Risk Prediction of Pancreatic Ductal Adenocarcinoma using AI Analysis of Abdominal CT Scans

Touseef Ahmad Qureshi, PhD

Assistant Professor | Research Scientist
Biomedical Imaging Research Institute,
Cedars-Sinai Medical Center Los Angeles, CA

Adjunct Assistant Professor
David Geffen School of Medicine,
University of California, Los Angeles, CA

Co-Authors:
Sehrish Javed, MS, Srinivas Gaddam, MD, Stephen Jacob Pandol, MD, and Debiao Li, PhD
Biomedical Imaging Research Institute, Cedars-Sinai Medical Center, LA, CA
Pancreatic Cancer

• The Pancreas has two main jobs
  • Make digestive enzymes that help the intestines break down food
  • Regulate body’s use of Sugars and starches

• Pancreatic Cancer, also known at the “Silent Killer”, has been diagnosed in many prominent figures
Pancreatic Ductal Adenocarcinoma (PDAC)

- Accounts for 3% of all cancers and is 4th leading cause of all cancer deaths
- Expectedly becoming the 2nd most by 2030.
- ~60,430 new cases and 48,220 deaths are expected this year.
- ~80% of diagnosis occurs at late stage of cancer.
- The 5-year survival rate is 11% but can be as high as 50% with early-stage diagnosis.
- Early-stage diagnosis of PDAC is challenging:
  - Lack of specific symptoms (e.g., abdominal pain)
- Prediction (risk stratification) can assist improving early diagnosis!

- Key Statistics for Pancreatic Cancer, American Cancer Society,
5-Year Relative Survival Rates %

- Breast Cancer
- Colon Cancer
- Pancreatic cancer

Connor O’Malley
Can Early Detection Improve Pancreatic Cancer?

The graph shows the probability of overall survival over months from surgery for different T stages:
- T1 (n = 120)
- T2 (n = 334)
- T3 (n = 68)

- < 2 cm
- 2 – 4 cm
- > 4 cm

The p-value is less than 0.001.
Rationale, Objective, and Hypotheses

• **Motivation:**
  - Seven million ER visits per year due to abdominal reasons in the USA, where abdominal Computed Tomography (CT) scan is usually performed.
  - Pre-diagnostic CT scans may provide critical morphological information associated with pre-cancer or early cancer biological changes to predict PDAC risk.

• **Objective:**
  - To develop an artificial intelligence (AI) model to predict PDAC risks in 3 years using a combination of pre-diagnostic CT image features and non-imaging factors.

• **Hypotheses**
  - AI allows extraction of unique image features in pre-diagnostic CT images associated with pre-cancer or early cancer biological changes that are invisible to naked eyes.
  - The combination of pre-diagnostic image features and non-imaging factors improves the accuracy of PDAC risk stratification and prediction over that using conventional non-imaging factors alone.
Data Design

- **Diagnostic:**
  - Contrast-enhanced abdominal CT scan
  - Histopathologically established PDAC (visible tumor)

- **Pre-diagnostic:**
  - Contrast-enhanced abdominal CT scan
  - same patient as in the diagnostic scan,
  - acquired **up to 3 years** prior to the PDAC diagnosis

- **Healthy control:**
  - Contrast-enhanced abdominal CT scan
  - Different subject with pancreas is declared healthy on the imaging and non-imaging clinical reports.
  - Non-gastrointestinal disorders or accidents.
  - This group didn’t develop PDAC in the following 3 years.
  - Gender, age, and CT scan time are matched with pre-diagnostic imaging.
Radiomic Analysis

108 CT scans (36 from Healthy, Pre-Diagnostic, and Diagnostic)
Pilot analysis: Healthy vs Pre-diagnostic scans
Healthy Control vs Pre-Diagnostic

Healthy

Pre-diagnostic
Identification of PDAC predictors

| Intensity | Variations in tissue **intensity** of the whole pancreas and pancreatic subregions. For example, the affected region starts to become darker during the development of PDAC.  
Pancreatic calcifications may appear within the tissue or duct during PDAC development, resulting in variations in subregional **intensity**. |
|-----------|--------------------------------------------------------------------------------------------------|
| Texture   | Changes in the cellular mechanism and vasculature cause **textural** changes, increasing tissue heterogeneity.  
Ductal complications, such as distal parenchymal atrophy, intraductal papillary mucinous neoplasms of the pancreas (IPMNs), and intraductal calculi (pancreatolithiasis) are often associated with PDAC. Such complications are demonstrated as **textural** changes in PD. |
| Size      | Pancreatic inflammation is often associated with PDAC, which may cause changes in the **size** of the pancreas.  
Cancerous cells lining the PD distract the normal behavior of the duct during PDAC development. This can be observed as dilatation or increased **size** of the duct. |
| Shape     | PD tends to turn tortuous during PDAC development (deformed **shape**). |

Preliminary studies and results

- Bayesian Classification model
- Classification accuracy **85%** in predicting that a patient will develop PDAC within 3 years.
- Confusion matrix for classification of 28 CT scans of the external set consisting of 14 from each of Healthy control and Pre-diagnostic group. Numbers in the green blocks show true positives.

<table>
<thead>
<tr>
<th></th>
<th>True Healthy</th>
<th>True Pre-diagnostic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Predicted Healthy</td>
<td>13</td>
<td>3</td>
</tr>
<tr>
<td>Predicted Pre-diagnostic</td>
<td>1</td>
<td>11</td>
</tr>
</tbody>
</table>
Progress

- **Patent** - LI/PANDOL/QURESHI/WU/GADDAM [PREDICTION OF PANCREATIC DUCTAL ADENOCARCINOMA USING COMPUTED TOMOGRAPHY IMAGES OF PANCREAS]; CSMC Ref. li001287; Our Ref.: 065472-000797WO00


- **NIH Award**: NIH R01 CA260955, 09/01/21 - 08/31/26 “Predicting Pancreatic Ductal Adenocarcinoma (PDAC) Through Artificial Intelligence Analysis of Pre-Diagnostic CT Images”
  - 8 centers will collaborate
  - Large dataset (>3000 CT scans)
  - Extensive validation of the prediction model
PDAC Risk prediction model

• **Aim 1 (Technical Development):**
  To develop automated segmentation techniques for the pancreas, pancreatic subregions, and pancreatic duct (PD) in abdominal CT scans.

• **Aim 2 (Analysis):**
  To derive pre-diagnostic CT image features that are predictive of PDAC.

• **Aim 3 (Prediction modelling):**
  To develop and evaluate the PDAC prediction model using image features and non-imaging factors.
Target Enrollment

• **Target enrollment**: <2000 subjects (~3000 CT scans)

• **8 Centers will collaborate**
  1. Cedars-Sinai Medical Center (CSMC)
  2. Greater LA Veterans Affairs Healthcare System (GLA VA)
  3. University of Michigan (UM)
  4. Massachusetts General Hospital (MGH)
  5. Northwestern University (NU)
  6. University of Southern California (USC)
  7. University of California, Irvine (UCI)
  8. Rutgers University (RU)
Aim 1: Data harmonization

Target enrollment: > 2,000 subjects, 3000 scans
Challenge: variability in voxel size, contrast, etc.
1. Linear transformation
2. Generative adversarial network
3. Image discretization
4. ComBat correction

Aim 1: Data labelling application

- **Key features of the application:**
  - **Interactive predefined procedure to outline subregions**
    - Baseline segmentation are defined already (using regional ratios)
    - None-to-slight change might be required
  - **Relatively simpler than commercial applications**
    - Color schemes predefined
    - File formatting predefined
    - Less complex interface
  - **Reduced time and workload**
    - Saves ~60-70% of subregion labelling efforts and time.
Aim 1: Segmentation of Pancreas

- Preliminary segmentation results
- NIH public data (n=83)
- Compared with manual labelling as ground truth
- Sørensen–Dice coefficient (DSC) score > 89%
- Manuscript under review by Journal of Medical Imaging
Aim 1: Segmentation of pancreatic subregions

## Significance of subregional analysis

### Discrepancies among pancreatic subregions

<table>
<thead>
<tr>
<th>Tumor Structure</th>
<th>Head</th>
<th>Body</th>
<th>Tail</th>
<th>Application</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Histology:</strong></td>
<td>Non-Squamous tumors</td>
<td>Squamous tumors</td>
<td>Less Aggressive</td>
<td>Variable tumor biology among sub regions leads to tumor heterogeneity across subregions</td>
</tr>
<tr>
<td><strong>Genetics:</strong></td>
<td>Less Aggressive</td>
<td>More Aggressive, High Grade, Poorly or undifferentiated</td>
<td>Low Grade, Well differentiated</td>
<td></td>
</tr>
<tr>
<td>Symptoms</td>
<td>Jaundice, Dark urine, Light stool, and Weight loss</td>
<td>Back and upper abdominal pain</td>
<td>Lower abdominal pain</td>
<td>Helps correlating unique clinical factors to specific subregions</td>
</tr>
<tr>
<td>Drug Response</td>
<td>More sensitive to Gemcitabine-based regimen</td>
<td>More sensitive to Fluorouracil-based regimen</td>
<td>Better treatment plan and prediction of treatment response</td>
<td></td>
</tr>
<tr>
<td>Metastasis</td>
<td>Low (42%)</td>
<td>Moderate (68%)</td>
<td>Extreme (84%)</td>
<td>Helps identifying high risk organs</td>
</tr>
<tr>
<td>Incidence Rate</td>
<td>High (71%)</td>
<td>Low (13%)</td>
<td>Low (16%)</td>
<td></td>
</tr>
<tr>
<td>Survival Rate</td>
<td>44%</td>
<td>27%</td>
<td>27%</td>
<td></td>
</tr>
<tr>
<td>Resection Rate</td>
<td>17%</td>
<td>4%</td>
<td>7%</td>
<td></td>
</tr>
</tbody>
</table>

Segmentation of Pancreatic subregions

- Developed Anatomy-guided deep learning model
  - Incorporating Structural constraints
    - Width ratio, shape analysis
    - Convolution Neural Network
- 2D segmentation
  - NIH 82 CT images
  - 82% Dice score
- 3D segmentation
  - In process
Aim 1: Federated Learning

Privacy preserved decentralized training without data sharing!

- EDRN, MD Anderson [Collaboration]
- Address critical issues:
  - Data privacy
  - Data security
  - Data access rights
- Eliminate data transfer
  - Only model updates are shared
  - Different ways to updates
- Utilize large amount of unused data from centers

Aim 1: Federated Learning pancreas segmentation

- Three datasets
  - Medical Segmentation Decathlon
  - NIH 82 Pancreas CT
  - Beyond the Cranial Vault (BTCV) Abdomen dataset
- Existing methods only average updates based on training size
- Our goal: Automatically adapted aggregation scheme based on history of loss and aggregation weight
- Better and stable performance across multiple datasets

<table>
<thead>
<tr>
<th>Training method</th>
<th>Test dataset</th>
<th>Case-wise Avg (130 cases)</th>
<th>Site-wise Avg (3 sites)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Local-Decathlon</td>
<td>76.80%</td>
<td>76.11%</td>
</tr>
<tr>
<td></td>
<td>Local-NIH-82</td>
<td>64.04%</td>
<td>66.39%</td>
</tr>
<tr>
<td></td>
<td>Local-BTCV</td>
<td>31.40%</td>
<td>38.30%</td>
</tr>
<tr>
<td></td>
<td>DWA</td>
<td>76.60%</td>
<td>77.35%</td>
</tr>
<tr>
<td></td>
<td>FedAvg</td>
<td>77.32%</td>
<td>75.94%</td>
</tr>
<tr>
<td>RNN-Aggr(ours)</td>
<td></td>
<td><strong>78.74%</strong></td>
<td><strong>78.90%</strong></td>
</tr>
</tbody>
</table>
Risk prediction based on subregional analysis

Aim 2: Identification of PDAC predictors

<table>
<thead>
<tr>
<th>Category</th>
<th>Factors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Demographic</td>
<td>Age at imaging, sex, race/ethnicity</td>
</tr>
<tr>
<td>Epidemiologic risk factors</td>
<td>Smoking history</td>
</tr>
<tr>
<td>Anthropometry</td>
<td>Weight, weight change, height, BMI</td>
</tr>
<tr>
<td>Clinical comorbidities</td>
<td>pancreatitis, liver disease, alcoholism, chronic obstructive pulmonary disease</td>
</tr>
<tr>
<td>Laboratory tests</td>
<td>Creatinine, hemoglobin A1c, cholesterol, bilirubin</td>
</tr>
</tbody>
</table>

- **Intensity**: Variations in tissue intensity of the whole pancreas and pancreatic subregions. For example, the affected region starts to become darker during the development of PDAC. Pancreatic calcifications may appear within the tissue or duct during PDAC development, resulting in variations in subregional intensity.

- **Texture**: Changes in the cellular mechanism and vasculature cause textural changes, increasing tissue heterogeneity. Ductal complications, such as distal parenchymal atrophy, intraductal papillary mucinous neoplasms of the pancreas (IPMNs), and intraductal calculi (pancreatolithiasis) are often associated with PDAC. Such complications are demonstrated as textural changes in PD.

- **Size**: Pancreatic inflammation is often associated with PDAC, which may cause changes in the size of the pancreas. Cancerous cells lining the PD distract the normal behavior of the duct during PDAC development. This can be observed as dilatation or increased size of the duct.

- **Shape**: PD tends to turn tortuous during PDAC development (deformed shape).

Significance test
Information gain ranking
Trend identification
Aim 3: Prediction model development

Integrated machine learning for risk prediction
Combining radiomics, Deep learning and clinical features

End-to-End risk prediction model
Integrating deep learning and clinical features

Training
Validation
Progress summary

• **Collaboration with partners**
  - IRBs approved from most centers, Data agreements documents are prepared
  - Schemes for data mining and data collection are being designed
  - 500 cases are obtained, analysis is undergoing.

• **Technical Development**
  - Labelling application
  - Pancreas segmentation model
  - Pancreatic subregional segmentation model
  - Federated learning framework
  - Distributed system for efficient processing

• **Publications**
Patients come into the ER with abdominal pain. Perform CT scan and obtain clinical indicators. Input scan to our model. If high risk, follow up patient and monitor closely to detect cancer early. If normal, surgical resection and medical treatment at the earliest stage to save patients’ lives.
Major Novelties of This Project

- Data structure – pre-diagnostic time point never explored before due to data pool limitation
- Product software will significantly improve prognosis and decrease treatment cost
- Prediction of PDAC will save patients’ lives and early detection will allow for effective medical intervention with ability to cure the disease
- AI massive population screening tool will allow for large data collection which will enable researchers to explore many questions about this disease not previously possible due to data limitation
Future direction and other projects

• Including liver organ as part of analysis for early changes.
• Replicating model on liver cancer, bladder cancer, etc.
• Predicting PDAC treatment response.
One of the most comprehensive imaging research center in the world

Biomedical Imaging Research Institute at Cedars-Sinai Medical Center

Areas of Research
- Pancreatic Cancer Imaging
- Breast Cancer Imaging
- Cardiovascular Imaging
- Quantitative Image Analysis
- Neuroimaging
- Oncologic Radiation Therapy Imaging
- Abdominal Computerized Tomography

25 Faculty Members
40 Research Scientists
20 Supporting Staff and Management