memory consumption was approximately 27 GB. The inference time and GPU memory consumption remained well within acceptable limits for gigapixel whole slide images.

7.2. Implementation Details

We preprocessed each WSI by segmenting it into 224×224 nonoverlapping patches at a 20× magnification level, excluding background regions. We implemented our model using the Xtuner [11] toolkit and trained it across two stages on 8 × NVIDIA A100 GPUs. The training process consists of an alignment phase followed by instruction fine-tuning: Stage 1: We freeze the LLM and train the Projection and Slide Encoder with WSI-caption data for 3 epochs, using a learning rate of 0.001. Stage 2: We unfreeze the LLM, Slide Encoder, and Projection, training the model on WSI instruction-following data for 1 epoch, with a learning rate of 0.00002. Both stages are optimized using AdamW.

7.3. Ability Showcase

7.3.1. Captioning Ability

The examples shown in Fig. 7 illustrate the capability of our model, SlideChat, to effectively perform whole-slide image captioning tasks. SlideChat demonstrates its proficiency in generating detailed and contextually accurate summaries for complex pathological whole-slide images, accurately capturing key clinical findings and pathological features. Whether summarizing broad findings, explaining pivotal details, or highlighting core results, SlideChat showcases an advanced understanding of whole-slide images, providing concise yet informative reports that align with clinical terminology and expectations.

7.3.2. VQA Ability

Fig. 8 showcases the conversational examples of SlideChat, demonstrating its ability to accurately answer a range of questions based on WSIs, covering diverse aspects such as histological classifications, tumor grading, lymph node involvement, and treatment decisions. SlideChat effectively interprets complex pathological data, engages in nuanced question-and-answer exchanges, and delivers clinically relevant responses. This reflects its potential as an intelligent assistant capable of supporting pathologists in diagnostic decision-making by providing insightful, context-aware dialogue grounded in visual pathology data.

7.3.3. Comparing Model Outputs

Fig. 9 presents a comparative analysis of the outputs from SlideChat and other models within SlideBench. The examples illustrate SlideChat's remarkable capacity to precisely classify tumors, identify distinct histological features, and describe the structural organization of tumor cells from WSIs. SlideChat demonstrates a unique proficiency in capturing both local and global features—seamlessly integrating detailed microscopic characteristics with broader contextual understanding to deliver accurate and clinically meaningful interpretations. In contrast, existing models are limited to processing small pathology images, often yielding ambiguous or incorrect classifications. This underscores SlideChat's advanced capability in comprehending whole-slide images by incorporating both intricate details and a comprehensive visual perspective.

7.4. Detailed Test Performance

7.4.1. Performance on SlideBench-VQA (TCGA)

The results presented in the tables demonstrate a comprehensive evaluation of SlideChat's performance on SlideBench-VQA (TCGA) in comparison to other existing models across microscopy, diagnosis, and clinical tasks. In

Method	Input	SlideBench-VQA(TCGA) **Microscopy**				
		Tissue Architecture and Arrangement	Tumor Characteristics	Cytomorphological Characteristics	Histopathological Changes	Overall
Random	Text	23.70	22.42	23.63	27.80	24.44
GPT-4		40.83	40.28	41.71	37.46	39.62
GPT-4o	Patch	65.94	66.20	60.10	59.23	62.89
MedDr		75.04	75.78	70.10	72.23	73.30
LLaVA-Med		50.04	40.63	40.38	56.95	47.34
Quilt-LLaVA		65.26	54.04	50.66	59.55	57.76
GPT-4o	Slide (T)	37.07	38.76	39.93	37.60	38.28
MedDr		71.58	71.27	69.87	69.05	70.48
LLaVA-Med		51.80	45.02	36.27	49.01	45.82
Quilt-LLaVA		53.59	45.37	43.09	53.24	49.12
SlideChat	Slide	88.07 (+13.03)	87.01 (+11.23)	88.02 (+17.92)	87.36 (+15.13)	87.64 (+14.34)

Method	Input	SlideBench-VQA(TCGA) **Diagnosis**					
		Disease Detection	Disease Classification	Staging	Grading	Differential Diagnosis	Overall
Random	Text	25.82	24.06	24.14	26.12	24.40	24.91
GPT-4		27.12	31.07	22.27	27.45	38.70	29.09
GPT-4o	Patch	50.27	55.94	39.94	39.66	49.66	46.69
MedDr		59.11	61.11	48.66	52.97	68.83	57.78
LLaVA-Med		37.25	28.57	30.41	20.71	47.27	32.78
Quilt-LLaVA		40.74	39.3	32.32	28.96	39.52	35.96
GPT-4o	Slide (T)	22.95	26.76	18.06	21.06	27.82	23.10
MedDr		54.29	56.40	48.66	43.52	61.61	52.47
LLaVA-Med		27.87	25.19	24.07	24.96	36.18	27.58
Quilt-LLaVA		32.47	28.25	20.18	22.96	32.25	26.97
SlideChat	Slide	80.90 (+21.79)	76.12 (+15.01)	68.41 (+19.75)	68.39 (+15.42)	73.72 (+4.89)	73.27 (+15.49)

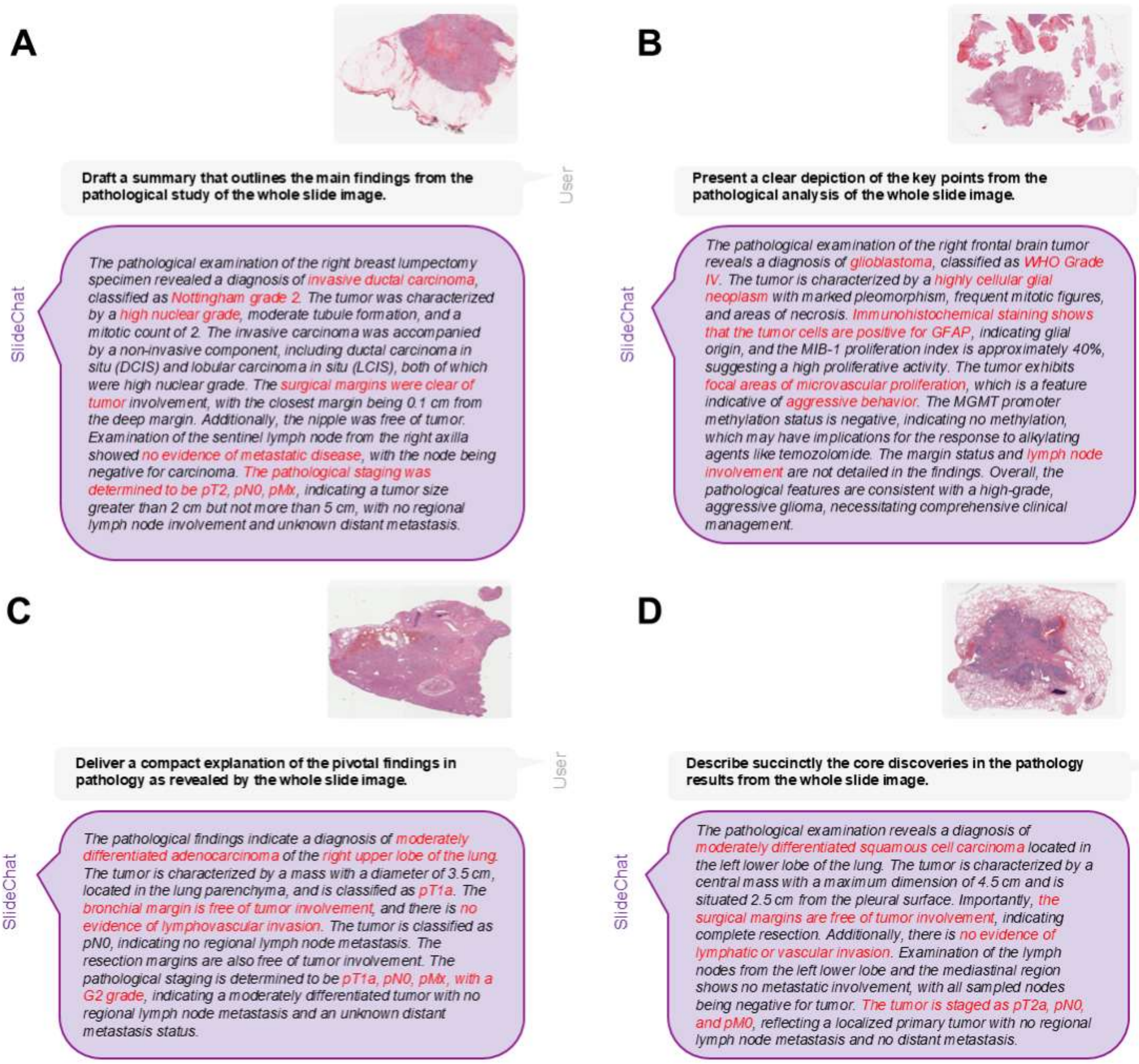
microscopy, SlideChat significantly outperforms its counterparts, achieving a notable overall accuracy improvement of 14.34 points over the nearest model. This strong performance is consistent across sub-tasks, such as tissue architecture analysis, tumor characteristics identification, and cytomorphological assessment, showcasing SlideChat's advanced capability to analyze both detailed cellular structures and broader histopathological changes. In the diagnostic tasks, SlideChat also demonstrates superior accuracy, with an overall gain of 15.49 points, excelling in disease detection, classification, staging, grading, and differential diagnosis. The clinical analysis results further validate the model's strength, with SlideChat outperforming other methods by 10.01 points overall, particularly excelling in treatment guidance, biomarker analysis, and risk factor assessment. These results illustrate SlideChat's capability to seamlessly handle complex medical data and deliver reliable insights across multiple clinical and diagnostic domains, indicating its potential as a robust tool for comprehensive pathology analysis.

7.4.2. Performance on SlideBench-VQA (BCNB)

The evaluation of SlideChat on SlideBench-VQA (BCNB), a real-world dataset designed for zero-shot testing, further underscores its ability to generalize effectively to unseen data. SlideChat demonstrates an overall accuracy improvement of 12.71 points compared to other models, showcasing its ability to generalize well across diverse and complex breast cancer-related tasks. SlideChat's performance is particularly strong in identifying tumor type, ER status, PR status, and HER2 status, demonstrating a nuanced understanding of critical histopathological features. Nevertheless, in the more complex tasks of HER2 Expression, Histological Grading, and Molecular Subtype classification, SlideChat still exhibits potential for improvement, highlighting specific areas that warrant further refinement to enhance its overall performance.

Method	Input	SlideBench-VQA(TCGA) **Clinical**				Overall
		Treatment Guidance	Biomarker Analysis	Risk Factors	Prognostic Assessment	
Random	Text	23.62	31.87	24.36	24.33	26.44
GPT-4		49.98	44.63	46.46	39.64	45.00
GPT-4o	Patch	64.18	57.99	76.99	66.64	66.77
MedDr		74.18	82.99	82.43	60.66	74.25
LLaVA-Med		62.04	53.98	53.04	26.54	47.96
Quilt-LLaVA		64.79	40.42	63.40	43.06	53.07
GPT-4o	Slide (T)	50.00	50.08	44.16	32.64	43.42
MedDr		71.43	84.51	78.92	60.24	72.80
LLaVA-Med		50.50	48.01	48.90	19.88	40.84
Quilt-LLaVA		47.38	32.93	47.71	48.64	44.76
SlideChat	Slide	83.42 (+9.24)	89.04 (+4.53)	91.71 (+9.28)	74.93 (+8.29)	84.26 (+10.01)

Method	Input	SlideBench-VQA(BCNB)							Overall
		Tumor Type	ER Type	PR Type	HER2 Type	HER2 Expression	Histological Grading	Molecular Subtype	
Random	Text	23.82	24.48	25.05	25.05	24.39	24.41	23.63	24.40
GPT-4		0	0	0	0	0	0	0	0
GPT-4o	Patch	34.69	77.50	63.51	36.95	23.95	28.63	23.15	41.43
MedDr		45.46	23.53	25.99	71.81	22.73	30.28	15.49	33.67
LLaVA-Med		23.95	36.62	40.19	50.76	23.72	18.99	15.05	30.10
Quilt-LLaVA		77.14	68.58	42.63	58.17	23.18	18.23	19.82	44.43
GPT-4o	Slide (T)	0	0	0	0	0	0	0	0
MedDr		28.92	45.84	25.71	72.68	20.65	29.96	23.88	35.48
LLaVA-Med		0.01	0	0.01	0.02	0	0	0	0.01
Quilt-LLaVA		67.41	66.73	36.58	62.67	15.97	22.89	16.27	41.55
SlideChat	Slide	90.17 (+44.71)	78.54 (+1.04)	68.81 (+5.3)	71.93 (-0.75)	25.05 (+0.66)	23.11 (-7.17)	17.49 (-6.39)	54.14 (+12.71)



Draft a summary that outlines the main findings from the pathological study of the whole slide image.

The pathological examination of the right breast lumpectomy specimen revealed a diagnosis of invasive ductal carcinoma, classified as Nottingham grade 2. The tumor was characterized by a high nuclear grade, moderate tubule formation, and a mitotic count of 2. The invasive carcinoma was accompanied by a non-invasive component, including ductal carcinoma in situ (DCIS) and lobular carcinoma in situ (LCIS), both of which were high nuclear grade. The surgical margins were clear of tumor involvement, with the closest margin being 0.1 cm from the deep margin. Additionally, the nipple was free of tumor. Examination of the sentinel lymph node from the right axilla showed no evidence of metastatic disease, with the node being negative for carcinoma. The pathological staging was determined to be pT2, pN0, pMx, indicating a tumor size greater than 2 cm but not more than 5 cm, with no regional lymph node involvement and unknown distant metastasis.

Present a clear depiction of the key points from the pathological analysis of the whole slide image.

The pathological examination of the right frontal brain tumor reveals a diagnosis of glioblastoma, classified as WHO Grade IV. The tumor is characterized by a highly cellular glial neoplasm with marked pleomorphism, frequent mitotic figures, and areas of necrosis. Immunohistochemical staining shows that the tumor cells are positive for GFAP, indicating glial origin, and the MIB-1 proliferation index is approximately 40%, suggesting a high proliferative activity. The tumor exhibits focal areas of microvascular proliferation, which is a feature indicative of aggressive behavior. The MGMT promoter methylation status is negative, indicating no methylation, which may have implications for the response to alkylating agents like temozolomide. The margin status and lymph node involvement are not detailed in the findings. Overall, the pathological features are consistent with a high-grade, aggressive glioma, necessitating comprehensive clinical management.

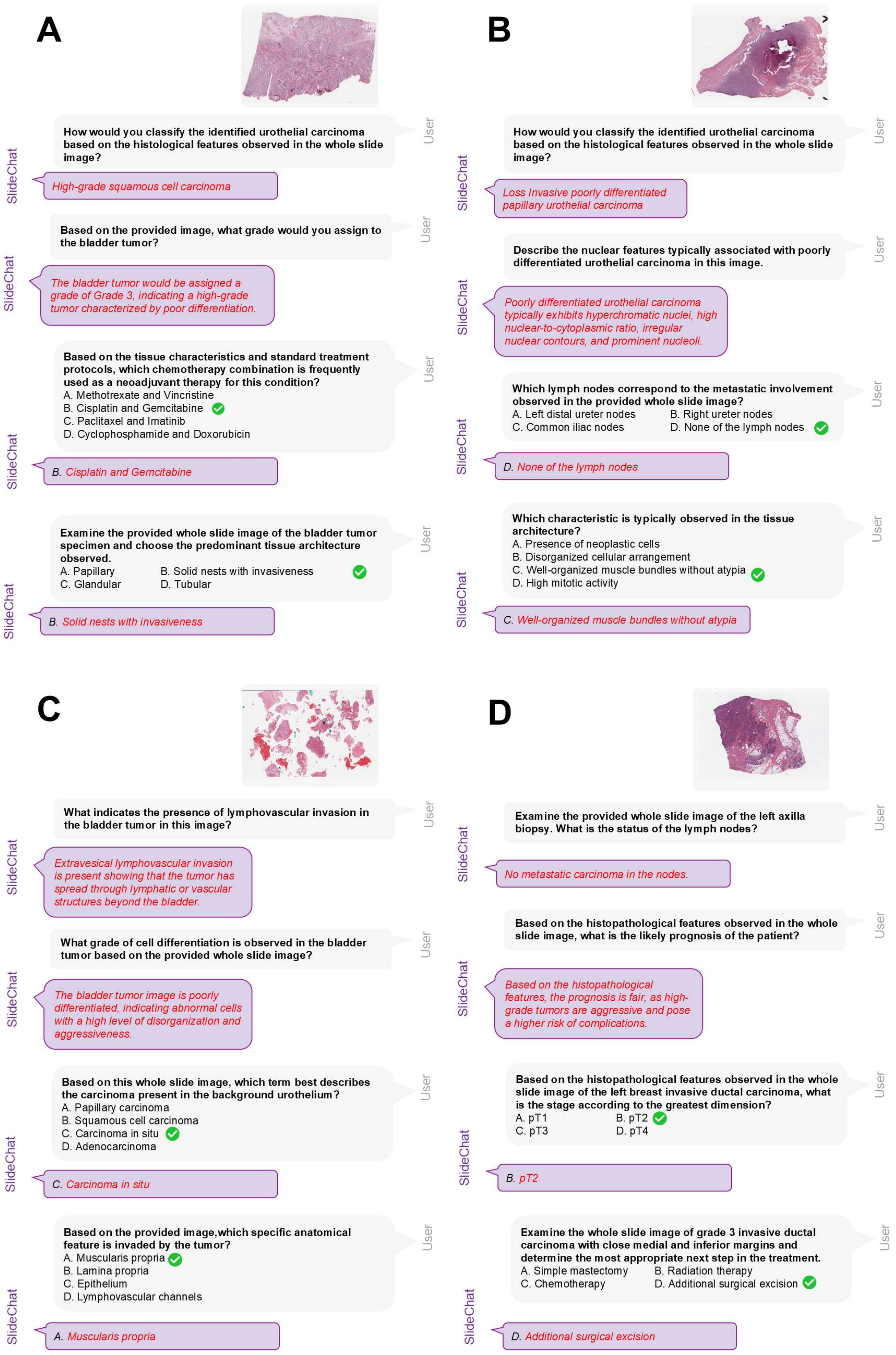
Deliver a compact explanation of the pivotal findings in pathology as revealed by the whole slide image.

The pathological findings indicate a diagnosis of moderately differentiated adenocarcinoma of the right upper lobe of the lung. The tumor is characterized by a mass with a diameter of 3.5 cm, located in the lung parenchyma, and is classified as pT1a. The bronchial margin is free of tumor involvement, and there is no evidence of lymphovascular invasion. The tumor is classified as pN0, indicating no regional lymph node metastasis. The resection margins are also free of tumor involvement. The pathological staging is determined to be pT1a, pN0, pMx, with a G2 grade, indicating a moderately differentiated tumor with no regional lymph node metastasis and an unknown distant metastasis status.

Describe succinctly the core discoveries in the pathology results from the whole slide image.

The pathological examination reveals a diagnosis of moderately differentiated squamous cell carcinoma located in the left lower lobe of the lung. The tumor is characterized by a central mass with a maximum dimension of 4.5 cm and is situated 2.5 cm from the pleural surface. Importantly, the surgical margins are free of tumor involvement, indicating complete resection. Additionally, there is no evidence of lymphatic or vascular invasion. Examination of the lymph nodes from the left lower lobe and the mediastinal region shows no metastatic involvement, with all sampled nodes being negative for tumor. The tumor is staged as pT2a, pN0, and pM0, reflecting a localized primary tumor with no regional lymph node metastasis and no distant metastasis.

Figure 7. Demonstration of SlideChat's Capability in Whole-Slide Image Captioning.



A

User: **How would you classify the identified urothelial carcinoma based on the histological features observed in the whole slide image?**

SlideChat: *High-grade squamous cell carcinoma*

User: **Based on the provided image, what grade would you assign to the bladder tumor?**

SlideChat: *The bladder tumor would be assigned a grade of Grade 3, indicating a high-grade tumor characterized by poor differentiation.*

User: **Based on the tissue characteristics and standard treatment protocols, which chemotherapy combination is frequently used as a neoadjuvant therapy for this condition?**
A. Methotrexate and Vincristine
B. Cisplatin and Gemcitabine ✓
C. Paclitaxel and Imatinib
D. Cyclophosphamide and Doxorubicin

SlideChat: *B. Cisplatin and Gemcitabine*

User: **Examine the provided whole slide image of the bladder tumor specimen and choose the predominant tissue architecture observed.**
A. Papillary
B. Solid nests with invasiveness ✓
C. Glandular
D. Tubular

SlideChat: *B. Solid nests with invasiveness*

B

User: **How would you classify the identified urothelial carcinoma based on the histological features observed in the whole slide image?**

SlideChat: *Loss Invasive poorly differentiated papillary urothelial carcinoma*

User: **Describe the nuclear features typically associated with poorly differentiated urothelial carcinoma in this image.**

SlideChat: *Poorly differentiated urothelial carcinoma typically exhibits hyperchromatic nuclei, high nuclear-to-cytoplasmic ratio, irregular nuclear contours, and prominent nucleoli.*

User: **Which lymph nodes correspond to the metastatic involvement observed in the provided whole slide image?**
A. Left distal ureter nodes
B. Right ureter nodes
C. Common iliac nodes
D. None of the lymph nodes ✓

SlideChat: *D. None of the lymph nodes*

User: **Which characteristic is typically observed in the tissue architecture?**
A. Presence of neoplastic cells
B. Disorganized cellular arrangement
C. Well-organized muscle bundles without atypia ✓
D. High mitotic activity

SlideChat: *C. Well-organized muscle bundles without atypia*

C

User: **What indicates the presence of lymphovascular invasion in the bladder tumor in this image?**

SlideChat: *Extravesical lymphovascular invasion is present showing that the tumor has spread through lymphatic or vascular structures beyond the bladder.*

User: **What grade of cell differentiation is observed in the bladder tumor based on the provided whole slide image?**

SlideChat: *The bladder tumor image is poorly differentiated, indicating abnormal cells with a high level of disorganization and aggressiveness.*

User: **Based on this whole slide image, which term best describes the carcinoma present in the background urothelium?**
A. Papillary carcinoma
B. Squamous cell carcinoma
C. Carcinoma in situ ✓
D. Adenocarcinoma

SlideChat: *C. Carcinoma in situ*

User: **Based on the provided image,which specific anatomical feature is invaded by the tumor?**
A. Muscularis propria ✓
B. Lamina propria
C. Epithelium
D. Lymphovascular channels

SlideChat: *A. Muscularis propria*

D

User: **Examine the provided whole slide image of the left axilla biopsy. What is the status of the lymph nodes?**

SlideChat: *No metastatic carcinoma in the nodes.*

User: **Based on the histopathological features observed in the whole slide image, what is the likely prognosis of the patient?**

SlideChat: *Based on the histopathological features, the prognosis is fair, as high-grade tumors are aggressive and pose a higher risk of complications.*

User: **Based on the histopathological features observed in the whole slide image of the left breast invasive ductal carcinoma, what is the stage according to the greatest dimension?**
A. pT1
B. pT2 ✓
C. pT3
D. pT4

SlideChat: *B. pT2*

User: **Examine the whole slide image of grade 3 invasive ductal carcinoma with close medial and inferior margins and determine the most appropriate next step in the treatment.**
A. Simple mastectomy
B. Radiation therapy
C. Chemotherapy
D. Additional surgical excision ✓

SlideChat: *D. Additional surgical excision*

Figure 8. Demonstration of our SlideChat for answering various questions based on the WSI.

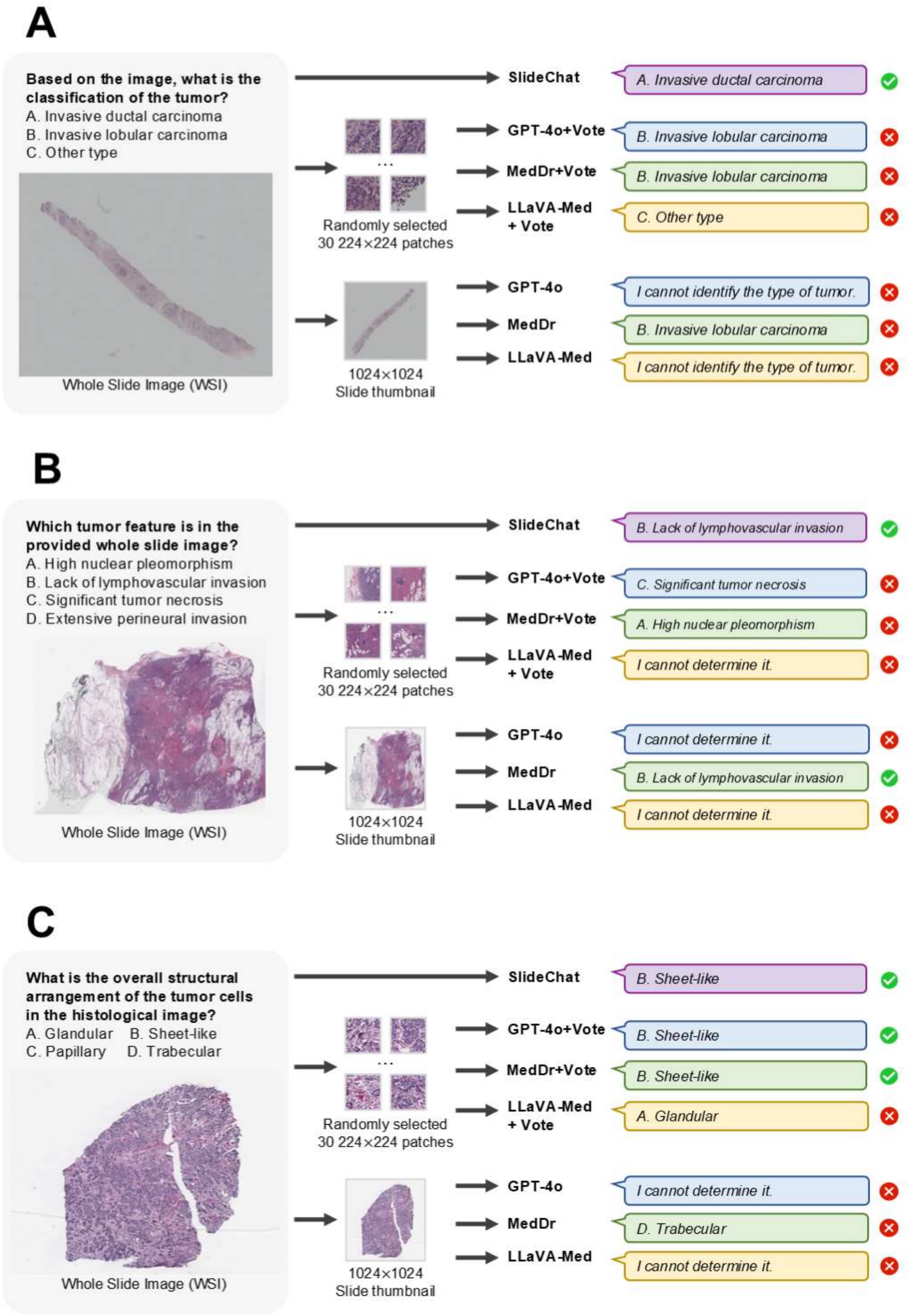

A
Based on the image, what is the classification of the tumor?
A. Invasive ductal carcinoma
B. Invasive lobular carcinoma
C. Other type
Whole Slide Image (WSI)
SlideChat
A. Invasive ductal carcinoma
Randomly selected 30 224×224 patches
GPT-4o+Vote
B. Invasive lobular carcinoma
MedDr+Vote
B. Invasive lobular carcinoma
LLaVA-Med + Vote
C. Other type
1024×1024 Slide thumbnail
GPT-4o
I cannot identify the type of tumor.
MedDr
B. Invasive lobular carcinoma
LLaVA-Med
I cannot identify the type of tumor.
B
Which tumor feature is in the provided whole slide image?
A. High nuclear pleomorphism
B. Lack of lymphovascular invasion
C. Significant tumor necrosis
D. Extensive perineural invasion
Whole Slide Image (WSI)
SlideChat
B. Lack of lymphovascular invasion
Randomly selected 30 224×224 patches
GPT-4o+Vote
C. Significant tumor necrosis
MedDr+Vote
A. High nuclear pleomorphism
LLaVA-Med + Vote
I cannot determine it.
1024×1024 Slide thumbnail
GPT-4o
I cannot determine it.
MedDr
B. Lack of lymphovascular invasion
LLaVA-Med
I cannot determine it.
C
What is the overall structural arrangement of the tumor cells in the histological image?
A. Glandular B. Sheet-like
C. Papillary D. Trabecular
Whole Slide Image (WSI)
SlideChat
B. Sheet-like
Randomly selected 30 224×224 patches
GPT-4o+Vote
B. Sheet-like
MedDr+Vote
B. Sheet-like
LLaVA-Med + Vote
A. Glandular
1024×1024 Slide thumbnail
GPT-4o
I cannot determine it.
MedDr
D. Trabecular
LLaVA-Med
I cannot determine it.

Figure 9. Comparing model outputs on SlideBench.