

1 **PanCanAID – Pancreas Cancer Artificial Intelligence Driven** 2 **Diagnosis in CT Scan Imaging: A Protocol for a Multicentric** 3 **Ambispective Diagnostic Study**

4 **Running Title:** Protocol of PanCanAID – Pancreas Cancer AI-Driven Diagnosis

5 Seyed Amir Ahmad Safavi-Naini^{1,2}, Armin Behnamnia², Faezeh Khorasanizadeh³, Ali Soroush⁴,
6 Farhad Zamani⁵, Faeze Salahshour³, Amir Sadeghi¹, Seyedmahdi Mirtajaddini⁶, Ashkan Zandi⁷,
7 Fatemeh Shojaeian⁸, Maryam Saeedi², Azade Ehasni¹, Abdolhamid Chavoshi Khamneh⁴, Zhaleh
8 Mohsenifar⁹, Farid Azmoudeh Ardalan¹⁰, Kavous Firouznia³, Shabnam Shahrokh¹, Masoomeh
9 Raoufi¹¹, Pooneh Dehghan¹², Pardis Ketabi Moghadam¹, Alireza Mansour-Ghanaei¹³, Parinaz
10 Mellatdoust¹⁴, Habib Malekpour¹⁵, Alireza Rasekhi¹⁶, Fariborz Mansour-Ghanaei¹³, Masoudreza
11 Sohrabi^{5*}, Fariba Zarei^{16*}, Amir Reza Radmard¹⁷, Hossein Ghanaati^{3*}, Hamid Assadzadeh
12 Aghdaei^{1*}, Mohammad Reza Zali¹, Hamid R. Rabiee^{2*}

13 1- Gastroenterology and Liver Diseases Research Center, Research Institute for
14 Gastroenterology and Liver Diseases, Shahid Beheshti University of Medical Sciences, Tehran,
15 Iran

16 2- Data Science and Machine Learning (DML) Lab, Department of Computer Engineering,
17 Sharif University of Technology, Tehran, Iran

18 3- Advanced Diagnostic and Interventional Radiology Research Center (ADIR), Imam Khomeini
19 Hospital, Tehran University of Medical Science, Tehran, Iran

20 4- Division of Data Driven and Digital Health (D3M), The Charles Bronfman Institute for
21 Personalized Medicine, Icahn School of Medicine at Mount Sinai, New York, NY, USA

22 5- Gastrointestinal and Liver Diseases Research Center, Iran University of Medical Sciences,
23 Tehran, Iran

24 6- Valiasr International Hospital, Tabriz, Iran

25 7- School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta,
26 GA, 30332, USA

27 8- Sidney Kimmel Comprehensive Cancer Research Center, Johns Hopkins School of Medicine,
28 Baltimore, USA

29 9- Department of Pathology, Ayatollah Taleghani Educational Hospital, Faculty of Medicine,
30 Shahid Beheshti University of Medical Sciences, Tehran, Iran

31 10- Pathology and Laboratory Medicine Department, Imam Khomeini Hospital Complex, Tehran
32 University of Medical Sciences, Tehran, Iran

33 11- Department of Radiology, Imam Hossein Hospital, School of Medicine, Shahid Beheshti
34 University of Medical Sciences, Tehran, Iran

- 35 12- Imaging Department, Taleghani Hospital; Shahid Behesti University of Medical Sciences,
36 Tehran, Iran
- 37 13- Gastrointestinal and Liver Diseases Research Center, Guilan University of Medical Sciences,
38 Rasht, Iran
- 39 14- Computer Science and Engineering, Dipartimento di Elettronica Informazione e
40 Bioingegneria, Politecnico di Milano, Milan, Italy
- 41 15- Department of Adult Gastroenterology and Hepatology, Imam Hossein Hospital, Shahid
42 Beheshti University of Medical Sciences, Tehran, Iran
- 43 16- Medical Imaging Research Center, Shiraz University of Medical Sciences, Shiraz, Iran
- 44 17- Department of Radiology, Shariati Hospital, Tehran University of Medical Sciences, Tehran,
45 Iran
- 46 * Corresponding to rabiee@sharif.edu (HRR) & hamid.assadzadeh@gmail.com (HAA) &
47 ghanaati@yahoo.com (HG) & zareifari@sums.ac.ir (FZ) & sohrab_r@yahoo.com (MS) &
48 fmansourghanai@gmail.com (FMG)

49

50 **Abstract**

51 **Introduction:** Pancreatic cancer is thought to have an extremely dismal prognosis. Most cancer-
52 related deaths occur from metastasis rather than the primary tumor, although individuals with
53 tumors smaller than 1 cm in diameter have more than 80% 5-year survival. Thus, the current
54 protocol introduces PanCanAID project which intends to develop several computer-aided-
55 diagnosis (CAD) systems to enhance pancreatic cancer diagnosis and management using CT
56 scan imaging.

57 **Methods and analysis:** Patients with pathologically confirmed pancreatic ductal
58 adenocarcinoma (PDAC) or pancreatic neuroendocrine tumor (PNET) will be included as
59 pancreatic cancer cases. The controls will be patients without CT evidence of abdominal
60 malignancy. A data bank of contrast-enhanced abdominopelvic CT scans, survival data, and
61 demographics will be collected from ten medical centers in four provinces. Endosonography
62 images and clinical data, if available, will be added to the data bank. Annotation and manual
63 segmentation will be handled by radiologists and confirmed by a second expert radiologist in
64 abdominal imaging. PanCanAID intelligent system is designed to (1) detect abdominopelvic CT
65 scan phase, (2) segment pancreas organ, (3) diagnose pancreatic cancer and its subtype in arterial
66 phase CT scan, (4) diagnose pancreatic cancer and its subtype in non-contrast CT scan, (5) carry
67 out prognosis (TNM stage and survival) based on arterial phase CT scan, (6) and estimate tumor
68 resectability. A domain adaptation step will be handled to use online data and provide pancreas
69 organ segmentation to reduce the segmentation time. After data collection, a state-of-the-art deep
70 learning algorithm will be developed for each task and benchmarked against rival models.

71 **Conclusion:** PanCanAID is a large-scale, multidisciplinary AI project to assist clinicians in
72 diagnosing and managing pancreas cancer. Here, we present the PanCanAID protocol to assure
73 the quality and replicability of our models. In our experience, the effort to prepare a detailed
74 protocol facilitates a positive interdisciplinary culture and the preemptive identification of errors
75 before they occur.

76 **Keyword:** Machine Learning; Pancreatic Neoplasms; Tomography, X-Ray Computed;
77 Endosonography; Artificial Intelligence

78 Introduction

79 Among all cancer types, pancreatic cancer has an especially dismal prognosis (1). At the time of
80 presentation, only 11% of patients are at an early enough stage to qualify for curative surgery (1-
81 3). In particular, individuals with tumors smaller than 1 cm in diameter showed a relatively
82 favorable average long-term survival rate of 80.4% at five years (4). Therefore, effective early
83 detection of pancreatic cancer is critical for increasing the proportion of individuals who can
84 qualify for treatments that reduce mortality (1). Current methods of pancreatic cancer detection
85 include abdominal computed tomography (CT), endoscopic ultrasonography (EUS), endoscopic
86 retrograde cholangiopancreatography (ERCP), and magnetic resonance imaging (MRI) (5).

87 **Fig 1. a) Pancreatic cancer can metastasize to several secondary sites, including the liver,**
88 **lung, peritoneum, bone, and brain. The most common secondary site is the liver, which is**
89 **affected in more than half of the cases of metastatic pancreatic cancer. The lung is the**
90 **second most common site, followed by the peritoneum (6). b) Cancer progression in the**
91 **pancreas. The tumor is initially confined to the pancreas in stage #1, but it spreads beyond**
92 **the pancreas to involve nearby lymph nodes in stage #2. By stage #3, the tumor has invaded**
93 **either the celiac axis or the mesenteric artery. In stage #4, cancer involves other organs**
94 **outside the pancreas (7). ci) Pancreatic cancer often results from a sequence of genetic**
95 **mutations that transform normal pancreatic mucosa into an invasive malignancy through**
96 **precursor lesions. The three most widely recognized precursor lesions are Pancreatic**
97 **Intraepithelial Neoplasia (PanIN), Intraductal Papillary Mucinous Neoplasm (IPMN), and**
98 **Mucinous Cystic Neoplasm (MCN). PanIN is the most common precursor lesion for**
99 **pancreatic ductal adenocarcinoma (PDAC), and genetic abnormalities found in these**
100 **lesions are common in adjacent PDAC. The histological pattern of PanIN progression also**
101 **reflects the accumulation of mutations in cancerous tissue. KRAS mutations and shortened**
102 **telomeres characterize low-grade PanIN lesions. At the same time, high-grade PanIN and**
103 **PDAC tissues display mutations in p16, p53, CDNK27, and SMAD4, along with a higher**
104 **frequency of KRAS mutation (8). cii) Image with a green border shows normal acini and**
105 **normal ducts. The base of acinar cells turns blue due to the abundance of RNA and nuclei,**
106 **while the cells' apex (or luminal aspect) appears pink due to the high presence of zymogen**
107 **proteins that function as digestive enzymes. Intralobular ductule in cross section is obvious.**
108 **The ductule's lumen contains a granular proteinaceous precipitate that appears pink due**
109 **to pancreatic juice. The nuclei in images with orange borders are enlarged,**
110 **hyperchromatic, and show moderate to severe nuclear atypia, with prominent nucleoli**
111 **representing PanIN. The cytoplasm may be abundant and mucin filled. PanIN, a pre-**
112 **neoplastic lesion of the pancreas, is classified into low-grade and high-grade based on the**
113 **degree of dysplasia. The epithelial cells display severe nuclear atypia and anaplasia, with**

114 **loss of polarity and increased mitotic activity. Image with a red border, PDAC is the most**
115 **common form of pancreatic cancer. On H&E staining, PDAC lesions typically exhibit the**
116 **following features: desmoplastic reaction, hyperchromatic nuclei with irregular contours**
117 **and clumped chromatin tumor cells, mitotic figures, and invasion.**

118
119 CT scans and EUS are the commonly used imaging examinations for pancreas cancer (9, 10).
120 EUS offers an excellent spatial resolution of the pancreas, and CT scans give information about
121 the tumor and its relationship to surrounding structures. However, EUS is an invasive procedure,
122 and its performance depends on the endoscopist's skill (11, 12). Radiologists also require a
123 considerable amount of training to identify early-stage pancreatic tumors. In a retrospective
124 analysis of pancreatic cancer cases, tumors were detectable in CT images three years before
125 clinical diagnosis (13). Even with expert radiologists, exhaustion and negligence can additionally
126 lead to missed diagnoses (14). These findings suggest an improved review of CT scan exams
127 could increase the proportion of pancreatic cancers diagnosed early.

128 Contrast-enhanced CT scan (CECT) is the preferred technique for pancreas imaging since it
129 characterizes the tumor and surrounding tissue. After the injection of intravenous contrast (IV),
130 the operator takes sequential CT scans at 45 seconds (late arterial phase) and 60 seconds
131 (portovenous phase). Some centers obtain more detailed triple-phase "pancreas protocol"
132 imaging (arterial, venous, and portal) to improve visualization of the tumor and characterization
133 of invasion (15). The effective interpretation of multiple collections of images in different phases
134 requires significant experience and attention.

135 Optimism has steadily grown over the potential of computer-assisted radiology techniques to
136 facilitate early diagnosis and timely management (16, 17). Machine learning (ML) models can
137 explicitly explore hidden patterns in the data and have produced groundbreaking results in
138 almost all fields of medical imaging (18). Expert radiologists often outperform ML models.

139 However, the cost and limited availability of such expertise impair the early detection of
140 pancreatic cancer (18). ML-driven computerized clinical decision support systems (CDSS) can
141 help less experienced clinicians decrease time-to-diagnosis, increase accuracy, reduce
142 interobserver variability, promote equitable healthcare access, and enhance cost-effectiveness
143 (14, 18).

144 A sizeable multicenter image dataset and interdisciplinary framework are required to develop a
145 generalizable and practical CDSS. Both requirements are especially challenging for pancreatic
146 cancer as the disease is rare, and its management is a winding journey involving many points of
147 care (17, 19). The challenge of obtaining data could explain why few high-performance but data-
148 demanding deep learning models have been published for pancreatic cancer (20). As of 2023, 2
149 out of 13 studies of ML-based pancreatic cancer detection on CT scan imaging had more than
150 300 participants (21), some of which have been summarized in **Table 1**. Although artificial
151 intelligence (AI)-assisted pancreatic cancer detection is rapidly growing, studies have been
152 hampered by small sample sizes and lack of external validation. Besides, new approaches such as
153 Segment Anything Model (SAM) may improve model performance and facilitate data labeling
154 (20-22).

155 **Table 1. Literature of review for diagnostic and prognostic machine learning algorithms**
156 **for pancreas CT scan imaging (22).** Footnote: PDAC: Pancreatic ductal adenocarcinoma; RF:
157 random forest; CNN: convolutional neural network; ML: Machine learning; Acc: Accuracy;
158 AUC: Area under the receiver operating characteristic curve; Internal validation (In); External
159 validation (Ex).

Author, Year	Aim	Algorithm	Data Set	Evaluation
Si et al., 2021 (23)	Pancreas cancer detection and segmentation and subtype detection	Deep CNN and federated learning	319 cases	Ex (AUC): 0.87; Ex (dice score): 83%

Liu et al., 2020 (24)	Pancreas cancer detection	CNN	370 cases and 320 controls (2 External validation sets with 101 cases and 281 cases)	Ex: 98%
Kambakamba et al, 2019 (25)	Prediction of post-operative fistula	ML and Texture analysis	101 cases	In (AUC): 0.72
Mu et al., 2020 (26)	Prediction of pancreatoenteric fistula	CNN	95 cases and 303 controls	In (AUC): 0.85 Ex (AUC): 0.78
Watson et al., 2021 (27)	Prediction of Response to Neoadjuvant Therapy	CNN and LeNet	81 cases	In (AUC): 0.738
Zhang et al., 2020 (28)	Survival Prediction	Transfer learning; CNN	68 cases and 422 controls and external validation (30 cases)	Ex (Concordance index): 0.65
Kaissis et al, 2020 (29)	Quasi mesenchymal identification of PDAC	RF; Radiomic Feature extraction (Pyradiomics)	207 cases	In (AUC): 0.93
Ma et al., 2020 (30)	pancreatic cancer	CNN	222 cases and 190 controls	In (Acc): 95%

160

161 ML-guided tools' design and desired outputs must be tailored toward implementation in
 162 healthcare systems (14). Thus, the current study describes the protocol for developing several
 163 computer-aided diagnoses (CAD) models to facilitate pancreatic cancer management using CT
 164 scan images. We sought to develop generalizable CAD systems to aid clinicians with a
 165 pancreatic cancer diagnosis (classification, segmentation, cancer subtype detection) and
 166 prognosis (cancer resectability, survival, staging) based on pancreas protocol CT scan images
 167 from 10 medical centers.

168 **Material and methods**

169 **Ethical consideration**

170 The Institutional Review Board of Research Institute for Gastroenterology and Liver Diseases
171 (RIGLD), Shahid Beheshti University of Medical Sciences review board approved this
172 ambispective study after consideration of data anonymization and security (code:
173 IR.SBMU.RIGLD.REC.1401.043; link:
174 <https://ethics.research.ac.ir/EthicsProposalViewEn.php?id=323598>). This protocol and future
175 studies adhere to Helsinki Declaration of 1975, as revised in 2008, which provides ethical
176 guidelines for medical research involving human subjects. We have taken necessary measures to
177 protect the privacy and confidentiality of all participants and their personal data. Patients will be
178 included prospectively from 1 December 2022 until March 2024 and retrospectively from 21
179 March 2015 to 23 October 2022. Informed consent will be collected through a phone call from
180 the patient or their legal representative, providing a detailed description of the research aim and
181 use. However, informed consent collection has been waived for patients collected retrospectively
182 or in cases where access to the patient is not possible.

183 **Reporting guidelines and checklist**

184 PanCanAID studies will be conducted adhering to the Standards for Reporting Diagnostic
185 Accuracy (2015-STARD and STARD- AI) and Checklist for Artificial Intelligence in Medical
186 Imaging (CLAIM) (31-33). The STARD checklist, flow diagram, and CLAIM is presented in **S1**
187 **Table**, **S1 Fig**, and **S2 Table**, respectively.

188 **Interdisciplinary Team Building**

189 Starting in January 2021, a biweekly session was extended, with invitations sent to GI referral
190 centers and healthcare providers. The study design involved a team of radiologists,
191 gastroenterologists, surgeons, and computer science experts collaborating to develop the study.

192 Together, they discussed the data-gathering process, labeling techniques, and the desired
193 machine learning tasks, which resulted in the current study design.

194 **Study design**

195 This multicentric observational ambispective study will be conducted in ten medical centers in
196 Tehran, Tabriz, and Guilan provinces of Iran: Taleghani Hospital in Tehran (T-H), Emam
197 Khomeini Hospital in Tehran (EK-H), Firozgar Hospital in Tehran (F-H), Emam Hossein
198 Hospital in Tehran (EH-H), Razi Hospital at Guilan province (R-H), Valiasr International
199 Hospital at Tabriz province (V-H), Shariati Hospital in Tehran (S-H), Namazi Hospital at Shiraz
200 Province (N-H), Shahid Faghihi Hospital at Shiraz (SF-H), Behboud Specialized Clinic for
201 Gastroenterology Diseases (B-C), and the Research Institute of Gastroenterology clinic (RIG-C).
202 Patients will be included prospectively from 1 December 2022 until March 2024 and
203 retrospectively from 21 March 2015 to 23 October 2022. An ethical review board waived
204 gathering informed consent.

205 Patient demographic data will be collected from the electronic hospital information system
206 (HIS). Image data will be obtained from the Picture Archiving and Communication Systems
207 (PACS). EUS images will be obtained from a dedicated system (EndoPACS) at each local
208 hospital system if available. CT scan and EUS images will be gathered in the “.dicom” and
209 “.jpg” series. The medical team conducting the study will call all enrolled patients within two
210 weeks of the enrollment to evaluate the survival and outcome of the cancer. **Fig 2** demonstrates
211 the workflow and aims of PanCanAID.

212 **Fig 2. Workflow and aims of PanCanAID, a multicentric study to facilitate diagnosis and**
213 **management of pancreas cancer.** Footnote: The black circle represents a contrast-enhanced CT
214 scan, and the half-black circle represents a non-contrast CT scan

215 **Patient eligibility, identification, and validation**

216 Potential pancreatic cancer cases will be defined accord to using the following criteria: (1)
217 international classification of diseases (ICD) code C25, (2) a histological diagnosis of pancreatic
218 ductal adenocarcinoma (PDAC) OR pancreatic neuroendocrine tumor (PNET) or (3) a radiologic
219 diagnosis of a pancreas mass OR pancreas tumor. Prospective enrollment of potential cases will
220 occur at the time of referral to a radiology center or gastroenterology clinic.

221 Benign and premalignant lesions of the pancreas will be excluded from the case group. Patients
222 without any valid CT scan before the initiation of treatment (chemotherapy or surgery) will be
223 excluded. The treatment initiation will be obtained during a follow-up call or review of patient
224 HIS records. Cases between 20 and 80 years old, with valid CT scan imaging and histologic
225 confirmation of PDAC and PNET from pancreas specimens collected during surgery or FNA
226 biopsy, will be included in the study. The suited CT scan for inclusion of cases and controls is
227 triple phase CT scan or with and without contrast enhanced CT scan.

228 The control group will comprise three subgroups of patients aged 20 to 80 undergoing
229 abdominopelvic CT scan and EUS examination but have no evidence of abdominopelvic
230 malignancy or history of mass resection. The first subgroup will consist of patients without any
231 pancreatic neoplasms. The second subgroup will consist of patients with premalignant lesions,
232 primarily Intraductal papillary mucinous neoplasms (IPMNs), along with pancreatic
233 intraepithelial neoplasia (PanIN) and mucinous cystic neoplasms (MCN) confirmed by pathology
234 report (10). The third subgroup will consist of patients with acute or chronic pancreatitis
235 confirmed by the radiologist. The selection of three subgroups aim to enrich the control group to
236 represent the real-world challenge of diagnosing pancreatic cancer.

237 **Phone Interview for Survival and clinical data**

238 Our prognostic model aims to predict the 6-month and 1-year survival of patients with pancreatic
239 cancer. To achieve this, the medical team will conduct follow-up calls to both cases and controls,
240 enrolled either retrospectively (over the past year) or prospectively, to collect survival time and
241 relevant risk factors (10). The phone interview form used during these calls is attached in **S1**
242 **Appendix**. It includes information on the patient's demographics, blood group, symptoms,
243 diagnostic exams, treatments, smoking/alcohol consumption, diabetes/pancreatitis history, family
244 history of cancer/pancreatitis, first symptom-diagnosis interval, and diagnosis-death interval. In
245 addition to the phone interview data, two previously collected datasets from F-Hospital (with a
246 6-year follow-up) and Ekh-Hospital (with a 1-year follow-up) will be used, along with available
247 imaging data retrieved from the PACS system.

248 **Sample size**

249 Sample size estimation in machine learning projects in medical imaging requires an initial set of
250 annotated data, which in our case, was unavailable (34). Moreover, a reliable method for
251 estimating the sample size of biomedical ML research is unclear, especially with rapidly
252 evolving modeling techniques (34). Figueroa et al. have proposed that between 80 and 560
253 annotated samples in each class are needed to achieve a root mean squared error lower than 0.01
254 (35). Similar studies using CT scan images achieved satisfactory results, with about 250 cases in
255 each class of PDAC and non-PDAC (**Table 1**). The primary aim of PanCanAID is to collect data
256 from 300 PDAC cases and 300 controls. We will examine the sample size using post hoc curve-
257 fitting and the Figueroa method once the first 150 PDAC cases have been collected (75 cases and
258 75 controls) (35).

259 **Hospitals and imaging devices**

260 Data will be collected from ten medical centers in Tehran. The center attributes, including the
 261 CT scanner model and weekly incidence of pancreatic cancer scans, are described in **Table 2**.
 262 Different imaging devices and technical variations may result in batch effect, so each patient's
 263 imaging device and hospital sources will be recorded.

264 **Table 2. Information of ten medical centers that will participate in patient enrollment.**
 265 Footnote: PDAC: Pancreatic ductal adenocarcinoma; PNET: pancreatic neuroendocrine tumor.

Hospital	GI referral center	Number of beds	CT scan Device	Estimated Workload of Pancreas cancer per week	Previously available data
Ekh-H	Yes	2100	16 Detector Siemens SOMATOM Emotion	15	Yes (100 PNET Cases with CT scan)
EH-H	No	500	16 Detector Siemens Somatom	5	No
T-H	Yes	500	16 Decotor Siemenc Somatom	10	No
F-H	Yes	554	16 Detector Siemens Somatom	10	Yes (200 PDAC patients with 2-year follow-up)
R-H	Yes	500	-	7	No
V-H	No	1000	-	3	No
Sh-H	Yes	1000	2-MDCT Siemens Somatom Volume Zoom, Siemens	15	No
S-H	Yes	850	16-detector Somatom Emotion, Siemens	7	Yes (42 PDAC)
B-C	Yes	-	-	3	No
RIG-C	Yes	-	-	5	Yes (150 PDAC patients)

266 **Online open datasets**

267 We reviewed all previously published open-source pancreatic cancer CT imaging datasets,
 268 presented in **Table 3**. Previously segmented images from datasets such as WORD and

269 AbdomenCT-1k will segment the pancreas on local images. This segmentation will be revised by
 270 the first investigator and confirmed by an experienced radiologist.

271 **Table 3. Online open dataset for pancreas cancer CT scan imaging.** Footnote: *Data used in
 272 this publication were generated by the National Cancer Institute Clinical Proteomic Tumor
 273 Analysis Consortium (CPTAC). Abbreviations: NCI: national cancer institute (of the United
 274 States of America)

Name	Source	Data	License
AbdomenCT-1k (36)	GitHub (JunMa11/AbdomenCT-1k)	abdominal CT organ segmentation dataset with 1000+ CT scans by augmenting the existing single organ datasets	Apache-2.0 license
WORD (37)	GitHub (HiLab-git/WORD)	abdominal CT organ segmentation dataset	GNU General Public License v3.0
Vindr (38)	Vindr.ai	Dataset of 1188 scans for phase recognition in abdominal contrast-enhanced CT	One can use the dataset without charge for non-commercial research purposes only
CPTAC-PDA (39)	The Cancer Imaging Archive (TCIA)* (40)	The NCI Clinical Proteomic Tumor Analysis Consortium collected CT scan images after pathological confirmation of 107 pancreas cancer cases	TCIA data usage policy
Pancreas-CT (41, 42)	The Cancer Imaging Archive (TCIA)* (40)	The National Institutes of Health Clinical Center performed 82 abdominal contrast-enhanced 3D CT scans in the portovenous phase	TCIA data usage policy
Pancreatic-CT-CBCT-SEG (43, 44)	The Cancer Imaging Archive (TCIA)* (40)	Breath-hold CT and cone-beam CT images with expert manual organ-at-risk segmentations from radiation treatments of 40 locally advanced pancreatic cancer patients	TCIA data usage policy
Our future dataset (PanCanAID)	The Cancer Imaging Archive or other imaging repositories	We plan to provide biphasic CT scans of 500 pancreas cancer cases with segmentation on arterial phase and patient outcome in 200 cases	GNU Affero General Public License v3.0

275 Manual segmentation

276 We used a panel discussion and Chu et al.'s experience to decide on the segmentation strategy
 277 (17). Six radiologists will annotate and classify each axial plane of the abdominopelvic CT scan

278 images in arterial phases of CECT using an approaches, an offline 3D slicer 5.0.3 program or an
279 online XNAT server (45). The 3D slicer software will be used on a Windows-based local
280 computer, and annotations will be made using lase pen (XP pen Deco 01 v2). “Brush” and “pen”
281 tools will be used after setting the editable intensity Hounsfield range using the “threshold” tool.
282 Using “threshold” will prevent selecting surrounding elements with different Hounsfield units. In
283 the second approach, and for ease of access, an XNAT application on a Linux server with 200
284 Gb of storage and a two core 8Gb ram (46).

285 A second radiologist with expertise on abdominal imaging will review and confirm the
286 annotations (not blinded to previous segmentations). In case of conflict, the data will be tagged
287 as controversial, and conflicts will be resolved in a dedicated conflict resolution panel with two
288 radiologists. Instruction for radiologists in the Persian language will be available before
289 annotation, and five dedicated cases for educational purposes have been designed to ensure
290 labeling and annotation uniformity.

291 **Active learning for segmentation**

292 Providing ground truth annotations for medical images, especially in the case of pancreas
293 images, is very time-consuming and requires limited expert resources. This is especially so for
294 segmentation, where pixel-wise annotation is needed. Hence, we utilize active learning to
295 interact with the annotator. Active learning is a technique in which a machine learning algorithm
296 can improve its accuracy using less labeled training data by selecting the most informative data
297 to learn from. Instead of being given a fixed set of labeled data, an active learner can ask an
298 oracle to label additional instances that are most useful for improving its performance (47).
299 Active learning has been shown to be effective in radiology AI studies (48). We propose an
300 automated system to carefully select the most representative data samples for annotation. We

301 also consider the model’s uncertainty and approximate error probability on the new data sample
 302 for the selection. The selected samples are then given to the model for initial annotation. The
 303 annotated image is given to the radiologist to correct the annotation of the model and then added
 304 to the set of labeled datasets. After several steps, the newly labeled dataset is given to the model
 305 for retraining. We continue until the performance improvement stops or a pre-defined proportion
 306 of the dataset is labeled. The process is summarized in **Fig 3**.

307 **Fig 3. Workflow for manual segmentation using active learning approach.**

308

309 **Mass characteristics**

310 Radiologists will evaluate tumor characteristics, and this data will be used in the future phases of
 311 PanCanAID. These characteristics can be used to utilize automated reporting of pancreas mass in
 312 the future. **Table 4** shows pancreas cancer characteristics.

313 **Table 4. The routine report of mass characteristics and source of labeling in PanCanAID.**

314 Footnote: CBD: common bile duct; PD: pancreatic duct

Cancer Feature	Classes	Source of labeling
Location	Periampullary-head-body-tail	Radiologist
Morphology	Solid-cystic-mixed	Radiologist
Mass size		Manual segmentation
CBD and PD diameter	Largest diameter in mm in any plane	Radiologist
CBD dilation	Yes-no	Radiologist
Stent in situ	Yes-no	Radiologist
TNM stage	Size (T), regional lymph node metastasis (N), non-regional metastasis (M)	Radiologist- manual segmentation (size)
DPCG criteria for classification	Resectable-borderline- irresectable	Radiologist- manual segmentation (size)

315

316 **Resectability definition and staging**

317 The National Comprehensive Cancer Network (NCCN) and Dutch Pancreatic Cancer Group
318 (DPCG) guidelines can help physicians in assessing the resectability of tumors (49, 50). This
319 clinical decision is usually challenging for even expert radiologists. Both over- and under-
320 treating can significantly impact a patient’s quality of life. We choose the DPCG criteria (**Table**
321 **5**) because of its simplicity and lower classification workload. An expert radiologist will classify
322 CT scan images in the arterial phase, and this labeled data will be used to predict the resectability
323 of mass. For TNM staging, a segmented pancreas mass will inform the tumor size (T-stage). A
324 tumor with its longest diameter of less than 2 centimeters in an axial CT scan is defined as T1.
325 Tumors with a diameter of 2-4 centimeters and those wider than 4 centimeters correspond to the
326 T2 and T3 stages, respectively. The T4 stage consists of tumors involving vessels such as the
327 celiac trunk, hepatic artery, and superior mesenteric artery. Radiologists will classify lymph node
328 metastasis (N-stage) as a distant or regional invasion.

329 **Table 5. The Dutch Pancreatic Cancer Group (DPCG) criteria to assess the resectability of**
330 **pancreatic cancer.** Footnote: CA: celiac artery; SMA: superior mesenteric artery; CHA:

331 common hepatic artery; SM: superior mesenteric vein; PV: portal vein.

	Resectable	Borderline	Irresectable
Arteries (CA, SMA, CHA)	No contact	< 90 degrees	>90 degrees
Veins (SMV, PV)	<90 degrees (contact)	90-270 degrees (no occlusion)	>270 degrees (or occlusion)
Metastasis	no	no	yes
N-Stage	Locoregional	Locoregional	Extra regional

332

333 **Data bank storage and computer processors**

334 Data including “.dicom” files of CT scan images, “.jpg” image of EUS, “.csv” files with
335 metadata (including patient characteristics, labels, hospital source, and notes as attached in

336 Supplementary File 1), “.nrrd” files of 3D slicer with manual segmentations will be stored in
337 three external hard drives with two terabyte storage. Processing will be handled using a pair of
338 GTX 1080Ti GPUs. Performance and inference time will be evaluated on both GPU and CPU.

339 **CAD systems and tasks**

340 As **Fig 2** depicts, different CAD models will be developed using a databank with specified aims,
341 including:

342 1- Phase detection in abdominopelvic CT scan: classification (non-contrast phase, arterial phase,
343 venous phase, portal phase, delay phase)

344 2- Pancreas organ segmentation (pancreas organ segmentation in CECT and non-contrast
345 enhanced abdominopelvic CT scan)

346 3- Diagnosis of pancreas cancer in CECT scan: Classification (cancerous vs. non-cancerous),
347 segmentation (pixel perfect pancreas organ and mass), and cancer subtype (PDAC and PNET) in
348 CECT scan images (arterial phase)

349 4- Diagnosis of pancreas cancer in non-contrast CT scan: Classification (cancerous vs. non-
350 cancerous), segmentation (pixel perfect pancreas organ and mass), and cancer subtype
351 classification (PDAC and PNET) in non-contrast abdominopelvic CT scan images

352 5- Prognosis and survival of pancreas cancer in CECT: TNM stage classification including (T:
353 size, N: lymph node metastasis, M: distant metastases), and survival (months) in contrast-enhanced
354 CT scan images

355 6- DPCG resectability classification (resectable, borderline, and irresectable) in CECT scan images
356 (arterial phase)

357 Future Direction: Multimodal approach for pancreas cancer ML tasks: Classification (cancerous
358 vs. non-cancerous), cancer subtype (PDAC and PNET), resectability (resectable vs. irresectable),
359 and survival (month) using demographics, EUS, and CT scan images

360 **Experiments and model development**

361 Each input image is contrast-enhanced and denoised for all tasks. Normalization steps are
362 conducted to mitigate the variability in data samples due to experimental conditions and
363 measurement device configuration. For CT images, unnecessary slices are removed from the
364 beginning and end of CT image sequences. We design and train a separate classification model
365 that detects slices that have the pancreas visible, using the labels already gathered for our data.
366 We consider slices with a minimum amount of detected pancreas area as positive and otherwise
367 as negative. This filtering allows our models to be guided by useful cues and removes the
368 computational cost of processing extra slices.

369 **Fig 4.** shows a brief overview of the model development workflow. The tasks are solved
370 according to the following procedure:

371 **Fig 4. A brief overview of PanCanAID model development workflow.**

372 1- A multi-instance learning approach is deployed to find substantial slices of data and make an
373 initial detection of cancer. The features extracted in this step are also used for detecting and
374 segmenting cancerous tissue in the following steps.

375 2- To extend our dataset, we conduct data augmentation. We use the variations between data
376 samples to generate novel data. We train a spatial and an intensity registration network.
377 Registration is a method to match images with the same structure but is distorted and hence not
378 pixel-by-pixel matched (51). It finds a transformation that can map corresponding pixels to each
379 other. We sample from the detected intensity and spatial maps learned from these models
380 throughout the dataset and combine them to create new maps to generate new data.

381 3- A segmentation of the pancreas alongside any suspicious tissue mass is carried out in the
382 CECT image. Due to the lack of labeled data and the cost of manual segmentation, we follow a
383 domain adaptation approach. Domain adaptation has been extensively used in medical data and
384 signals (52), especially in CT images (53)It is used when a model is supposed to be adapted to an
385 unlabeled external domain with the help of the labeled internal dataset. We train a model on
386 public datasets with segmentation labels and adapt the trained model to our dataset. We deploy
387 different reconstruction methods to utilize the unlabeled data in our dataset. Reconstruction is
388 mainly used to force latent features to contain as much information as possible to recreate the
389 input data patterns (54). We gradually add labeled data and update the model according to an
390 active learning framework to carry out the segmentation more accurately.

391 4- The cancer classification task uses segmentation from the original and reconstructed images.
392 The delta image, the original image, and the features extracted in step 1 are processed through
393 another Convolutional Network to detect cancer.

394 5- Upon cancer detection, we perform an additional classification task (PDAC or PNET) by
395 processing a neighbor of the segmented mass in relation to the whole image. Local features are
396 obtained by processing the area around the pancreas as well as the pancreas itself. Global
397 features are computed based on the global attention of the whole image, considering the

398 relationship between all regions in the image, providing a global representation of the structure
399 of the abdominal CT image (55). We use local features around the segmented area and their
400 relationship along with the global features of the entire image to model both the local
401 information and the significance of this information relative to the global structure of the
402 abdominal area. Local features

403 6- For the resectability classification task, we follow an approach similar to detecting the cancer
404 type. We estimate the chance of successful tumor resection based on local-global feature
405 extraction.

406 7- Segmentation estimation for plain non-contrast CT images is done by matching them to the
407 label assigned to their corresponding CE-CT images. As we have segmentation maps for CE-CT
408 images, we temporally align CT slices of patients with their CE-CT image to use the CE-CT
409 labels as ground truth annotation for the plain CT images. The rest of the procedure is similar to
410 the one for CE-CT images. Other tasks performed on the basic CT images are done similarly, as
411 the true values of cancer labels, resectability, and prognosis results are the same for CE and plain
412 CT images.

413 8- For prognosis, we model survival time as a conditioned normal variable, with mean and
414 variance predicted by a 3D-CNN applied on the combination of the segmentation and original
415 image. The mean head estimates the average survival time, and the (log) variance head estimates
416 the uncertainty of the prediction. We infer mean and variance by maximizing the likelihood of
417 the data.

418 **Batch effect removal**

419 We designed a multi-level multi-site batch normalization (MMBN) architecture to remove the
420 batch effect. We aim to remove batch effect at both the data- and feature-level.

421 **Data-level batch effect removal**

422 To remove the effect of different measurement devices, we normalize the intensity and contrast
423 of the CT images. We also apply affine normalization to remove geometrical biases caused by
424 the experimental conditions.

425 **Feature-level batch effect removal**

426 Only some of the variations in the data can be detected by low-level analysis of raw input
427 images. We utilize our multi-site dataset to remove the feature-level batch effect that occurs in
428 higher data abstractions. We deploy a Multi-Site Batch Normalization Layer (MSBNL) that
429 consists of a batch normalization layer for each site in our dataset. The data sample is normalized
430 according to the normalization parameters of its site. For a target site, we first estimate the mean
431 and variance of the samples. Then, for each target-site sample, we pass it to each site-specific
432 batch normalization layer and aggregate them using the weights defined as the KL divergence of
433 the distribution of target site data and the corresponding source site data. To lessen the
434 computational complexity, we assume a normal distribution for the data in each site. More
435 concretely, assume site statistics (mean and variance) estimates from batch normalization
436 parameters are (μ_i, σ_i^2) for each site-specific batch normalization layer and (μ_t, σ_t^2) are the
437 statistics of the target site. The weight of each layer for a sample from a target sample is
438 calculated according to Equations 1 and 2:

439
$$\text{Equation 1: } \hat{w}_i = \log \frac{\sigma_i}{\sigma_t} + \frac{1}{2\sigma_i^2} (\sigma_t^2 + (\mu_t - \mu_i)^2) - \frac{1}{2}$$

440
$$\text{Equation 2: } w_i = \frac{\hat{w}_i}{\sum_{j=1}^n \hat{w}_j}$$

441

442 If we don't have enough data from the target site, we also add the likelihood of the target data
443 according to the distribution of each site using Equation 3:

444
$$\text{Equation 3: } \tilde{w}_i = \frac{1}{\sqrt{2\pi\sigma_i^2}} \exp\left[-\frac{1}{2\sigma_i^2}(x_t - \mu_i)^2\right]$$

445 Where x_t , is a feature of the sample from the target data, μ_i and σ_i are expected value and
446 variance of the feature for each domain according to the batch normalization layer, and $\tilde{w} = \hat{w} +$
447 \tilde{w} which means that we add computed \hat{w} parameters to \tilde{w} to account for a small amount of data
448 in the target domain.

449 **Evaluation metrics and proposed model**

450 The model performance will be tested in internal validation (test set) and external validation
451 (from external hospital). For segmentation, we use IOU (intersection over union) to evaluate the
452 proportion of detected mass. We also measure pixel-wise sensitivity and specificity to assess the
453 model's power to correctly find the true segmented areas and discard the unsegmented areas. F1
454 concludes these concepts as a single number. For classification (cancer and resectability), we
455 measure the area under the receiver operating characteristic curve (AUROC), accuracy,
456 sensitivity, specificity, and F1 by using the K-fold cross-validation technique (56). In addition, a
457 calibration curve will be used to show how well the probabilistic predictions of a binary
458 classifier are calibrated (57). We perform a simple statistical test to evaluate the prognosis's
459 predicted mean and variance values. For each ground truth prediction in the test dataset, we

460 examine if it comes from the proposed normal distribution or not. The proportion of samples that
461 pass the test is defined as a measure of the model's performance.

462 **Explainable AI**

463 The adaptation of AI models in the clinical setting has been constrained by their “black-box”
464 nature, which makes it challenging for clinicians to comprehend and believe their predictions
465 (58). Herby, we will embed Explainable AI (XAI) approach to increase its use case in the
466 medical field. XAI is one of the branches of artificial intelligence concerned with building
467 models that can offer clear and understandable justifications for their predictions and choices. By
468 using architectures such as U-net, we can incorporate XAI techniques, such as Layer-wise
469 Relevance Propagation (LRP) or Local Interpretable Model-agnostic Explanations (LIME), and
470 we can provide visual explanations of the model's predictions. Also, by segmenting the
471 cancerous masses and pinpointing their location, medical professionals can better understand the
472 reasons behind the model's decisions and build trust in its predictions (58-60).

473 **Discussion**

474 The development of CDSS and CAD tools requires interdisciplinary teamwork (17). PanCanAID
475 is a multipurpose CDSS project addressing the current needs of pancreatic cancer care delivery
476 across multiple phases of the disease. In developing the PanCanAID protocol, we addressed
477 various aspects of team building, data collection and annotation, and model development. The
478 process evolved over months of collaborative sessions with medical and computational experts.
479 Many challenges were faced during protocol design, including the feasibility of data collection
480 and annotation, complex multidisciplinary collaboration, optimization of data storage capacity,

481 and the development of a state-of-the-art ML model. The protocol development process
482 addressed these challenges and fostered effective cross-disciplinary collaboration toward a
483 common goal.

484 CDSSs depend highly on a large, precisely-labeled dataset representing real-world data patterns
485 (14). However, the truly adequate dataset size is unknown without conducting pilot studies (34).
486 ML models of medical imaging may require even large datasets than models of tabular data due
487 to the complexities of imaging data. The lack of publicly available cases on The Cancer Imaging
488 Archive (TCIA) highlights the challenge of acquiring pancreatic cancer image data (**Table 3**).

489 We constructed a multi-institutional team and designed it to collect sufficient cases for our
490 models. Our initial aim is to gather 500 cases and 500 controls, but this number may be extended
491 to reach the desired model performance: an AUC of 0.85 for diagnosis and 0.80 for cancer
492 prognosis.

493 The data collection task for pancreatic cancer is more difficult for three reasons. First, a
494 pancreatic cancer diagnosis is a relatively rare event (2). Second, the life span of patients is short,
495 and many cases pass away within the first months, which makes patient identification even
496 harder (1, 61). The third and most important reason is the winding journey of pancreatic cancer
497 patients during diagnosis and management. Multiple medical centers and specialties manage
498 pancreatic cancer, and the data is stored in various sources (**Fig 5**). The collection of this data
499 needs rigorous amounts of time and a substantial amount of effort (17). We used an ambispective
500 design to collect more accessible cases retrospectively and precise data prospectively. However,
501 finding pathologically confirmed patients with survival outcomes can be unattainable in many
502 cases. We aim to collect the survival of patients, which is needed for some tasks, by calling
503 patients diagnosed in the last year or included prospectively.

504 **Fig 5. Dissemination, storage, patient data sources, and the journey of patients with**
505 **pancreas cancer.**

506 The next issue is the time and effort needed to label the data. In our case, tumor characteristics,
507 pancreas segmentation, tumor segmentation, and resectability assessment were warranted. We
508 assembled a team of junior radiologists with six years of experience who will handle
509 segmentation. An active learning model will also be developed using publicly available data for
510 pancreas organ segmentation and less than 100 local cases. Two senior radiologists who are
511 experts in the field will validate and confirm the labeled data. The interobserver variability will
512 be reported to show the use case of CADs.

513 In addition, we sought to collect other data, such as EUS images and clinical data (symptoms,
514 past medical history, and demographics). We will use the collected data to develop a cross-
515 modality platform, which is the future direction of PanCanAID. This data bank can overcome the
516 current bottleneck in the model development of pancreatic cancers.

517 Our project involves a comprehensive benchmarking of previous models and the development of
518 new algorithms. However, several challenges make our task more complex than a typical
519 classification/segmentation task. Firstly, labeling all CT scans, particularly those with narrow
520 imaging cuts, is impractical due to the high cost of segmentation labeling. To overcome this, we
521 have adopted an active learning approach where radiologists interact with our team to validate
522 and correct segmentation labels. This results in higher quality annotations and improved sample
523 efficiency. Secondly, the complex structure of abdominal images, especially the irregular shape
524 of the pancreas, can benefit from large-scale external data with extensive labeling. We aim to
525 deploy a domain adaptation framework to transfer the knowledge learned from the large-scale
526 external data into our dataset. Thirdly, we have designed a batch-effect removal protocol to

527 eliminate batch effects in our multi-source data, which we will extend using other domain
528 generalization techniques. Finally, we aim to build our model with unlabeled data, using semi-
529 supervised learning as a key component of our framework.

530 Although we tried to overcome several challenges during the protocol design process, several
531 unknown factors could still affect our future work. The presence of all three needed data
532 modalities (CT scans, pathology reports, and survival data) may be unachievable in many
533 pancreatic cancer cases. In addition, the quality of CT scan images, especially in the arterial
534 phase, may be insufficient. Developing a generalizable model will need rigorous effort for batch
535 effect removal. In addition, the segmentation and labeling of data require a vast amount of time
536 from radiologists, which may be exhausting. We hope to overcome upcoming challenges through
537 interdisciplinary teamwork.

538 **Conclusion**

539 PanCanAID is a large-scale AI project developing CADs and CDSSs using pancreatic cancer CT
540 scan images. In hopes of improving pancreatic cancer prognosis, it will tackle the current
541 bottleneck of model development and data shortage. We plan to collect good quality, sufficient
542 amounts, and precisely labeled data banks by creating a team of experts from various institutions.
543 Besides, in our model development, we utilize and expand different concepts according to the
544 challenges we have in our task, including active learning, semi-supervised learning, and domain
545 adaptation and generalization. Experts in medical and computational fields were involved in
546 protocol development, striving to describe the problem from all aspects. The protocol design
547 lasted for months, but it fostered the replicability of the method and cross-disciplinary teamwork.

548 **Acknowledgment**

549 The authors want to thank the National Elite Foundation, Research Institute of Gastroenterology
550 and Liver Diseases (Shahid Beheshti University of Medical Sciences), Data Science and
551 Machine Learning (DML) Lab (Sharif University of Technology), Advanced Diagnostic and
552 Interventional Radiology Research Center (Tehran University of Medical Sciences),
553 Gastroenterology and Liver Diseases Research Center (Iran University of Medical Sciences),
554 Gastrointestinal and Liver Diseases Research Center (Guilan University of Medical Sciences),
555 and Medical Imaging Research Center (Shiraz University of Medical Sciences) for providing
556 resources and support.

557 **References**

- 558 1. Khalaf N, El-Serag HB, Abrams HR, Thrift AP. Burden
559 of Pancreatic Cancer: From Epidemiology to Practice.
560 Clin Gastroenterol Hepatol. 2021;19(5):876-84. DOI:
561 10.1016/j.cgh.2020.02.054.
- 562 2. Siegel RL, Miller KD, Jemal A. Cancer statistics,
563 2019. CA Cancer J Clin. 2019;69(1):7-34. DOI:
564 10.3322/caac.21551.
- 565 3. Aier I, Semwal R, Sharma A, Varadwaj PK. A
566 systematic assessment of statistics, risk factors, and
567 underlying features involved in pancreatic cancer.
568 Cancer Epidemiol. 2019;58:104-10. DOI:
569 10.1016/j.canep.2018.12.001.
- 570 4. Egawa S, Toma H, Ohigashi H, Okusaka T, Nakao A,
571 Hatori T, et al. Japan Pancreatic Cancer Registry; 30th
572 year anniversary: Japan Pancreas Society. Pancreas.
573 2012;41(7):985-92. DOI: 10.1097/MPA.0b013e318258055c.
- 574 5. Kato S, Honda K. Use of Biomarkers and Imaging for
575 Early Detection of Pancreatic Cancer. Cancers (Basel).
576 2020;12(7). DOI: 10.3390/cancers12071965.
- 577 6. Tempero MA, Malafa MP, Al-Hawary M, Asbun H, Bain
578 A, Behrman SW, et al. Pancreatic Adenocarcinoma,
579 Version 2.2017, NCCN Clinical Practice Guidelines in

- 580 Oncology. *J Natl Compr Canc Netw*. 2017;15(8):1028-61.
581 DOI: 10.6004/jnccn.2017.0131.
- 582 7. Edge SB, Compton CC. The American Joint Committee
583 on Cancer: the 7th Edition of the AJCC Cancer Staging
584 Manual and the Future of TNM. *Annals of Surgical*
585 *Oncology*. 2010;17(6):1471-4. DOI: 10.1245/s10434-010-
586 0985-4.
- 587 8. Jones S, Zhang X, Parsons DW, Lin JC, Leary RJ,
588 Angenendt P, et al. Core signaling pathways in human
589 pancreatic cancers revealed by global genomic analyses.
590 *Science*. 2008;321(5897):1801-6. DOI:
591 10.1126/science.1164368.
- 592 9. Yoshida T, Yamashita Y, Kitano M. Endoscopic
593 Ultrasound for Early Diagnosis of Pancreatic Cancer.
594 *Diagnostics (Basel)*. 2019;9(3). DOI:
595 10.3390/diagnostics9030081.
- 596 10. McGuigan A, Kelly P, Turkington RC, Jones C,
597 Coleman HG, McCain RS. Pancreatic cancer: A review of
598 clinical diagnosis, epidemiology, treatment and
599 outcomes. *World J Gastroenterol*. 2018;24(43):4846-61.
600 DOI: 10.3748/wjg.v24.i43.4846.
- 601 11. Zakaria A, Al-Share B, Klapman JB, Dam A. The Role
602 of Endoscopic Ultrasonography in the Diagnosis and
603 Staging of Pancreatic Cancer. *Cancers (Basel)*.
604 2022;14(6). DOI: 10.3390/cancers14061373.
- 605 12. Maguchi H. The roles of endoscopic ultrasonography
606 in the diagnosis of pancreatic tumors. *J Hepatobiliary*
607 *Pancreat Surg*. 2004;11(1):1-3. DOI: 10.1007/s00534-002-
608 0752-4.
- 609 13. Gonoï W, Hayashi TY, Okuma H, Akahane M, Nakai Y,
610 Mizuno S, et al. Development of pancreatic cancer is
611 predictable well in advance using contrast-enhanced CT:
612 a case-cohort study. *European Radiology*.
613 2017;27(12):4941-50. DOI: 10.1007/s00330-017-4895-8.
- 614 14. Doi K. Computer-aided diagnosis in medical imaging:
615 historical review, current status and future potential.
616 *Comput Med Imaging Graph*. 2007;31(4-5):198-211. DOI:
617 10.1016/j.compmedimag.2007.02.002.

- 618 15. Francis IR. Role of CT and MR in detection and
619 staging of pancreatic adenocarcinoma. *Cancer Imaging*.
620 2003;4(1):10-4. DOI: 10.1102/1470-7330.2003.0026.
- 621 16. Tonozuka R, Itoi T, Nagata N, Kojima H, Sofuni A,
622 Tsuchiya T, et al. Deep learning analysis for the
623 detection of pancreatic cancer on endosonographic
624 images: a pilot study. *J Hepatobiliary Pancreat Sci*.
625 2021;28(1):95-104. DOI: 10.1002/jhbp.825.
- 626 17. Chu LC, Park S, Kawamoto S, Wang Y, Zhou Y, Shen W,
627 et al. Application of Deep Learning to Pancreatic
628 Cancer Detection: Lessons Learned From Our Initial
629 Experience. *J Am Coll Radiol*. 2019;16(9 Pt B):1338-42.
630 DOI: 10.1016/j.jacr.2019.05.034.
- 631 18. Aggarwal R, Sounderajah V, Martin G, Ting DSW,
632 Karthikesalingam A, King D, et al. Diagnostic accuracy
633 of deep learning in medical imaging: a systematic
634 review and meta-analysis. *NPJ Digit Med*. 2021;4(1):65.
635 DOI: 10.1038/s41746-021-00438-z.
- 636 19. Granata V, Fusco R, Setola SV, Castelguidone ELD,
637 Camera L, Tafuto S, et al. The multidisciplinary team
638 for gastroenteropancreatic neuroendocrine tumours: the
639 radiologist's challenge. *Radiol Oncol*. 2019;53(4):373-
640 87. DOI: 10.2478/raon-2019-0040.
- 641 20. Janssen BV, Verhoef S, Wesdorp NJ, Huiskens J, de
642 Boer OJ, Marquering H, et al. Imaging-based Machine-
643 learning Models to Predict Clinical Outcomes and
644 Identify Biomarkers in Pancreatic Cancer: A Scoping
645 Review. *Ann Surg*. 2022;275(3):560-7. DOI:
646 10.1097/sla.0000000000005349.
- 647 21. Jan Z, El Assadi F, Abd-Alrazaq A, Jithesh PV.
648 Artificial Intelligence for the Prediction and Early
649 Diagnosis of Pancreatic Cancer: Scoping Review. *J Med*
650 *Internet Res*. 2023;25:e44248. DOI: 10.2196/44248.
- 651 22. Hayashi H, Uemura N, Matsumura K, Zhao L, Sato H,
652 Shiraishi Y, et al. Recent advances in artificial
653 intelligence for pancreatic ductal adenocarcinoma.
654 *World J Gastroenterol*. 2021;27(43):7480-96. DOI:
655 10.3748/wjg.v27.i43.7480.
- 656 23. Si K, Xue Y, Yu X, Zhu X, Li Q, Gong W, et al.
657 Fully end-to-end deep-learning-based diagnosis of

- 658 pancreatic tumors. *Theranostics*. 2021;11(4):1982-90.
659 DOI: 10.7150/thno.52508.
- 660 24. Liu KL, Wu T, Chen PT, Tsai YM, Roth H, Wu MS, et
661 al. Deep learning to distinguish pancreatic cancer
662 tissue from non-cancerous pancreatic tissue: a
663 retrospective study with cross-racial external
664 validation. *Lancet Digit Health*. 2020;2(6):e303-e13.
665 DOI: 10.1016/s2589-7500(20)30078-9.
- 666 25. Kambakamba P, Mannil M, Herrera PE, Müller PC,
667 Kuemmerli C, Linecker M, et al. The potential of
668 machine learning to predict postoperative pancreatic
669 fistula based on preoperative, non-contrast-enhanced
670 CT: A proof-of-principle study. *Surgery*.
671 2020;167(2):448-54. DOI: 10.1016/j.surg.2019.09.019.
- 672 26. Mu W, Liu C, Gao F, Qi Y, Lu H, Liu Z, et al.
673 Prediction of clinically relevant Pancreatico-enteric
674 Anastomotic Fistulas after Pancreatoduodenectomy using
675 deep learning of Preoperative Computed Tomography.
676 *Theranostics*. 2020;10(21):9779-88. DOI:
677 10.7150/thno.49671.
- 678 27. Watson MD, Baimas-George MR, Murphy KJ, Pickens RC,
679 Iannitti DA, Martinie JB, et al. Pure and Hybrid Deep
680 Learning Models can Predict Pathologic Tumor Response
681 to Neoadjuvant Therapy in Pancreatic Adenocarcinoma: A
682 Pilot Study. *Am Surg*. 2021;87(12):1901-9. DOI:
683 10.1177/0003134820982557.
- 684 28. Zhang Y, Lobo-Mueller EM, Karanicolas P, Gallinger
685 S, Haider MA, Khalvati F. CNN-based survival model for
686 pancreatic ductal adenocarcinoma in medical imaging.
687 *BMC Med Imaging*. 2020;20(1):11. DOI: 10.1186/s12880-
688 020-0418-1.
- 689 29. Kaissis GA, Ziegelmayr S, Lohöfer FK, Harder FN,
690 Jungmann F, Sasse D, et al. Image-Based Molecular
691 Phenotyping of Pancreatic Ductal Adenocarcinoma. *J Clin*
692 *Med*. 2020;9(3). DOI: 10.3390/jcm9030724.
- 693 30. Ma H, Liu ZX, Zhang JJ, Wu FT, Xu CF, Shen Z, et
694 al. Construction of a convolutional neural network
695 classifier developed by computed tomography images for
696 pancreatic cancer diagnosis. *World J Gastroenterol*.
697 2020;26(34):5156-68. DOI: 10.3748/wjg.v26.i34.5156.

- 698 31. Sounderajah V, Ashrafian H, Golub RM, Shetty S, De
699 Fauw J, Hooft L, et al. Developing a reporting
700 guideline for artificial intelligence-centred
701 diagnostic test accuracy studies: the STARD-AI
702 protocol. *BMJ Open*. 2021;11(6):e047709. DOI:
703 10.1136/bmjopen-2020-047709.
- 704 32. Mongan J, Moy L, Kahn CE. Checklist for Artificial
705 Intelligence in Medical Imaging (CLAIM): A Guide for
706 Authors and Reviewers. *Radiology: Artificial
707 Intelligence*. 2020;2(2):e200029. DOI:
708 10.1148/ryai.2020200029.
- 709 33. Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA,
710 Glasziou PP, Irwig L, et al. STARD 2015: an updated
711 list of essential items for reporting diagnostic
712 accuracy studies. *Bmj*. 2015;351:h5527. DOI:
713 10.1136/bmj.h5527.
- 714 34. Balki I, Amirabadi A, Levman J, Martel AL, Emersic
715 Z, Meden B, et al. Sample-Size Determination
716 Methodologies for Machine Learning in Medical Imaging
717 Research: A Systematic Review. *Can Assoc Radiol J*.
718 2019;70(4):344-53. DOI: 10.1016/j.carj.2019.06.002.
- 719 35. Figueroa RL, Zeng-Treitler Q, Kandula S, Ngo LH.
720 Predicting sample size required for classification
721 performance. *BMC Medical Informatics and Decision
722 Making*. 2012;12(1):8. DOI: 10.1186/1472-6947-12-8.
- 723 36. Ma J, Zhang Y, Gu S, Zhu C, Ge C, Zhang Y, et al.
724 AbdomenCT-1K: Is Abdominal Organ Segmentation a Solved
725 Problem? *IEEE Transactions on Pattern Analysis and
726 Machine Intelligence*. 2022;44(10):6695-714. DOI:
727 10.1109/TPAMI.2021.3100536.
- 728 37. Luo X, Liao W, Xiao J, Chen J, Song T, Zhang X, et
729 al. WORD: A large scale dataset, benchmark and clinical
730 applicable study for abdominal organ segmentation from
731 CT image. *Medical Image Analysis*. 2022;82:102642. DOI:
732 <https://doi.org/10.1016/j.media.2022.102642>.
- 733 38. Dao BT, Nguyen TV, Pham HH, Nguyen HQ. Phase
734 recognition in contrast-enhanced CT scans based on deep
735 learning and random sampling. *Med Phys*.
736 2022;49(7):4518-28. DOI: 10.1002/mp.15551.

- 737 39. National Cancer Institute Clinical Proteomic Tumor
738 Analysis Consortium (CPTAC) The Clinical Proteomic
739 Tumor Analysis Consortium Pancreatic Ductal
740 Adenocarcinoma Collection (CPTAC-PDA) (Version 13)
741 [Data set]. In: The Cancer Imaging Archive, editor.
742 2018.
- 743 40. Clark K, Vendt B, Smith K, Freymann J, Kirby J,
744 Koppel P, et al. The Cancer Imaging Archive (TCIA):
745 Maintaining and Operating a Public Information
746 Repository. *Journal of Digital Imaging*.
747 2013;26(6):1045-57. DOI: 10.1007/s10278-013-9622-7.
- 748 41. Roth HR, Lu L, Farag A, Shin H-C, Liu J, Turkbey
749 EB, et al., editors. *DeepOrgan: Multi-level Deep
750 Convolutional Networks for Automated Pancreas
751 Segmentation. Medical Image Computing and Computer-
752 Assisted Intervention -- MICCAI 2015; 2015 2015//;*
753 Cham: Springer International Publishing.
- 754 42. Roth HR, Farag A, Turkbey EB, Lu L, Liu J, Summers
755 RM. Data From Pancreas-CT. In: *Archive TCI*, editor.
756 2016.
- 757 43. Hong J, Reyngold M, Crane C, Cuaron J, Hajj C, Mann
758 J, et al. Breath-hold CT and cone-beam CT images with
759 expert manual organ-at-risk segmentations from
760 radiation treatments of locally advanced pancreatic
761 cancer [Data set]. In: *Archive TCI*, editor. 2021.
- 762 44. Han X, Hong J, Reyngold M, Crane C, Cuaron J, Hajj
763 C, et al. Deep-learning-based image registration and
764 automatic segmentation of organs-at-risk in cone-beam
765 CT scans from high-dose radiation treatment of
766 pancreatic cancer. *Medical Physics*. 2021;48(6):3084-95.
767 DOI: <https://doi.org/10.1002/mp.14906>.
- 768 45. Pieper S, Halle M, Kikinis R, editors. *3D Slicer*.
769 2004 2nd IEEE international symposium on biomedical
770 imaging: nano to macro (IEEE Cat No 04EX821); 2004:
771 IEEE.
- 772 46. Herrick R, Horton W, Olsen T, McKay M, Archie KA,
773 Marcus DS. XNAT Central: Open sourcing imaging research
774 data. *NeuroImage*. 2016;124:1093-6.
- 775 47. Settles B. Active learning literature survey. 2009.

- 776 48. Bangert P, Moon H, Woo JO, Didari S, Hao H. Active
777 learning performance in labeling radiology images is
778 90% effective. *Frontiers in radiology*. 2021;1:748968.
- 779 49. Murakami Y, Satoi S, Sho M, Motoi F, Matsumoto I,
780 Kawai M, et al. National comprehensive cancer network
781 resectability status for pancreatic carcinoma predicts
782 overall survival. *World journal of surgery*.
783 2015;39:2306-14.
- 784 50. Van Laethem J-L, Verslype C, Iovanna J, Michl P,
785 Conroy T, Louvet C, et al. New strategies and designs
786 in pancreatic cancer research: consensus guidelines
787 report from a European expert panel. *Annals of*
788 *oncology*. 2012;23(3):570-6.
- 789 51. Balakrishnan G, Zhao A, Sabuncu MR, Guttag J, Dalca
790 AV. VoxelMorph: a learning framework for deformable
791 medical image registration. *IEEE transactions on*
792 *medical imaging*. 2019;38(8):1788-800.
- 793 52. Sarafraz G, Behnamnia A, Hosseinzadeh M, Balapour
794 A, Meghrazi A, Rabiee HR. Domain Adaptation and
795 Generalization on Functional Medical Images: A
796 Systematic Survey. *arXiv preprint arXiv:221203176*.
797 2022.
- 798 53. Guan H, Liu M. Domain adaptation for medical image
799 analysis: a survey. *IEEE Transactions on Biomedical*
800 *Engineering*. 2021;69(3):1173-85.
- 801 54. Roels J, Hennies J, Saeys Y, Philips W, Kreshuk A,
802 editors. Domain adaptive segmentation in volume
803 electron microscopy imaging. 2019 *IEEE 16th*
804 *International Symposium on Biomedical Imaging (ISBI*
805 *2019)*; 2019: IEEE.
- 806 55. Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones
807 L, Gomez AN, et al. Attention is all you need. *Advances*
808 *in neural information processing systems*. 2017;30.
- 809 56. Hicks SA, Strümke I, Thambawita V, Hammou M,
810 Riegler MA, Halvorsen P, et al. On evaluation metrics
811 for medical applications of artificial intelligence.
812 *Scientific Reports*. 2022;12(1):5979. DOI:
813 10.1038/s41598-022-09954-8.
- 814 57. Niculescu-Mizil A, Caruana R, editors. Predicting
815 good probabilities with supervised learning.

816 Proceedings of the 22nd international conference on
817 Machine learning; 2005.
818 58. Chaddad A, Lu Q, Li J, Katib Y, Kateb R, Tanougast
819 C, et al. Explainable, domain-adaptive, and federated
820 artificial intelligence in medicine. IEEE/CAA Journal
821 of Automatica Sinica. 2023;10(4):859-76.
822 59. Karar ME, El-Fishawy N, Radad M. Automated
823 classification of urine biomarkers to diagnose
824 pancreatic cancer using 1-D convolutional neural
825 networks. Journal of Biological Engineering.
826 2023;17(1):28. DOI: 10.1186/s13036-023-00340-0.
827 60. Montavon G, Binder A, Lapuschkin S, Samek W, Müller
828 K-R. Layer-Wise Relevance Propagation: An Overview. In:
829 Samek W, Montavon G, Vedaldi A, Hansen LK, Müller K-R,
830 editors. Explainable AI: Interpreting, Explaining and
831 Visualizing Deep Learning. Cham: Springer International
832 Publishing; 2019. p. 193-209. DOI: 10.1007/978-3-030-
833 28954-6_10.
834 61. Moffat GT, Epstein AS, O'Reilly EM. Pancreatic
835 cancer-A disease in need: Optimizing and integrating
836 supportive care. Cancer. 2019;125(22):3927-35. DOI:
837 10.1002/cncr.32423.

838

839

840

841 **Supporting Information**

842 **S1 Table.** The Standards for Reporting Diagnostic Accuracy (STARD 2015)

843 **S2 Table.** Checklist for Artificial Intelligence in Medical Imaging (CLAIM)

844 **S1 Figure.** STARD 2015 patient flow diagram for PanCanAID

845 **S1 Appendix.** The phone interview form for collecting survival and clinical data

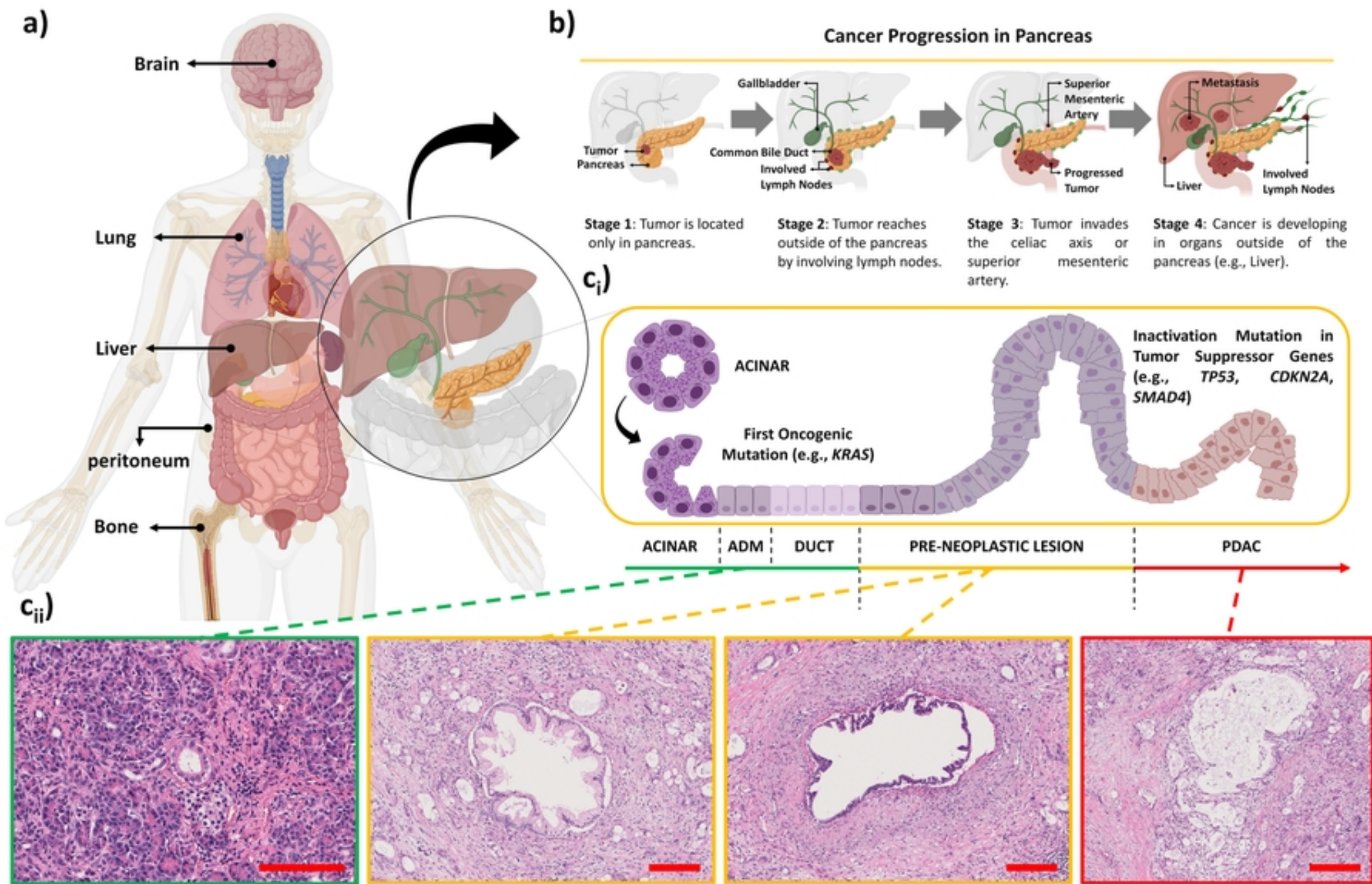


Fig 1

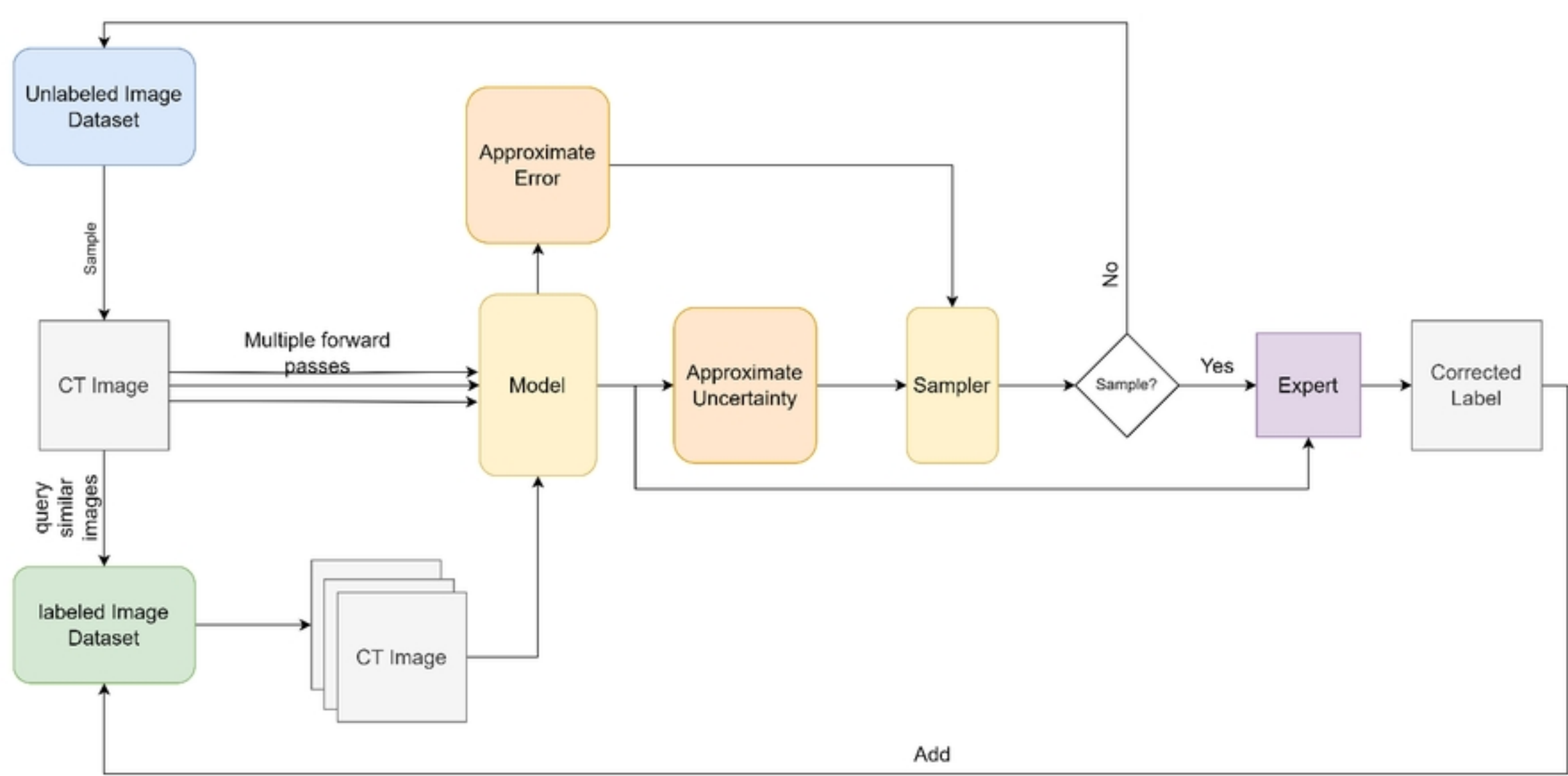


Fig 3

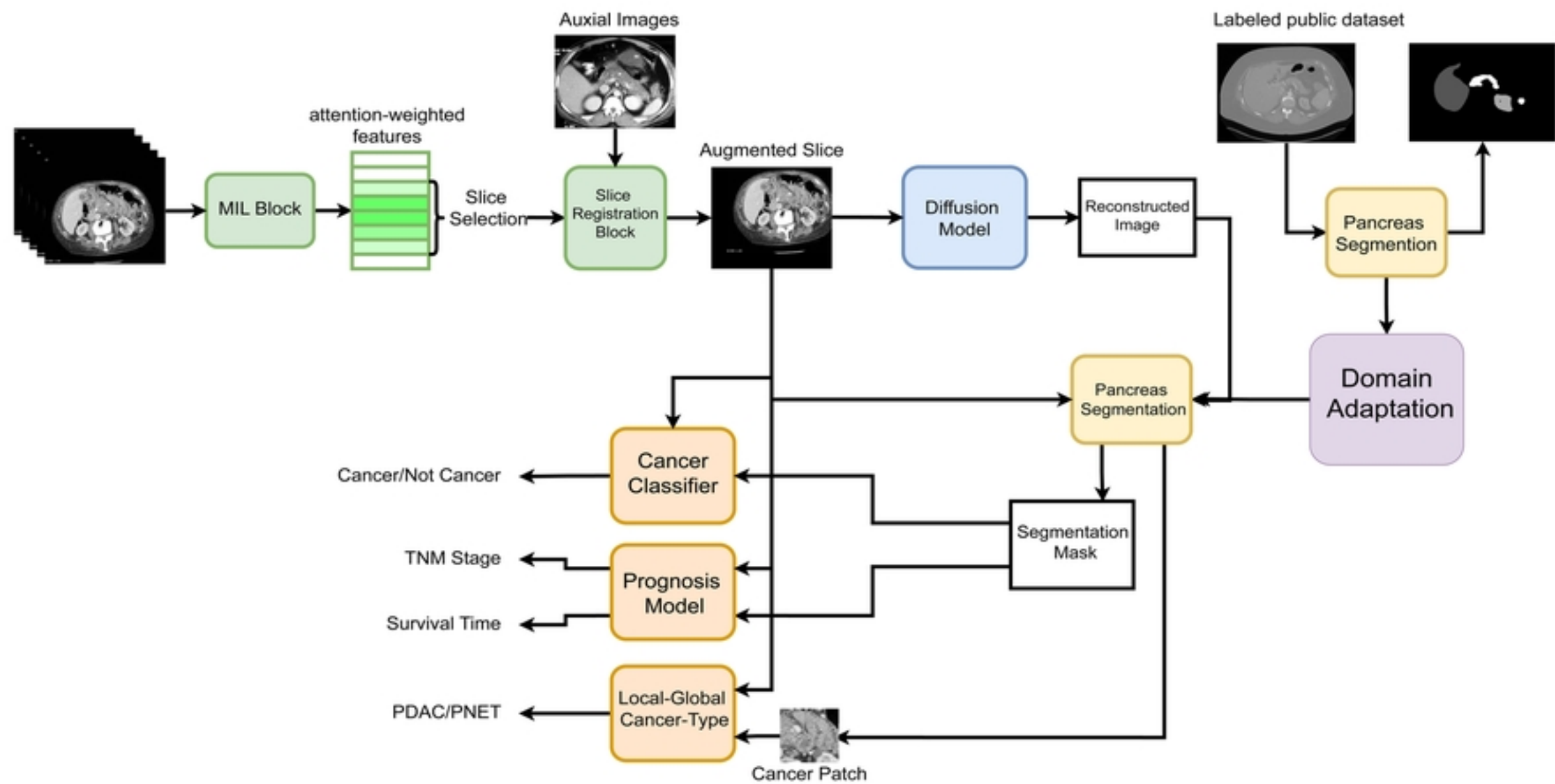


Fig 4

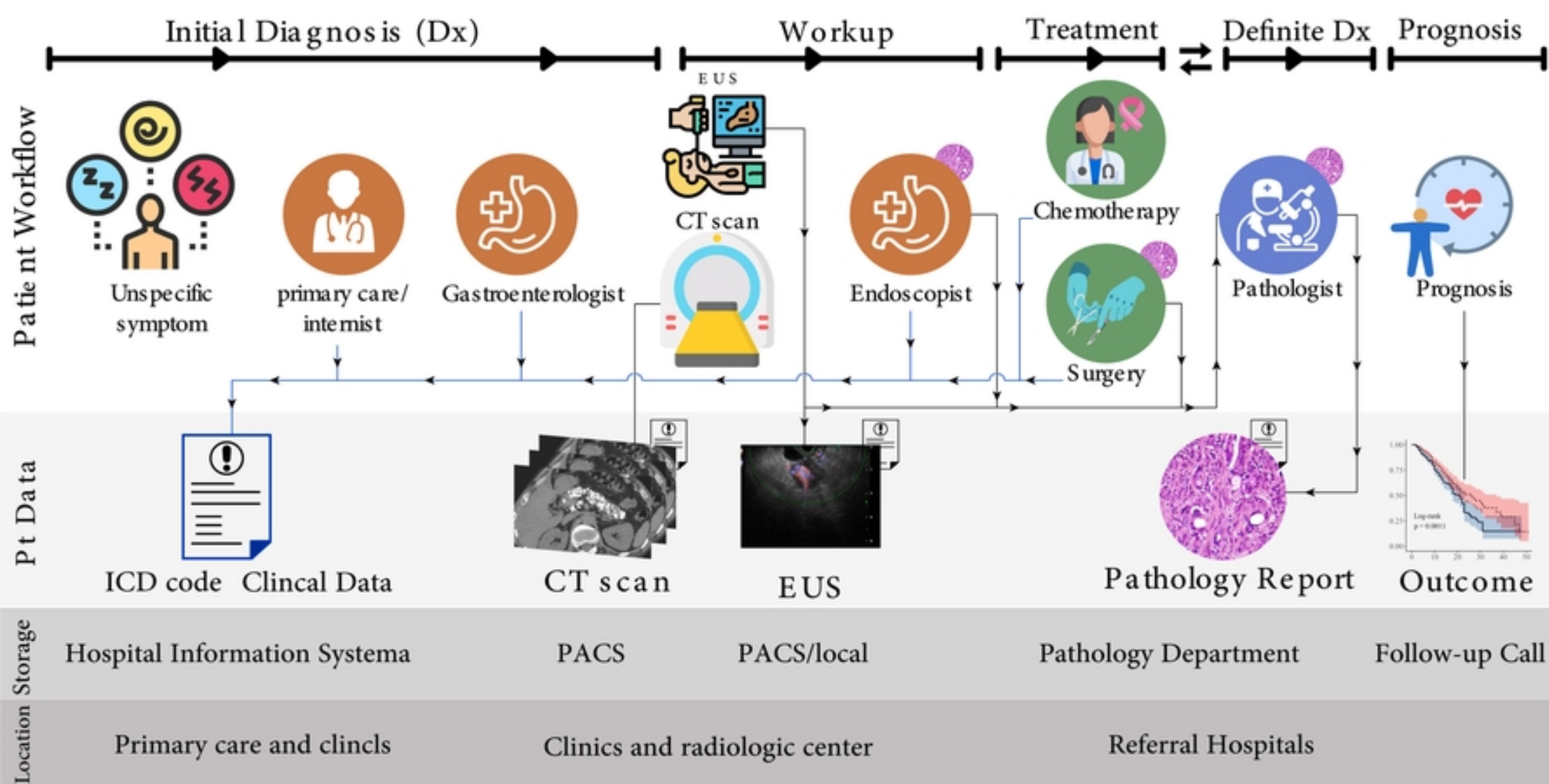


Fig 5

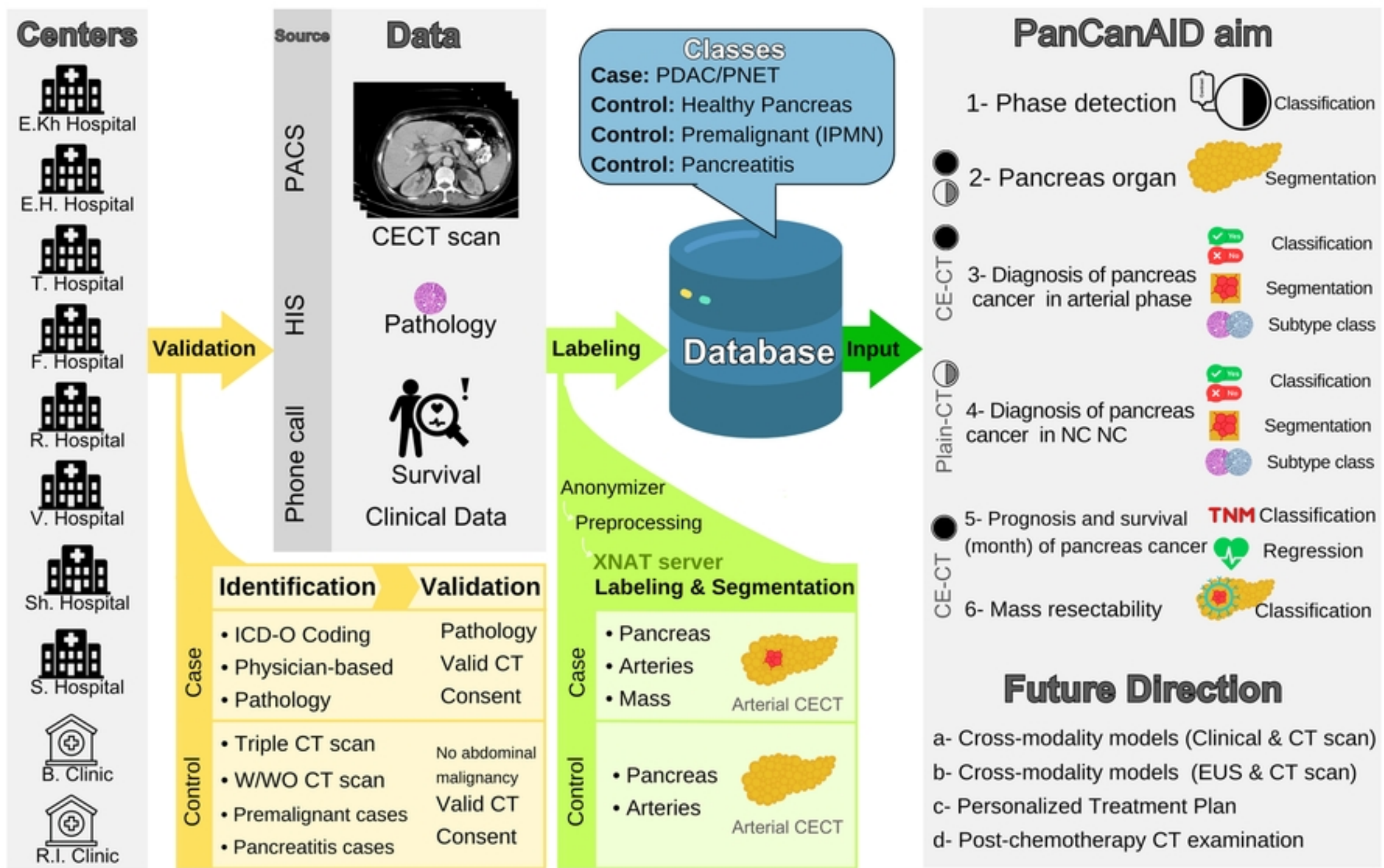


Fig 2