

# Automated Synoptic CT Reporting of IPMN Imaging Features: A Large Language Model Approach

---

Emir A. Syailendra, Felipe Lopez-Ramirez, Florent Tixier, Alejandra Blanco,  
Satomi Kawamoto, Linda C. Chu, Elliot K. Fishman



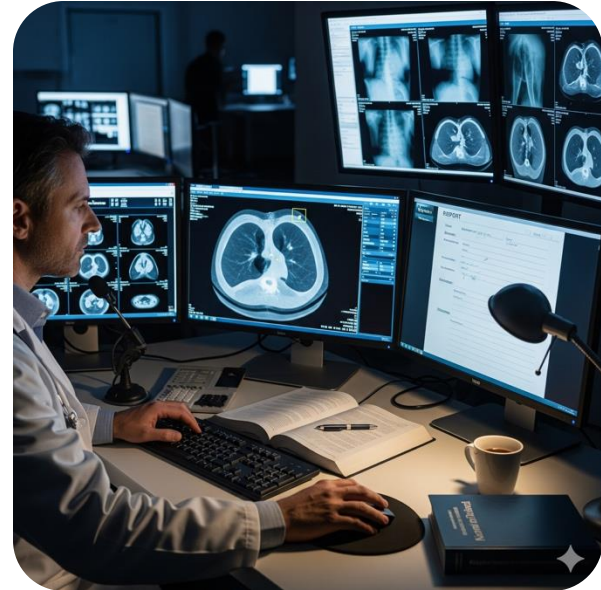
JOHNS HOPKINS  
M E D I C I N E

Presented by: Emir A. Syailendra, MD, MS  
Department of Radiology, Johns Hopkins University School of Medicine  
**SABI 2025 Annual Meeting** – October 12, 2025



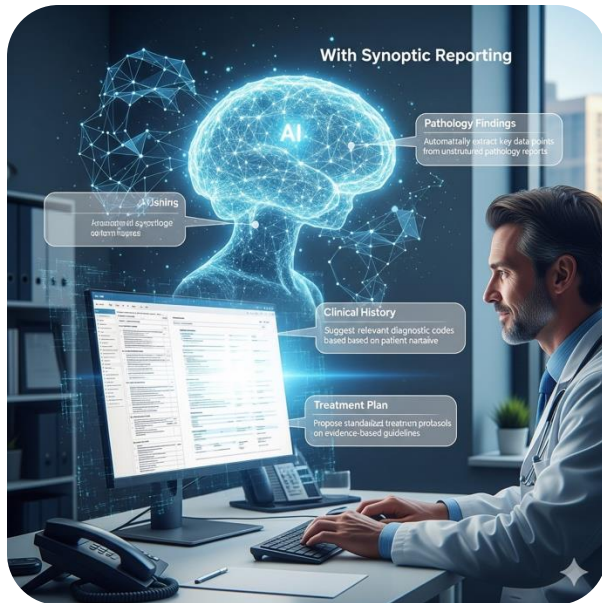
# Background

- Intraductal papillary mucinous neoplasms (IPMNs) are cystic pancreatic lesions with malignant potential  
→ *requiring careful preoperative imaging assessment*
- CT reports are the cornerstone of diagnosis and surgical planning and are typically written in free-text, narrative style.
- Narrative reports often lack standardization, completeness, and consistency, limiting their utility for:
  - Clinical decision-making
  - Multidisciplinary tumor boards
  - Research and database curation
- **Synoptic reporting**, which uses structured templates aligned with guidelines (e.g., Society of Abdominal Radiology, SAR), improves completeness and standardization  
→ *rarely adopted*



*\*generated by LLM*

# Background



*\*generated by LLM*

- Recent advances in large language models (LLMs) offer the ability to transform unstructured clinical text into structured outputs automatically.
- Applying LLMs to radiology reports could provide a scalable solution for generating synoptic reports without adding radiologist workload.

## Aims:

1. Evaluate whether a large language model (LLM) can convert free-text CT reports of intraductal papillary mucinous neoplasms (IPMNs) into structured synoptic templates  
→ Society of Abdominal Radiology (SAR) recommendations

# Methods

## Dataset

- 535 preoperative contrast-enhanced CT reports
- Random sample: 20 cases (pathology-confirmed IPMN) for manual evaluation

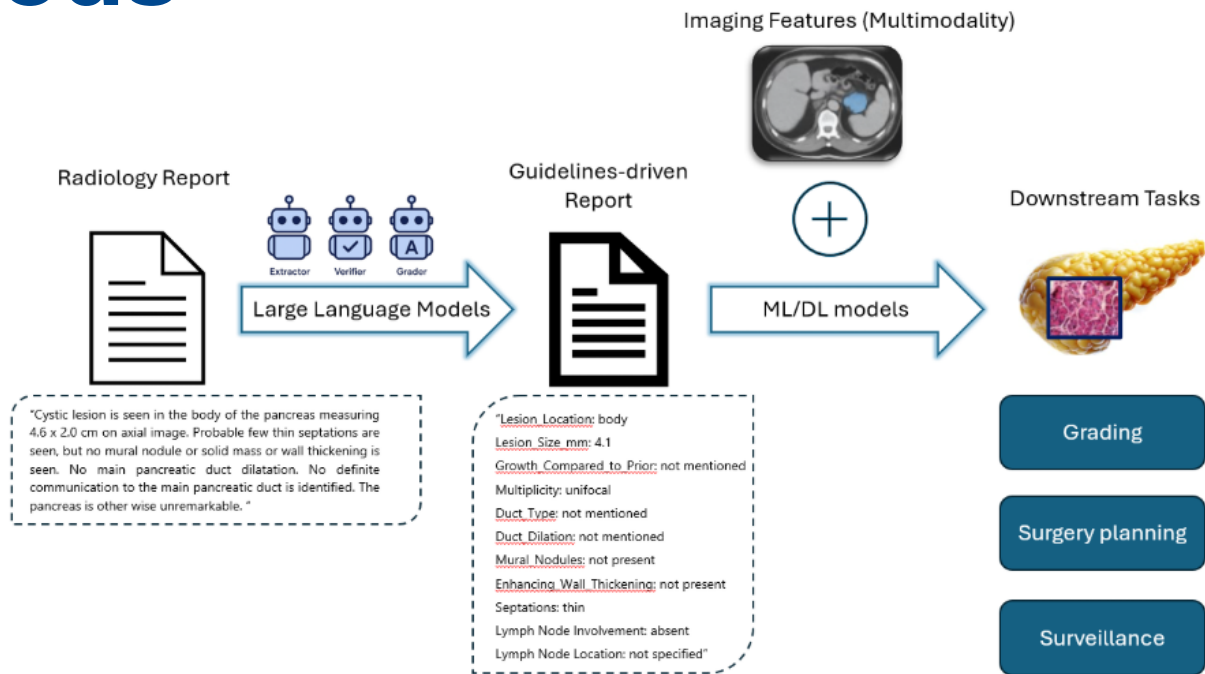
## Model & Framework

- Locally deployed LLaMA3 (domain-adapted)
- Multi-agent architecture: Extractor → Verifier → Grader
- Features: lesion location, size, growth, multiplicity, duct type, ductal dilation, mural nodules, wall thickening, septations, lymph node involvement (SAR guidelines<sup>1</sup>)

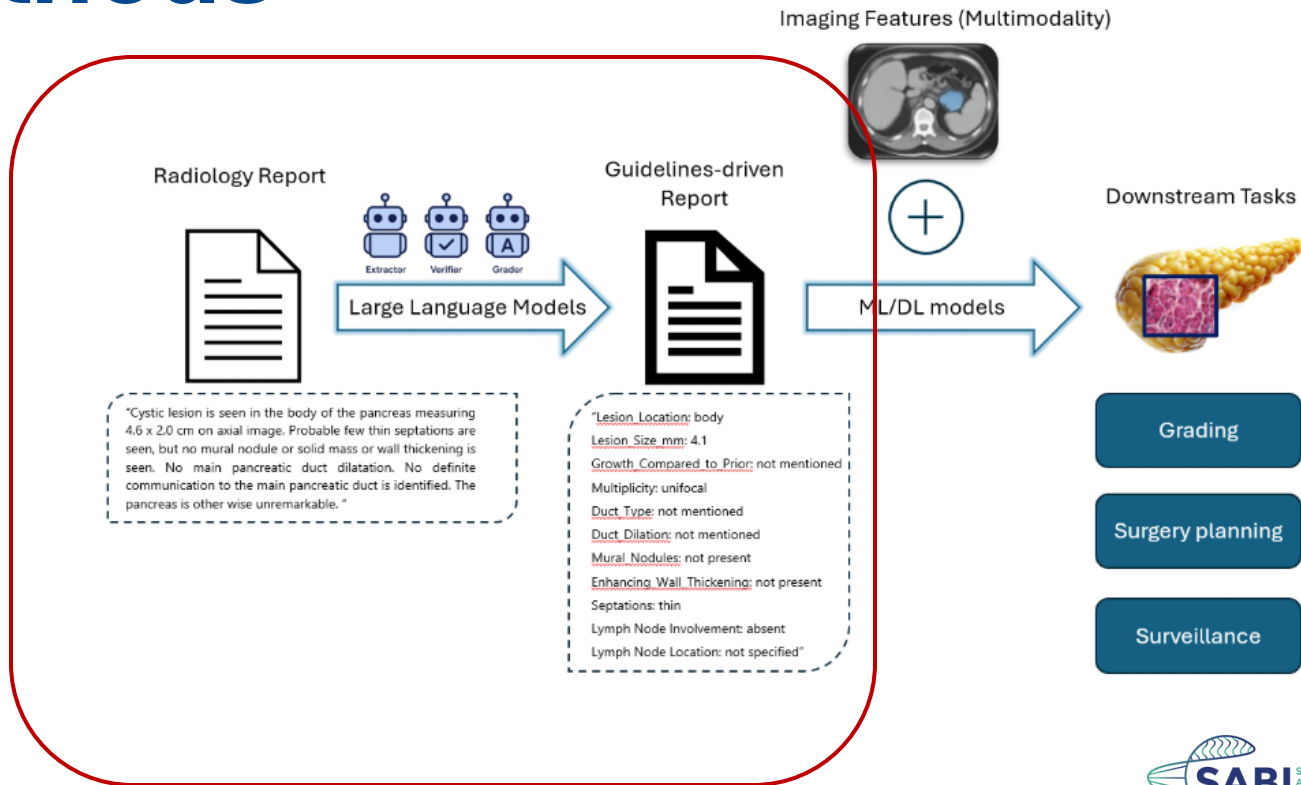
## Evaluation Metrics

- Completeness  
→ % of features filled when supported by report
- Accuracy  
→ correctness of extracted feature vs ground truth

# Methods



# Methods



# Results

Multiple cystic lesions distributed throughout the pancreatic parenchyma, without evidence of associated enhancement, as detailed below:

1. 1.4 x 0.9 cm cystic lesion within the anterior aspect of the pancreatic head (8/32).
2. 1.2 x 0.9 cm cystic lesion in the anterior aspect of the proximal pancreatic body (8/31). On previous exam this lesion and the lesion listed above appeared to connect and form a single lesion. Overall size and appearance has not substantially changed prior exam.
3. 1.0 x 0.8 cm cystic lesion in the posterior aspect of the pancreatic mid tail (8/37), unchanged in size and appearance since prior exam.
4. 0.8 x 0.6 cm cystic lesion in the anterior aspect of the distal pancreatic tail (8/42), also unchanged in size appeared since prior exam.
5. 0.8 x 0.8 cm cystic lesion within the posterior aspect of the uncinate process (8/49), also unchanged in size and appearance since prior exam.

These lesions likely represent side branch intraductal papillary mucinous neoplasms. On previous exam, unchanged in size and appearance since prior exam. Pancreatic parenchyma is otherwise normal without evidence of pancreatic ductal dilatation.

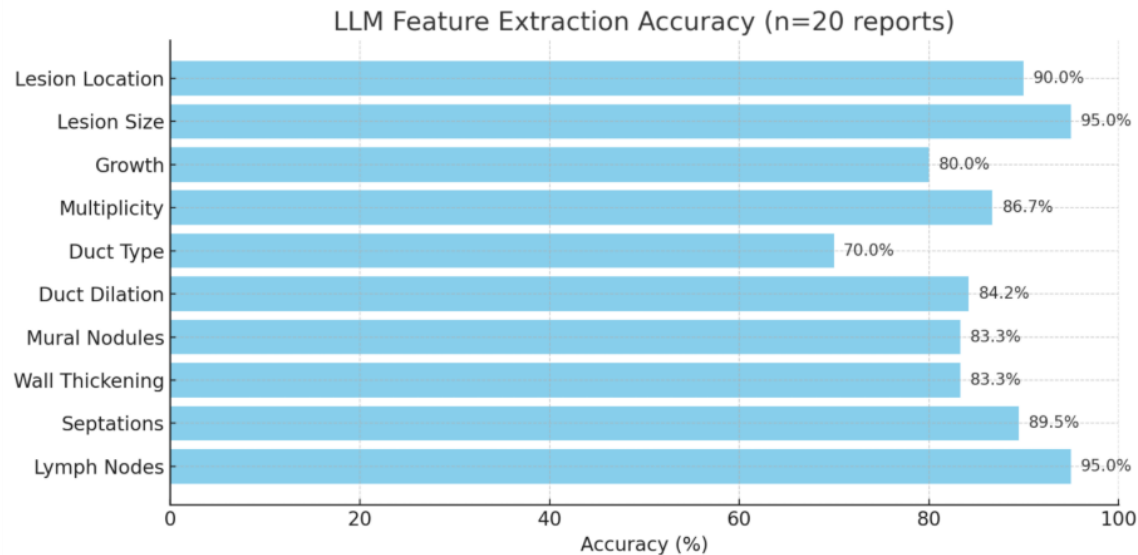
## IMPRESSION:

1. Multiple cystic lesions scattered throughout the pancreas, as detailed above, without evidence of associated enhancement, pancreatic atrophy, or pancreatic ductal dilatation. Findings most compatible with multiple side branch intraductal papillary mucinous neoplasms, unchanged in size and appearance since prior exam.



```
{
  "Lesion_Location": [
    "pancreatic head",
    "proximal pancreatic body",
    "pancreatic mid tail",
    "distal pancreatic tail",
    "uncinate process" ],
  "Lesion_Size_cm":
  [ 1.4 x 0.9 cm,
    1.2 x 0.9 cm,
    1.0 x 0.8 cm,
    0.8 x 0.6 cm,
    0.0 x 0.8 cm ],
  "Growth_Compared_to_Prior":
  [ "unchanged" ],
  "Multiplicity": "multiple",
  "Duct_Type": "intraductal papillary mucinous neoplasms",
  "Duct_Dilatation": "no evidence of pancreatic ductal dilatation",
  "Mural_Nodules": "no evidence",
  "Enhancing_Wall_Thickening": "no evidence",
  "Septations": "no evidence",
  "involvement_of_LNs": "no evidence", "
}
```

# Results



- All 20 cases successfully processed
- 97% completeness of fields filled
- 86% mean accuracy across 10 features

## High Performance Features

- Lesion size: 95% accuracy
- Lesion location: consistently identified (100%)
- Model retrieved up to 5 cysts per case with supporting sentences

# Error Analysis

## Error Types

- 62% = misinterpretation/misattribution
- 38% = hallucinations

**Duct type** (main duct/side duct/mixed duct) → lowest accuracy at **70%**

Feature	Error Example	Cause
Duct type	Inferred classification	Not explicitly stated on report
Mural nodule	Hallucination	Ambiguous wording

# Discussion

## Key Findings

- LLMs (domain-adapted LLaMA3) can reliably extract most guideline-driven IPMN features
- High completeness (97%) and accuracy (86%) → feasibility for real-world use.
- Strongest performance for lesion size and location, weaker for duct type due to inference from incomplete documentation.

## Significance

- Demonstrates that AI-assisted synoptic reporting is scalable, reducing the burden on radiologists.
- Structured reports improve data quality, multidisciplinary communication, and research readiness.
- LLM approach captures nuances like multiple cysts per case and retrieval of supporting text, adding interpretability.

# Discussion

## Limitations

- Small evaluation sample (20 reports) → may not fully reflect diverse report styles
- Errors from:
  - Ambiguous/underspecified language in original reports
  - Model hallucinations when key features were absent
- Requires human oversight to ensure clinical safety

## Future Directions

- Validate across larger, multi-institutional datasets
- Improve error handling (reduce misinterpretation & hallucination).
  - Explore integration into PACS/reporting systems with human-in-the-loop feedback
  - Assess downstream impact on clinical decisions, surgical planning, and outcomes

# Conclusion

- LLMs **can** be used to transform unstructured CT reports to structured templates for IPMN synoptic reporting
  - Scalable, standardized reporting of IPMNs
- Accuracy strong overall
  - Often limited by ambiguous or missing details
- **Human oversight still essential** in nuanced clinical interpretation

# Thank You!

Presented by: **Emir A. Syailendra, MD, MS**  
Department of Radiology, Johns Hopkins University School of Medicine

Contact:

 [msyaile1@jhu.edu](mailto:msyaile1@jhu.edu)  [thefelixlab.jhmi.edu](http://thefelixlab.jhmi.edu)



JOHNS HOPKINS  
MEDICINE

