UNIVERSITA' VITA-SALUTE SAN RAFFAELE

CORSO DI DOTTORATO DI RICERCA INTERNAZIONALE IN MEDICINA MOLECOLARE

CURRICULUM IN MEDICINA CLINICA E SPERIMENTALE

NON INVASIVE EARLY PREDICTION OF TUMOUR BIOLOGICAL BEHAVIOUR, AGGRESSIVENESS AND TREATMENT RESPONSE: A RADIOMIC TRANSLATIONAL APPROACH FOR IDENTIFICATION OF NOVEL IMAGING BIOMARKERS IN PANCREATIC/OESOPHAGEAL NEOPLASMS

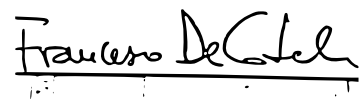DoS: Prof. Francesco De Cobelli

Second Supervisor: Prof. Andrea Laghi

Tesi di DOTTORATO di RICERCA di Palumbo Diego

matr. 015650

Ciclo di dottorato XXXV

SSD MED/36

Anno Accademico 2021/2022

CONSULTAZIONE TESI DI DOTTORATO DI RICERCA

Il/la sottoscritto/*a*   Diego Palumbo

Matricola / *registration number*   015650

nato a/ *born at*   Crema (CR)

il/*on*   05/07/1989

autore della tesi di Dottorato di ricerca dal titolo / *author of the PhD Thesis titled*

NON INVASIVE EARLY PREDICTION OF TUMOUR BIOLOGICAL BEHAVIOUR, AGGRESSIVENESS AND TREATMENT RESPONSE: A RADIOMIC TRANSLATIONAL APPROACH FOR IDENTIFICATION OF NOVEL IMAGING BIOMARKERS IN PANCREATIC/OESOPHAGEAL NEOPLASMS

AUTORIZZA la Consultazione della tesi / AUTHORIZES *the public release of the thesis*

X NON AUTORIZZA la Consultazione della tesi per 12 mesi /DOES NOT AUTHORIZE *the public release of the thesis for* 12 *months*

a partire dalla data di conseguimento del titolo e precisamente */ from the PhD thesis date, specifically*

Dal / *from* 01/12/2022 Al / *to* 01/12/2023

Poiché /*because*:

□ l'intera ricerca o parti di essa sono potenzialmente soggette a brevettabilità/ *The whole project or part of it might be subject to patentability;*

X ci sono parti di tesi che sono già state sottoposte a un editore o sono in attesa di pubblicazione/ *Parts of the thesis have been or are being submitted to a publisher or are in press;*

□ la tesi è finanziata da enti esterni che vantano dei diritti su di esse e sulla loro pubblicazione/ *the thesis project is financed by external bodies that have rights over it and on its publication.*

E' fatto divieto di riprodurre, in tutto o in parte, quanto in essa contenuto / *Copyright the contents of the thesis in whole or in part is forbidden*

Data /Date 21/11/2022 Firma /Signature

DECLARATION

This thesis has been: *i*) composed by myself and has not been used in any previous application for a degree; *ii*) throughout the text I use both 'I' and 'We' interchangeably; *iii*) it has been written according to the editing guidelines approved by the University.

Permission to use images and other material covered by copyright has been sought and obtained.

All the results presented here were obtained by myself, except for:

- Samples from peripheral and portal vein blood, as well as from neoplasm cystic fluid of IPMN patients (chapter "**Prediction of the characteristics of aggressiveness of pancreatic neuroendocrine neoplasms based on CT radiomic features**"), which have been systematically collected during surgery by experienced pancreatic surgeons (with overall coordination of prof. Stefano Crippa, Pancreas Translational and Clinical Research Centre, San Raffaele Scientific Institute, Milan, Italy). Levels of cytokines/chemokines, apolipoproteins, adipokines, hormones, and other biochemical parameters related to inflammation and lipid metabolism were evaluated from the same samples by Dott. Benedetta Ferrara (Diabetes Research Institute, San Raffaele Scientific Institute, Milan, Italy).

The following chapter contains data that have been already published:

- "**Prediction of early distant recurrence in upfront resectable pancreatic adenocarcinoma: a multidisciplinary, machine learning-based approach**" Palumbo et al. 2021, DOI: https://doi.org/10.3390/cancers13194938.

Abstract, Introduction as well as *Materials and Methods* sections from chapters "**Prediction of the characteristics of aggressiveness of pancreatic neuroendocrine neoplasms based on CT radiomic features**" and "**Dynamic change of radiomic**

features within tumour environment predicts pathological response to neoadjuvant chemotherapy and disease relapse of pancreatic adenocarcinoma" also contain data that have been already published (Palumbo et al. 2021, DOI: https://doi.org/10.3390/cancers13194938).

A paragraph from *Introduction* as well as Discussion section from chapter "**Early assessment of pathological response to neoadjuvant chemoradiotherapy for oesophageal cancer using fully hybrid PET/MR: results from interim analysis**" contain material from Palumbo D. *et al*, 2020 "Imaging in evaluation of response to neo-adjuvant treatment. *Ann Esophagus*".

Statistical analysis paragraph from chapter "**Identification and biological significance of a novel radiological biomarker for IPMN with *high-grade* dysplasia/invasive carcinoma**" contains methodology already used in **Palumbo *et al*, 2021 DOI: 10.1007/s00330-021-07788-y.**

All sources of information are acknowledged by means of reference.

ABSTRACT

Aims: *i*) To non-invasively stratify pancreatic neoplasms based on their biological behaviour and prognosis, and *ii*) to identify imaging markers for early assessment of neoadjuvant treatment response for PDAC and oesophageal cancers.

Methods: Five observational studies were designed. Preliminarily, a common radiomic workflow was defined using *i*) minimum redundancy, *ii*) robustness against delineation uncertainty and *iii*) a machine learning bootstrap-based method. #1 – A retrospective set of PDAC patients (pts) who underwent upfront surgery was enrolled to develop a preoperative model to stratify the chance of early (<11 months) distant disease relapse. #2 – A prospective set of IPMN pts candidate for resection was enrolled to identify radiological marker(s) to distinguish *low-* vs. *high-risk* IPMN. Besides high-risk stigmata/worrisome features, pancreatic fat content was assessed. #3 – A retrospective set of resected panNEN pts was enrolled to train and validate radiomic models to predict pathological characteristics of aggressiveness. #4 – A retrospective set of PDAC pts who underwent surgery after nCT was enrolled to develop three clusters of models based on *a*) pre-/ *b*) post-nCT and *c*) Δ radiomics to predict disease recurrence after surgery, N2 and pathological response to nCT. #5 – A prospective set of oesophageal cancer pts scheduled to nCRT was enrolled. PET/MR was performed prior to, during and after treatment; for each patient, ERI was computed. TRG=1 stands for complete response.

Results: Models comprising radiomic features and clinicoradiological variables were effective in stratifying the risk of early distant relapse after upfront surgery for PDAC, as well as pathological characteristics of panNEN aggressiveness. Δ radiomics well stratifies the risk of relapse after nCT for PDAC. Non-radiomic findings were also intriguing. Pancreatic fat content was higher in IPMN pts with *high-grade* dysplasia/invasive carcinoma than in *low/moderate-grade* dysplasia (cohort#2). When dealing with cohort#5, pts with TRG=1 had significantly lower ERI values than those with TRG≥2; they also had a significant higher median increase in tumour ADC from baseline to intermediate scans.

Conclusions: Our findings could complement current diagnostic workflow and improve clinical decision-making for pts with pancreatic/esophageal neoplasms.

TABLE OF CONTENTS

- *Materials and Methods*
- *Results*
- *Discussion*

- *Materials and Methods*
- *Results*
- *Discussion*

- *Materials and Methods*
- *Results*
- *Discussion*

- *Materials and Methods*
- *Results*
- *Discussion*

- *Materials and Methods*

- *Results*

- *Discussion*

ACRONYMS AND ABBREVIATIONS

AIOM = Associazione Italiana Oncologia Medica; **AUC** = Area Under the Curve; **CI** = Confidence Interval; **CT** = Computed Tomography; **EDR** = Early Distant Recurrence; **ERI** = Early Regression Index; **G** = Tumour Grade; **HR** = Hazard Ratio; **HRS** = High Risk Stigmata; **HU** = Hounsfield Unit; **ICC** = Intraclass Coefficient; **IPMN** = Intraductal Papillary Mucinous Neoplasm; **IQR** = Interquartile Range; **M+** = Metastases; **MCP-1** = Monocyte Chemoattractant Protein-1; **MPD** = Main Pancreatic Duct; **MRCP** = Magnetic Resonance Cholangiopancreatography; **N+** = metastatic Lymphnodes; **NCCN** = National Comprehensive Cancer Network; **nCT** = neoadjuvant Chemotherapy; **nCRT** = neoadjuvant Chemoradiotherapy; **NPV** = Negative Predictive Value; **PanNEN** = pancreatic Neuroendocrine Neoplasm; **PD** = pancreaticoduodenectomy; **PDAC** = Pancreatic Ductal Adenocarcinoma; **PET** = Positron Emission Tomography; **PPV** = Positive Predictive Value; **RECIST** = Response Evaluation Criteria In Solid Tumours; **ROI** = Region of Interest; **RFs** = Radiomic Features; **SFA** = Subcutaneous Fat Area; **TAMA** = Total Abdominal Muscle Area; **TRG** = Tumour Regression Grade; **VFA** = Visceral Fat Area; **VI** = Microvascular Invasion; **WF** = Worrisome Features.

TABLES AND FIGURES

TABLE #1: Patients' characteristics (cohort #1). **Palumbo *et al*, DOI: 10.3390/cancers13194938**

TABLE #2: Overall performance in both training and validation cohorts of radiomic, clinicoradiological and combined models quantified in terms of area under the ROC curve (AUC), positive and negative predictive values, specificity and sensitivity. **Palumbo *et al*, DOI: 10.3390/cancers13194938**

TABLE #3: High Risk stigmata/Worrisome Features.

TABLE #4: Patients' characteristics (cohort #3).

TABLE #5: Overall performance of pure radiomic and clinicoradiological/radiomic models.

TABLE #6: Overall performance of pre nCT radiomic models.

FIGURE #10: Kaplan-Meier survival curve stratifying the chance of disease relapse after surgery based upon the prognostic index resulting from pre nCT radiomic model.

FIGURE #11: Box plot illustrating N2 prediction model (pre nCT radiomic model), and corresponding ROC curve.

FIGURE #12: Box plot illustrating TRG prediction model (pre nCT radiomic model), and corresponding ROC curve.

FIGURE #13: Kaplan-Meier survival curve stratifying the chance of disease relapse after surgery based upon the prognostic index resulting from post nCT radiomic model.

FIGURE #14: Box plot illustrating N2 prediction model (post nCT radiomic model), and corresponding ROC curve.

FIGURE #15: Box plot illustrating TRG prediction model (post nCT radiomic model), and corresponding ROC curve.

FIGURE #16: Kaplan-Meier survival curve stratifying the chance of disease relapse after surgery based upon the prognostic index resulting from Δ radiomic model.

FIGURE #17: Box plot illustrating N2 prediction model (Δ radiomic model), and corresponding ROC curve.

FIGURE #18: Box plot illustrating TRG prediction model (Δ radiomic model), and corresponding ROC curve.

FIGURE #19: Study flowchart (cohort #5).

INTRODUCTION

A significant proportion of the overall improvement in cancer outcomes in the recent years (Arnold *et al*, 2019; The Lancet, 2022) is attributable to earlier diagnosis: combining current breakthrough, personalized therapies with an effective early detection would have an unmatched impact on patients' survival and quality of life (Whitaker, 2020; Koo *et al*, 2020). Early detection has already been acknowledged as having a key role in the management of cervical (Sawaya *et al*, 2019) and breast (Welch *et al*, 2016) cancer, and is likely to become more important in preventing cancer related deaths, although with current inhomogeneous results and non standardized protocols, in the control of colorectal (Bretthauer *et al*, 2022), prostate (Eklund *et al*, 2021) and lung (National Lung Screening Trial Research Team *et al*, 2011) cancer by the implementation of dedicated screening programs and awareness raising campaigns.

However, some tumour types do not benefit yet from this approach: pancreatic cancer (having no recommended screening protocols for patients with no increased risk [lifetime risk of pancreatic cancer < 5%]) is usually diagnosed at late stages due to the lack of specific clinical manifestations, when disease has already spread, resulting in poor prognosis (Pereira *et al*, 2020; Mizrahi *et al*, 2020; Overbeek *et al*, 2022). Actually, even when considering highly selected *high-risk* patients undergoing strict surveillance with endoscopic ultrasonography and/or magnetic resonance, Overbeek *and colleagues* (Overbeek *et al*, 2022) found that almost half of the cancerous lesions are usually identified a median of 11 months after an unremarkable imaging; furthermore, more than 70% of those neoplastic lesions were > T1 N0, highlighting how narrow the temporal window of possible intervention is. Mukherjee *and colleagues* recently proposed instead a different approach, based upon a machine learning model able to detect pancreatic cancer on prediagnostic CT scans, beyond conventional human interrogation capability (Mukherjee *et al*, 2022); however, major concerns, mainly ethical and methodological, currently undermines this approach practicability.

Moreover, irrespective of clinical stage, each pancreatic neoplasm subtype has itself a wide prognostic range corresponding to intra tumoural biological heterogeneity (Mizrahi *et al*, 2020). It follows that, in the specific setting of pancreatic cancer, early morphological depiction is crucial, but proper assessment of the actual malignant potential is nothing but necessary.

On that note, a significant example consists of the definition of resectable pancreatic
adenocarcinoma (PDAC), which is indeed a highly debated issue. The different
descriptions proposed over the years are mainly based on the extent of vascular
involvement by the tumour (Zins *et al*, 2018), which is thought to be the most important
factor possibly undermining technical feasibility of resection. According to the 2021
NCCN (*National Comprehensive Cancer Network*) guidelines (Tempero *et al*, 2021),
resectability status should be determined by a multidisciplinary team (comprising
dedicated radiologists, surgeons and oncologists) that discusses findings on contrast
enhanced CT scan and determines if the tumour is *i)* resectable, *ii)* borderline resectable,
iii) locally advanced/unresectable or *iv)* metastatic pancreatic adenocarcinoma.
Obviously, different *resectability status* reflects different scheduled approach and
prognosis (Tempero *et al*, 2021). However, despite careful selection, approximately
40% of patients undergoing upfront surgery are found to experience distant disease
recurrence within 12 months from the index procedure (Petrelli *et al*, 2017), resulting in
poor prognosis (Petrelli *et al*, 2017; Matsumoto *et al*, 2015). These data suggest that
upfront surgery is not the best treatment approach for the vast majority of those patients
currently being claimed as primary resectable, which could instead benefit from
neoadjuvant chemotherapy (Lee *et al*, 2019; Barugola *et al*, 2009). There is indeed an
urgent, unmet need to expand the concept of what is a pancreatic resectable tumour;
along with anatomical definition criteria, some other clinical, pathological and
biological features may help in identifying patients who would not benefit from upfront
surgery, even when a radiological locally resectable disease is present.

Other examples illustrating the importance of biological characterization over
morphological depiction come from intraductal papillary mucinous neoplasms (IPMNs)
and pancreatic neuroendocrine neoplasms (PanNENs).

The former entity (IPMN), through a wide spectrum of dysplastic changes, is a well-
known precursor of PDAC but with no clear established progression-timing schedule
(Crippa *et al*, 2016; Makohon-Moore *et al*, 2018). Given that the detection of IPMN in
the early stage of carcinogenesis could provide a unique opportunity to perform curative

surgical treatment before full progression to invasive carcinoma, proper characterization of the malignant potential of each pancreatic IPMN is fundamental. Current guidelines identify high-risk stigmata and worrisome features as indications for surgical resection of IPMN (Tanaka *et al*, 2017). However, no reliable malignancy biomarker for IPMN with *high-grade* dysplasia is known (Crippa *et al*, 2022), and the overtreatment (surgery, with high morbidity [55.6% in a recent series published by our group]/mortality rate [> 2%] even in high-experienced centres (Aleotti *et al*, 2022; Capretti *et al*, 2018)) of benign IPMN remains a critical clinical issue, as well as under treatment (continuing follow-up in case of high-grade dysplasia/invasive carcinoma).

Pancreatic neuroendocrine neoplasms (PanNENs) comprise instead a wide range of entities too, both histologically and macroscopically (Partelli *et al*, 2017b; Klimstra, 2016). In this respect, histology is the only validated methodology currently available that allows tumor grading, distinguishing low grade (G1) from more aggressive (G2/G3) neoplasms, mainly based on Ki67 proliferative index, and to assess other biological characteristics important for prognostic assessment and treatment planning (follow-up *vs*. surgery, surgery *vs*. neoadjuvant treatment) (Pasaoglu *et al*, 2015). However, endoscopic ultrasound-guided fine-needle aspiration has limited accuracy in determining pathological characteristics (Rebours *et al*, 2015), so that reliable histology may be obtained only after surgery. On the other hand, preoperative macroscopic imaging features have been associated with the outcome of PanNENs, specifically size of the lesion > 2 cm (Partelli *et al*, 2017a), enhancement kinetics (Yano *et al*, 2017), presence of necrosis (Mapelli *et al*, 2022b), nodal involvement (Partelli *et al*, 2022) or distant metastases, but with inhomogeneous results. With regard to this last point, as pointed out by Partelli *and colleagues* in the setting of a multidisciplinary prospective trial, subjective assessment of nodal status in panNENs has poor sensitivity (26%, when dealing with CT) despite good (95%) specificity (Partelli *et al*, 2022). On the opposite, our group recently took a step forward a novel direction suggesting an intriguing explanation connecting, in non-functioning panNENs, microvessel density, radiological appearance in terms of HU values and biological behavior: in short, low microvessel density (assessed by CD34[+] staining), corresponding to ipoenhancement in arterial phase, has been found to be associated with pathological features of aggressiveness (Battistella *et al*, 2022). However, there is still a long way to go before validating these

preliminary results.

Thus, rationale exists for identifying novel tools able to provide non-invasive stratification of pancreatic neoplasms based on their biological behaviour and prognosis.

Furthermore, another matter has to be taken into account when dealing with pancreatic cancer: even when a relatively early diagnosis is achieved, surgical treatment implies high mortality (up to 5% for pancreaticoduodenectomy even in highly-experienced centres (Capretti *et al*, 2018)) and morbidity rates (Palumbo *et al*, 2021a), with inhomogeneous and, unfortunately, unpredictable oncologic outcomes (Petrelli *et al*, 2017). There is indeed an urgent need for identifying those patients who could actually benefit, in terms of disease free/overall survival and, hopefully, quality of life, from a surgical approach. On that note, it is worth noting that growing literature suggests that even early stage PDAC could benefit from a neoadjuvant treatment (Lee *et al*, 2019; Reni *et al,* 2018a; Reni *et al*, 2018b); the rationale of such an approach lays in the observation that recurrence occurs in almost 90% of these patients within two years after surgery, regardless of deployment of adjuvant chemotherapy (which instead lengthens the overall survival (Neoptolemos *et al*, 2017; Conroy *et al*, 2018)), implying that even early stage disease is, actually, a micro metastatic one. It follows that since most patients have (micro) metastatic disease at the time of diagnosis, primary surgery, only debulking the overall tumour cell burden, is less effective in providing disease control when compared to chemotherapy, which instead potentially prevents both local and distant tumour progression by reducing the number of neoplastic clones in their exponential growth stage (Haeno *et al*, 2012). Both AIOM (*Associazione Italiana Oncologia Medica*) and NCCN guidelines (Tempero *et al*, 2021) now include neoadjuvant chemotherapy (nCT) as a possible treatment option for resectable disease. However, several authors have recently claimed that neoadjuvant chemotherapy might impair the diagnostic accuracy of CT in the evaluation of tumour response mostly due to abundant desmoplastic reaction mimicking cancer growth (Cassinotto *et al*, 2014; Park *et al*, 2021). There are indeed no reliable criteria to assess treatment response of PDACs after nCT, ultimately undermining the chance for an optimal therapeutic path.

Another side of the same issue (need for treatment response evaluation) heavily burden another cancer type: oesophageal neoplasms.

The CROSS trial (van Hagen *et al*, 2012; Shapiro *et al*, 2015) demonstrated that
preoperative administration of neoadjuvant chemoradiotherapy (nCRT) doubles the
median overall survival of locally advanced oesophageal neoplasms in comparison to
surgery alone. In the cohort of patients enrolled in this multicentre randomized clinical
trial, the reported percentage of patients with a pathological complete response was
29%; many authors suggest that this subgroup of patients do not benefit from additional
surgery, also having regard to the fact that oesophageal resections are associated with
substantial morbidity and postoperative mortality rates (D'Journo *et al*, 2021). On the
other hand 18% of patients who underwent nCRT were deemed as being non-
responders; these patients do not benefit from nCRT and only suffer from its side
effects. Early identification of tumour response/non response to nCRT could therefore
enable tailored therapeutic plans, avoiding unnecessary treatment efforts and related
adverse effects, with major impact on patients' quality of life as well as health care
costs (Palumbo *et al*, 2020).

Thus, rationale also exists for identifying novel imaging biomarkers able to improve
non-invasive early identification of tumour response.

How can these two objectives be achieved? Histological characterization from
bioptic samples represents current gold standard in clinical practice; however, bioptic
procedures have some limitations, the main ones being their invasiveness (which limits
their feasibility) and, above all, the fact that they cannot provide a reliable depiction of
the entire tumour heterogeneity. Solutions exist. Radiomics is a novel tool consisting in
extraction of quantitative data from medical images in order to develop predictive
models relating imaging features to clinical outcomes (Lambin *et al*, 2017; Gillies *et al*,
2016; Palumbo *et al*, 2020); interestingly, given that radiomic features are derived from
volumes (rather than from samples), they are paradoxically more representative of the
heterogeneity and characteristics of the entire lesion than histology itself. However,
there are many challenges to be faced regarding translation of radiomics into clinical
practice (Pinto Dos Santos *et al*, 2021; Spadarella *et al*, 2022); radiomics is a tricky
multistep process, and within each step there are issues to exceed to ensure robustness

of model's findings, while reproducibility and deployment are often impaired, as pointed out by Mori, Palumbo *and colleagues* in a recent review article questioning whether a role exists for radiomics in the setting of gastroesophageal junction adenocarcinoma (Mori *et al*, 2022). It follows that the implementation of a methodologically robust radiomic workflow should be the unavoidable basis for any further research. Moreover, combining few highly selected, highly robust radiomic features together with highly relevant clinical/radiological variables could be a game-changing approach to fill the current translational gap undermining radiomics deployment. In this respect, our group has already published a number of manuscripts dealing with such methodological issues (*a*) reliance upon diverse imaging parameters (Mapelli *et al*, 2022), *b*) delineation uncertainty (Loi *et al*, 2020; Mori *et al*, 2019), *c*) intra/interscanner variability and *d*) need for clinical interpretation of selected radiomic features (Palumbo *et al*, 2021b)) and proposed a standardized, multidisciplinary workflow for features extraction, model development and clinical/biological interpretation of final results (Palumbo *et al*, 2021b).

That said, another crucial point to be taken into account when looking for novel imaging biomarkers for *i*) prognosis stratification and/or *ii*) treatment response early assessment is a proper characterization of tumour microenvironment, which could represent an outstanding source of information for identifying carcinogenesis promoters as well as mechanisms of treatment resistance (Ho *et al*, 2020). On that note, imaging has unmatched advantages (over bioptic samples) to achieve this aim, since it allows simultaneous assessment of the index lesion together with its surrounding microenvironment as well as patients' body composition, which account for a sort of a surrogate for metabolic compartments. Now, when referring to this last point our recent efforts were headed towards possible relationships between radiologically measured patients' body composition and the occurrence of postoperative complications (Pecorelli *et al*, 2022; Guarneri *et al*, 2022); we are now shifting our focus on oncological outcomes, as will be explained in the next chapters of the present manuscript.

Finally, a last challenge to be faced to answer both our research questions lies in the attempt to highlight the importance of dynamical change of the neoplastic environment in course of treatment over its stationary consideration (before or after the same

treatment); there is indeed an urgent need for adjusting conventional "single temporal frame" imaging to a more comprehensive one, in order to move alongside with currently available novel treatment possibilities.

Accordingly, here we present five chapters, corresponding to five separate (yet firmly complemented) studies, thought to be embedded into a more comprehensive oncological framework in order to properly answer, from five different perspectives, to the abovementioned research questions.

AIM OF THE WORK

Characterizing, non-invasively, biological aggressiveness of pancreatic/oesophageal neoplasms constitutes the major core of the present project; these two tumour types well illustrate, indeed, the importance of biological characterization over morphological depiction. Such major aim translates into two specific research questions:

- How to non-invasively stratify pancreatic neoplasms (pancreatic adenocarcinoma, intraductal papillary mucinous neoplasm and pancreatic neuroendocrine neoplasm) based on their biological behaviour and prognosis? This first question stems from the observation that, irrespective of clinical stage, each pancreatic neoplasm subtype has itself a wide prognostic range corresponding to intra tumoural biological heterogeneity. Such assumption heavily struggles against currently used staging systems, mostly relying on macroscopic imaging findings and ultimately hampered by biological inconsistency. On that note, characterization of tumour microenvironment could be a game changer.

- How to identify imaging biomarkers for early assessment of neoadjuvant treatment response for pancreatic adenocarcinoma and oesophageal cancer? This second question comes instead from the need for adjusting conventional "single temporal frame" imaging to a more comprehensive one, in order to move alongside with currently available novel treatment possibilities. To accomplish this specific aim it would be necessary to highlight the importance of dynamical change of the neoplastic environment in course of treatment over its stationary consideration (before or after the same treatment).

Properly addressing both these questions could have significant clinical impact, since it could provide substantial aid in complementing current diagnostic workflow, ultimately improving patients' therapeutic path.

Prediction of early distant recurrence in upfront resectable pancreatic adenocarcinoma: a multidisciplinary, machine learning-based approach.

Materials and Methods

Patients' cohort

This is a single-center retrospective study conducted at San Raffaele Scientific Institute (Milan, Italy); data were collected within the context of an Ethics Committee approved study (28/INT/2015) in patients who had signed an institutional procedure specific informed consent. From a prospectively acquired database, all consecutive patients with pancreatic adenocarcinoma who underwent upfront pancreaticoduodenectomy (PD) between January 2015 and December 2019 were identified (n = 652); within this database, patients who were evaluated with at least one multiphase, contrast-enhanced CT scan within 30 days before index surgery (n = 156) were enrolled into our study. Patients who died within 90 days after index surgery (n = 7) were excluded from further analysis; moreover, two patients had no sufficient follow-up information and were also excluded. The resulting population (n = 147) was then randomly split into training (n = 94) and validation (n = 53) cohorts according to the second level of the TRIPOD guidelines for the validation of predictive models in oncology (Moons *et al*, 2015). A detailed flowchart of this study design (comprehensive of inclusion and exclusion criteria) is shown in Figure 1.

Patients were finally divided into an early distant recurrence (EDR) group (disease free survival < 11 months) and a non-EDR group (disease free survival ≥ 11 months); the cut-off was in agreement with the median time to distant relapse observed in our cohort (11 months [IQR: 8 – 15.7]).

Surgical technique, pathology protocol, adjuvant therapy and follow-up data collection

A multidisciplinary team comprising radiologists, surgeons and oncologists

evaluated the included patients and had deemed all of them as *upfront resectable* according to the 2021 NCCN (*National Comprehensive Cancer Network*) guidelines (Tempero *et al*, 2021).

Both pylorus preserving and Whipple PDs were performed by six surgeons with at least 10 years of experience in pancreatic surgery. All patients were treated according to the principles of the *Enhanced Recovery After Surgery* (Braga *et al*, 2014).

After resection, pathologic tumor stage (according to the 8th edition of the *American Joint Committee on Cancer staging system* (van Roessell *et al*, 2018)), and disease grade were assessed. Perineural invasion was systematically described as present/absent and further classified according to the calibre and number of nerve trunks involved; lymphovascular invasion was also described. The number of metastatic lymph nodes and the ratio of positive to harvested lymph nodes were recorded.

Adjuvant treatment was always considered when sufficient recovery within 12 weeks after resection was achieved. All the patients were monitored every three months, until death, via outpatient clinic visits, which included imaging studies and laboratory examinations. Once a follow-up imaging study showed the emergence of any distant lesion, the recurrence was confirmed.

Clinical Variables

Retrospective chart review was used to obtain information on demographics (gender, age, eventual comorbidities), duration of symptoms, laboratory findings and eventual use of adjuvant chemotherapy. Of note, in order to lower possible confounding factors (Ballehaninna & Chamberlain, 2012), CA 19.9 serum levels were recorded, as a continuous variable, after eventual endoscopic/angiographic palliation.

Radiological Variables and Radiomic Features

In patients who underwent multiple preoperative CT scan, the last examination closest to the date of surgery was used for review.

CT protocol – All CT examinations were performed on multidetector CT scanners (scanner 1: SOMATOM Definition Flash Dual Source CT, Siemens Healthcare; scanner 2: BRILLIANCE, Philips medical system). CT protocol included administration of intravenous iodine contrast medium (Iopromide, Ultravist 370 mg iodine/ml (Bayer

HealthCare), 120 ml at a rate of 4 ml/s) and consisted of a multiphase acquisition (unenhanced, late arterial, portal venous and late axial scans of the abdomen).

Conventional Image based parameters – CT findings were selected for analysis by two radiologists and two senior consultants pancreatic surgeons on the basis of their clinical experience; variables previously described in the literature were also considered (including those proposed by the *Society of Abdominal Radiology* and the *American Pancreatic Association* in their dedicated reporting template (Al-Hawary *et al*, 2014)). Readers with different experiences in abdominal CT imaging were selected for image review: specifically, two residents in their last year of training and one radiologist with 10 years experience and a specific commitment in pancreatic imaging. They independently analysed all CT images, blinded to any pathological information. After image review completion, a consensus was established for each selected categorical CT finding.

Lesion delineation on CT images – The robustness of CT radiomic features (RFs) against interobserver contouring variability was preliminarily assessed on a subgroup of 29 patients by the same three readers. Then, two of these three reviewers contoured all tumour volumes on late arterial phase CT images, where tumour conspicuity was the most. Contours were transferred from the late arterial to the unenhanced images by means of a rigid registration, and then manually adjusted on the latter to correct minor anatomical discrepancies due to organ motion. Contouring was performed using the *MIM Software* (v. 6.8.2).

Radiomic features extraction – SPAARC Pipeline for Automated Analysis and Radiomics Computing complying with the Image Biomarker Standardization Initiative (IBSI) (Zwanenburg *et al*, 2020) was used to process images for RF extraction. All images were resampled at 1 mm cubic voxels with a bilinear interpolation. This procedure was implemented to reduce directional bias when voxel sizes were not already isotropic to allow comparison between image data from different samples, cohorts or batches. This is essential to compare final results because many RFs are based on the sum of the entire number of voxels in the lesion. Image rebinning was also

16

necessary, not only to speed up the process of RF extraction, but also to limit noise: we chose 64 bins, as reported in literature (Benedetti *et al*, 2021).

Subsequently, adjusted DICOM files were imported to MATLAB using the Computational Environment for Radiological Research. 182 RFs of first and higher order were extracted, belonging to the following families: Morphology, Statistical, Intensity Histogram, Grey Level Co-occurrence Matrix 3D_average (GLCM3D_avg), Grey Level Co-occurrence Matrix 3D_combined (GLCM3D_comb), Grey Level Run Length 3D_average (GLRL3D_avg), Grey Level Run Length 3D_combined (GLRL3D_comb), Grey Level Size Zone Matrix 3D, Neighbor Grey Tone Difference Matrix 3D (NGTDM3D), Grey Level Distance Zone Matrix 3D (GLDZM3D). Figure 2 summarizes the radiomic workflow.

Statistical analysis

As previously stated, the original population was randomly split into training (n = 94) and validation (n = 53) cohorts. According to the primary endpoint (EDR evaluated at 11 months), 25 "events" were recorded in the training cohort, making feasible to preferably include a maximum of three variables in the resulting multivariable models (Peduzzi *et al*, 1996).

Variables redundancy elimination – Since the large number of variables (clinical, pure radiologic and radiomic [n = 182]) considered, exceeding the number of patients, many variables were expected to be redundant, especially RFs owning to the same family. To limit the risk of redundancy, we applied a correlation-based filter: starting from the correlation matrix, a Spearman coefficient (S) threshold equal to 0.70 was arbitrarily fixed to select redundant (S > 0.70) and independent features (S < 0.70). Variables found to be independent were selected; differently, among the redundant variables, the ones with the best p values in univariate logistic regression were selected for further analysis (one for each group of correlated features).

Inter reader agreement – The robustness of CT RF against interobserver contouring variability was assessed using intraclass correlation coefficient (ICC); ICCs higher than 0.80 were considered to be in high agreement. RFs demonstrating an ICC < 0.80 were excluded from further analysis.

Multivariable model development – In order to assess the best combination of the previously selected clinical, radiologic and radiomic variables to predict the primary endpoint, a machine learning bootstrap-based method was used. Briefly, the training set, constituted by the original sample of variables selected, was bootstrapped 1000 times and a backward univariate logistic regression was run for each sample. The variables most significantly associated with the endpoint occurring in each sample were collected and ranked according to their frequency. Accordingly, three models were developed: a strictly radiomic model, a clinicoradiological model, and a combined model. For each model, the most frequent variables resulting from the bootstrap ranking procedure (taking those variables with p value < 0.05 in more than 500 cases on the 1000 bootstrapped samples) were included in a backward multivariable logistic regression for the prediction of EDR. Finally, a maximum number of variables to be retained equal to three, based on the endpoint events' number, was fixed.

A prognostic index (*P index*) was derived for each model according to the following logistic regression formula in the training cohort and then tested in the validation group:

$$\text{P index} = \frac{1}{1+\exp-(\sum(B_0*+B_i*Xi))}$$

Specifically, for each model the B_i coefficients are given applying the logistic regression to the training cohort, whereas X_i represents the values of predictors from each patient.

Model performance assessment – To assess the ability of the P index in stratifying patients according to the risk of developing EDR, a cut off value was derived as the best criterion according to the maximum value of the Youden index of the corresponding ROC curve. The P index was then dichotomized as greater or smaller than the cut-off value; finally, the separation of the survival curves of the two groups was tested with a Kaplan-Meier test. The P index resulting from the training set was then tested in the validation cohort. The performances of the models were quantified in terms of: area under the ROC curve (AUC), positive and negative predictive values (PPV, NPV), specificity and sensitivity. Analyses were performed using homemade Matlab codes.

Results

Patients' characteristics

Patients' characteristics are summarized in Table 1. The median overall survival and progression-free survival were 20 (IQR: 15-28) and 15 (IQR: 10-22) months, respectively.

85 patients out of 147 (57.8%) had a disease recurrence during the follow-up time period (median follow-up time: 19 months, [IQR: 14 - 32]). Distant recurrence (n = 76, 51.7%) was the most frequent pattern of disease relapse. Considering distant recurrences alone, median time to relapse was 11 months (IQR: 8 – 15.7), which has been found, in agreement with previous literature (Petrelli *et al*, 2017; Yamamoto *et al*, 2014), to be a consistent threshold value for distinguishing early and late distant recurrences; accordingly, 39 out of 147 patients (26.5%) included in the final cohort had EDR.

Between training and validation cohorts, no significant differences were found in terms of EDR rate (26.5% *vs*. 26.3%, p = 0.54). No differences were observed when considering both clinical and radiological variables, nor pathological data, except for lymphovascular invasion (93.6% *vs*. 76.8%, p = 0.015). Adjuvant treatment was implemented in 73% (n = 69) and 73.2% (n = 39) of patients after surgery, respectively per cohort. No significant differences were found in terms of EDR rate between those patients who underwent adjuvant treatment and those who did not (25.7% *vs*. 30.2%, p = 0.197); however, a statistical trend (p = 0.068) in favour of adjuvant treatment exists when considering overall disease free survival (16 months [IQR: 10.5 – 27] *vs*. 14 months [IQR: 9 – 22]).

Between EDR and non-EDR groups no statistically significant differences exist in terms of R status (R0 *vs*. R1) (52.2% *vs*. 42.2%, p = 0.185) and lymphovascular invasion (94.8% *vs*. 88.5%, p = 0.346); EDR group patients had higher lymph node ratio (0.23±0.18 vs. 0.14±0.13, p = 0.003).

Variables Selection

Of the 182 radiomic features extracted, those with ICC values higher than 0.80 (89/182, 48.9%) were considered for analysis. After further selection to limit the risk of redundancy 23 variables were retained and tested. Thereafter, the machine-learning bootstrap-ranking procedure identified the 10 most frequent variables: eight radiomic features (3 morphologic, 4 texture related and 1 statistical features) and two clinicoradiological variables. The details of the selected features are shown in Figure 3.

Training and Validation of the Radiomic Model

Amongst the eight most frequent RFs resulting from machine-learning bootstrap-ranking procedure, only *Surface to Volume ratio* was retained in the final model (p = 0.0097, AUC = 0.59), with a strong inverse relation (coefficient: - 3.82) to the primary endpoint considered. After calculating the corresponding P index, the model was confirmed in the validation cohort (p = 0.0244, AUC = 0.73). Further details are provided in Table 2. Corresponding ROC curves and *Kaplan-Meier* survival curves based on the P index best threshold are shown in Figure 4. With regard to this last point, the radiomic model demonstrated an overall good performance in stratifying the risk of EDR after upfront surgery (training cohort: HR = 2.05, 95% CI = 1.03 – 4.09; validation cohort: HR = 2.84, 95% CI = 1.12 – 7.21).

Training and Validation of the Clinicoradiological Model

Both the most frequent clinicoradiological variables resulting from machine-learning bootstrap-ranking procedure (presence of tumour necrosis at preoperative CT imaging, and CA 19.9 serum levels) were retained in the final model (p = 0.0018, AUC = 0.72). However, after computation of the corresponding P index, the model was not confirmed in the validation cohort (p = 0.9529, AUC = 0.54). Further details are provided in Table 2. Corresponding ROC curves and *Kaplan-Meier* survival curves based on the P index best threshold are shown in Figure 5: the clinicoradiological model failed to predict EDR in the validation set.

Training and Validation of the Combined Model

The variables retained in this model were found to be the same as those in the separate models: *Surface to Volume ratio*, presence of tumour necrosis at preoperative

20

CT imaging and CA 19.9 serum levels. The model developed in the training cohort demonstrated good overall performance (p = 0.0015, AUC = 0.75). After calculating the corresponding P index, the model was confirmed in the validation cohort (p = 0.00178, AUC = 0.76). Further details are provided in Table 2. Corresponding ROC curves and *Kaplan-Meier* survival curves based on the P index best threshold are shown in Figure 6. The combined model demonstrated an excellent performance in stratifying the risk of distant relapse, especially in the first months after upfront surgery (training cohort: HR = 3.58, 95% CI = 1.91 – 6.71; validation cohort: HR = 5.06, 95% CI = 1.75 – 14.58): at 12 months after surgery 50% of *high risk* patients experienced distant relapse of disease vs. 12% of *low risk* patients (p < 0.001).

Discussion

There is growing literature demonstrating the efficacy of neoadjuvant chemotherapy in patients with resectable pancreatic ductal adenocarcinoma (Lee et al, 2019; Reni *et al*, 2018a; Reni *et al*, 2018b); whether all these patients should receive preoperative chemotherapy remains, though, controversial. The main issue is that, currently, there is no clinically relevant tool able to accurately stratify patients in terms of early distant relapse (EDR) after upfront surgery. Previously proposed models have limited clinical utility mainly because they consist of pathologic data obtained after surgery and therefore are not applicable in a preoperative setting (Groot *et al*, 2018; Kudo *et al*, 2020; Strijker *et al*, 2019); another major limitation is the poor, inhomogeneous selection of the study cohorts (Kim *et al*, 2020; Eilaghi *et al*, 2017). In the present study we sought to develop a preoperative model to help identify patients with increased risk of EDR after upfront surgery for pancreatic head adenocarcinoma. To facilitate its use in a clinical setting, only three variables were retained in the final, internally validated combined model: one radiomic feature (*Surface to Volume ratio*), one conventional radiological variable (presence of *tumour necrosis* at preoperative CT imaging), and one clinical variable (*CA 19.9 serum levels*). According to these three variables, a prognostic index can easily be derived for each patient, being a surrogate for the risk of developing EDR after primary surgery. Of note, the combined model outperformed the separate ones (radiomic and clinicoradiological) in terms of *i)* overall performance, i*i)* robustness and reproducibility, and, above all, *iii)* outcome prediction.

Literature has widely described the importance of both radiological *tumour necrosis* and *CA 19.9 serum levels* in outlying the biological behaviour of pancreatic adenocarcinomas regardless of anatomical resectability. Kudo *and colleagues*, for instance, identified a worthwhile relation between radiological *tumour necrosis* and pathological lymph node metastasis and lymphovascular invasion, strongly affecting overall prognosis (Kudo *et al*, 2020). On the other hand, *CA 19.9 serum levels* have been reported to well correlate with disease burden, even besides what imaging can show (Petrelli *et al*, 2017). Our results corroborate this evidence. In our cohort, *CA 19.9 serum levels* have been found to be the most informative clinical predictor of EDR after primary surgery (35 U/mL [non-EDR group] *vs*. 106 U/mL [EDR group], p < 0.001).

22

On the contrary, the biological significance of the radiomic feature finally retained in our model, *Surface to Volume ratio*, has not been investigated. It belongs to the morphological family of the radiomic features, summing up the relationship between the surface area of a given object and its volume. Our data highlighted a strong inverse relation between this neoplastic feature and the occurrence of EDR after upfront surgery: in short, adenocarcinomas with low *Surface to Volume ratio* values were more prone to early relapse after primary surgery. With regard to this last point, one may argue that the assumption that a round shaped tumour with smooth contours (the geometrical object lowering at most *Surface to Volume ratio* is a sphere) should have a worse prognosis when compared to an ill-defined one is, at least, counterintuitive. In this respect, three reflections have to be done.

1. Limkin *and colleagues* (Limkin *et al*, 2019) demonstrated that the major determinant of *Surface to Volume ratio* is volume, and therefore it should not be considered as an immediate surrogate for tumour complexity, but rather as a precise tool for dimensional assessment. However, in our cohort, *Surface to Volume ratio* has been proven to be an extraordinary predictor of EDR, exceeding the other dimension-related variables considered. It follows that *Surface to Volume ratio* accounts for more information than raw dimensional data do.

2. According to Bribiesca, for similar values of volume, *Surface to Volume ratio* could be considered an indirect expression of geometrical compactness (Bribiesca, 2000). Based on our own data, a compact pancreatic tumour (low *Surface to Volume ratio*) has to be considered at high risk of EDR. Mori *and colleagues*, while developing a PET (positron emission tomography)-based radiomic signature to predict distant relapse free survival in patients with locally advanced pancreatic adenocarcinoma, identified a radiomic feature (*Centre of Mass shift*), which represents the distance between the geometrical and the metabolic centroids of a given object: the smaller the shift, the more homogeneous the uptake of the contoured lesion and, ultimately, the higher its compactness (Mori *et al*, 2020). The authors found that low values of *Centre of Mass shift* contributed to worse prognosis, which is in line with

our results.

3.   From a biological point of view, *Surface to Volume ratio* could be considered the major determinant of cell size, since a low ratio may undermine the rate of chemical exchange, resulting in cell death (Harris & Theriot, 2018). Moving to tumour perspective, a low *Surface to Volume ratio* implies impaired vascularization, possibly resulting in tumour hypoxia and necrosis.

Taken together, these observations allow considering *Surface to Volume ratio* far more than just a mere morphological feature, and give a novel insight into pancreatic tumour biological behaviour.

However, one may question which is the contribution of *Surface to Volume ratio* to the overall performance of our combined model. The prognostic index computed from the combined model allowed optimal dichotomization of the validation cohort with 50% of *high risk* patients experiencing distant relapse of disease within 11 months after surgery vs. 12% of *low risk* patients (p < 0.001). Interestingly, this result was obtained by including a single, extremely robust radiomic feature with a strong biological rationale selected throughout a highly reliable methodological approach. Our approach, hence, differs substantially from previously reported, similar studies (Eilaghi *et al*, 2017; Abunahel *et al*, 2021; Xie *et al*, 2020) relying instead on several radiomic features ultimately impairing the actual deployment of the resulting models in clinics. Straightforwardness, robustness and reproducibility are, on the contrary, key features of the combined model we propose, which is, moreover, entirely presurgical.

Clinically, our data suggest that those upfront resectable patients thought to be at *high risk* of EDR according to our combined model should be scheduled to neoadjuvant treatment; on the other hand, *low risk* patients might be considered a highly selected group possibly suitable for primary resection.

The present study has several limitations, the most important being its retrospective nature and the relatively small number of events observed. External validation is also warranted. Furthermore, our model has been thought not for a standalone usage but rather to be embedded in the multidisciplinary assessment of the patient, which remains

24

the absolute cornerstone in the battleground against pancreatic adenocarcinoma.

In conclusion, despite the abovementioned limitations, we have developed a robust, entirely preoperative tool to predict early distant relapse of disease after upfront surgery. This model redefines resectability status and provides a personalized tool for patients' management, identifying those upfront resectable patients at *high risk* of early recurrence who would benefit from neoadjuvant chemotherapy, as well as those at *low risk*, which may represent a highly selected group potentially suitable for primary resection. Independent validations of the model are warranted to further corroborate its validity.

		TOTAL (n=147)	TRAINING (n=94)	VALIDATION (n=53)	p-value
CLINICAL VARIABLES	Age at diagnosis (year)*	69.94 (44-88)	70.06 (43-87)	69.73 (43-88)	0.84
	Sex				0.13
	Female	61 (41.6%)	35 (37.6%)	27 (50.5%)	
	Male	86 (58.4%)	59 (63.4%)	26 (49.5%)	
	CA 19.9 (U/mL)*	40 (14-150)	48.5 (13.25-191)	35 (14.5 - 77.5)	0.21
	CA 19.9				0.15
	<200 U/ml	111 (75.6%)	69 (73.4%)	42 (78.6%)	
	≥200 U/ml	36 (24.4%)	25 (26.6%)	11 (20.8%)	
	Adjuvant Treatment	111 (75%)	69 (73.2%)	42 (78.6%)	0.38
	Adjuvant Chemotherapy	107 (73.0%)	68 (73.0%)	39 (73.2%)	0.45
	Adjuvant Radiotherapy	31 (21%)	21 (22.3%)	10 (19.6%)	0.11
PATHOLOGICAL DATA	Tumor Size (mm)^	27.33 (+/- 0.78)	28.36 (+/- 0.97)	25.48 (+/- 1.29)	0.78
	Final R status				0.08
	R1	65 (44.2%)	44 (46.8%)	21 (39.6%)	
	R0	82 (55.8%)	50 (53.2%)	32 (60.4%)	
	Lymph-vascular Invasion	129 (87.8%)	88 (93.6%)	41 (76.8%)	0.015
	Perineural Invasion	130 (87.1%)	86 (91.4%)	44 (82.1%)	0.06
	Peripancreatic Fat Invasion	135 (91%)	92 (97.8%)	43 (80.3%)	0.72
	Grading				0.42
	G1	3 (2.0%)	3 (3.1%)	0 (0%)	
	G2	66 (45.0%)	45 (47.8%)	26 (48.2%)	
	G3	78 (53.0%)	46 (49.1%)	27 (51.8%)	
	TNM				
	T				0.75
	T1	34 (22.7%)	20 (20.3%)	14 (26.8%)	
	T2	102 (69.2%)	68 (72.3%)	34 (64.3%)	
	T3	11 (7.7%)	7 (7.4%)	5 (8.9%)	
	N				0.065
	N0	23 (16.1%)	15 (15.0%)	9 (17.8%)	
	N1	50 (33.9%)	26 (27.6%)	24 (44.6%)	
	N2	74 (50%)	54 (57.4%)	20 (37.5%)	
	Lymphnode Ratio*	0.136 (0.04 - 0.25)	0.16 (0.06-0.26)	0.10 (0.03-0.2)	0.11
RADIOLOGICAL DATA	Dimension (mm)^	24.5 (+/- 7.2)	25.79 (+/- 6.8)	23.23 (+/- 7.8)	0.07
	Necrosis				0.34
	Present	27 (18.2%)	18 (19.1%)	9 (16.9%)	

		TOTAL (n=147)	TRAINING (n=94)	VALIDATION (n=53)	p-value
	Absent	120 (81.6%)	76 (85.2%)	44 (83.0%)	
	Hypodense on pancreatic phase	116 (74.4%)	77 (81.9%)	39 (73.5%)	0.48
	Hypodense on venous phase	93 (57.1%)	91 (96.8%)	2 (3.5%)	0.50
	Isodense on pancreatic phase	23 (21.8%)	21 (23.4%)	2 (3.5%)	0.58
VARIABLES OUTCOME	Early distant recurrence (EDR)				0.54
	EDR	39 (25.6%)	25 (26.5%)	14 (26.3%)	
	non-EDR	108 (74.4%)	69 (73.5%)	39 (73.7%)	
	Time to recurrence (months)*	15 (10 - 22)	15 (9 - 22)	16 (11 - 26)	0.55
	Overall survival (months)*	20 (15-28)	20 (15-28)	20 (16-28.5)	0.98
	Length of follow-up (months)*	19 (14 - 27)	19 (14 - 27)	19.5 (13.25 - 27.75)	0.94

Unless otherwise indicated, data are numbers of patients and data in parentheses are percentages. p values were determined by comparing characteristics between patients of Training and Validation cohort.

*Data are medians, data in parentheses are ranges
^ Data are means with standard deviations

TABLE #1: *Patients' characteristics. Palumbo et al, DOI: 10.3390/cancers13194938*

Table #2 — Overall performance of radiomic, clinicoradiological and combined models

Legend: Training (orange), Validation (blue)

Model	Variable	Coefficient	p-value	OR	95% CI	Overall Fit Model (p-value)		AUC		95% CI		Sensivity		Specifity		PPP		NPP	
						Training	Validation	Training	Validation	Training	Validation	Training	Validation	Training	Validation	Training	Validation	Training	Validation
Radiomic Model	Surface To Volume Ratio	−3.82224	0.0183	0.0219	0.0009 to 0.5237	0.0097	0.0244	0.599	0.733	0.493 to 0.699	0.593 to 0.846	90.9	100.0	23.08	26.2	62.5	24.4	64.3	100.0
Clinico Radiological Model	Ca19.9	0.001128	0.049	1.0011	1.0000 to 1.0023	0.0018	0.9529	0.72	0.536	0.614 to 0.811	0.392 to 0.675	68.9	20.0	76.27	73.8	58.8	15.4	83.3	79.5
	Necrosis	1.10839	0.0633	3.0295	0.9402 to 9.7611														
Combined Model	Ca19.9	0.000964	0.0946	1.011	0.9998 to 1.0021	0.0015	0.0178	0.736	0.76	0.648 to 0.838	0.618 to 0.865	58.6	60.0	86.44	78.5	68.0	40.0	81.0	89.2
	Necrosis	0.86465	0.1594	2.3742	0.7120 to 7.9167														
	Surface To Volume Ratio	−2.8972	0.1171	0.0552	0.0015 to 2.0684														

Table #2: *Overall performance in both training and validation cohorts of radiomic, clinicoradiological and combined models quantified in terms of area under the ROC curve (AUC), positive and negative predictive values, specificity and sensitivity. Palumbo et al, DOI: 10.3390/cancers13194938*
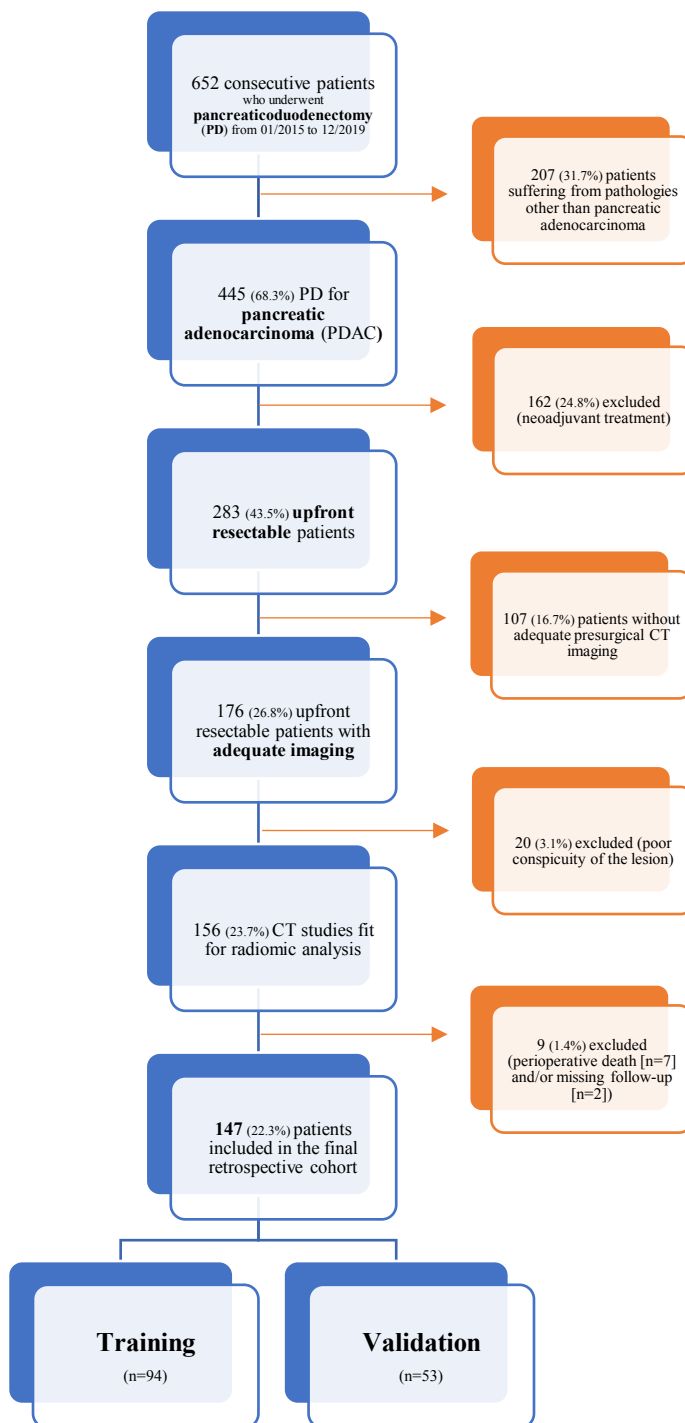
Figure #1: *Inclusion and exclusion criteria flowchart. Palumbo et al, DOI: 10.3390/cancers13194938*
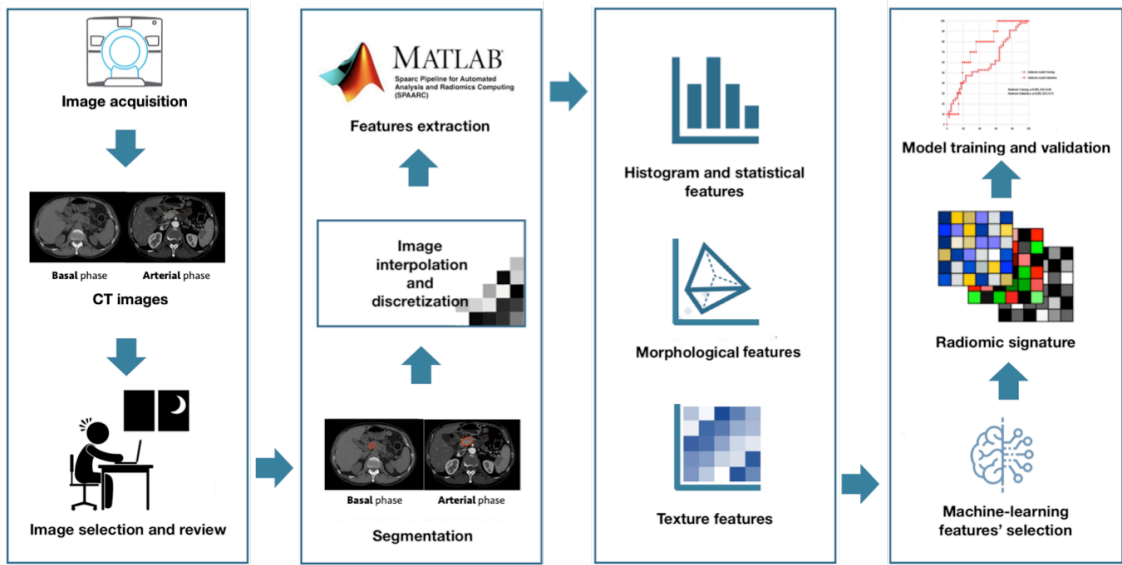
Figure #2: *Radiomic workflow for features extraction and models development.* Modified from *Palumbo et al, DOI: 10.3390/cancers13194938.*
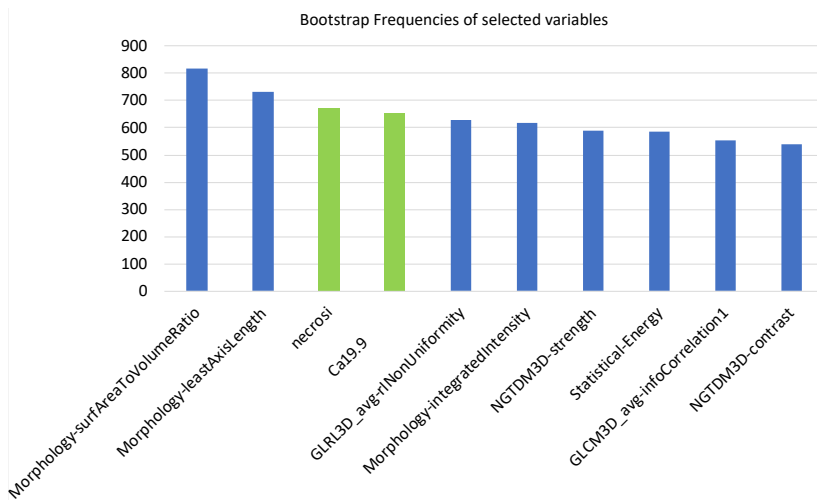

Figure #3: *10 most frequent variables identified through machine-learning bootstrap-ranking procedure: eight radiomic features (3 morphologic, 4 texture related and 1 statistical features) and two clinicoradiological variables (radiological tumour necrosis, serum level of CA19.9). Palumbo et al, DOI: 10.3390/cancers13194938.*
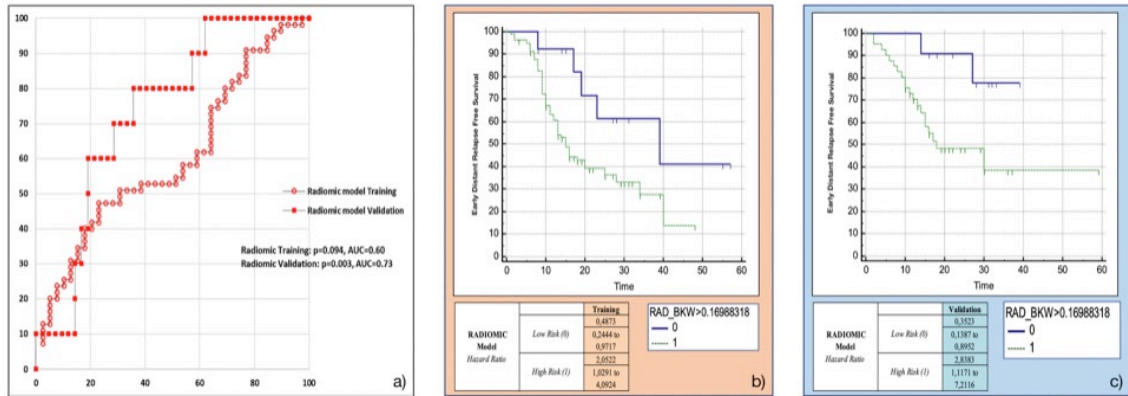
Figure #4: *Radiomic model overall performance in terms of [a] area under the ROC curve (AUC) for both training (red empty circles) and validation (red filled squares) cohorts, and outcome prediction in terms of Kaplan Meier curve separation between low and high-risk patients according to the computed prognostic index in both cohorts. Palumbo et al, DOI: 10.3390/cancers13194938*
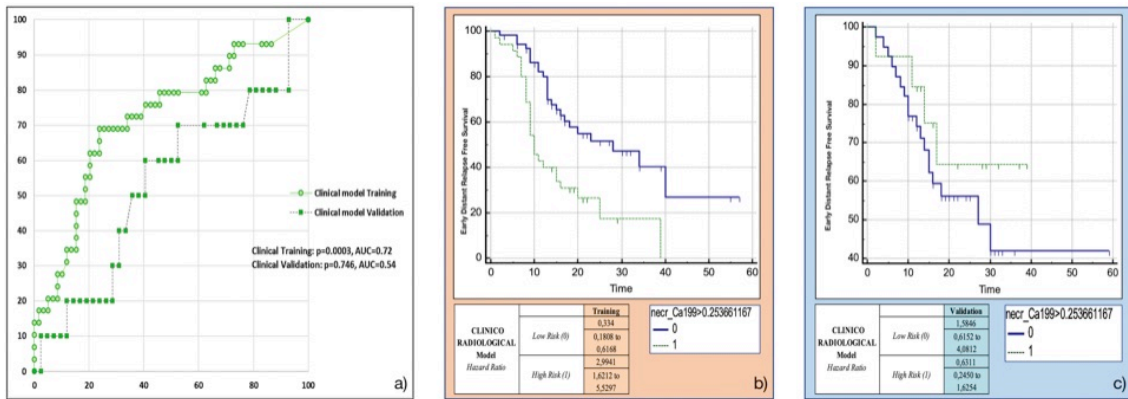
Figure #5: *Clinicoradiological model overall performance in terms of [a] area under the ROC curve (AUC) for both training (green empty circles) and validation (green filled squares) cohorts, and outcome prediction in terms of Kaplan Meier curve separation between low and high-risk patients according to the computed prognostic index in both cohorts. Palumbo et al, DOI: 10.3390/cancers13194938*

Figure #6: *Combined model overall performance in terms of [a] area under the ROC curve (AUC) for both training (blue empty circles) and validation (blue filled squares) cohorts, and*

31

outcome prediction in terms of Kaplan Meier curve separation between low and high-risk patients according to the computed prognostic index in both cohorts. Palumbo et al, DOI: 10.3390/cancers13194938
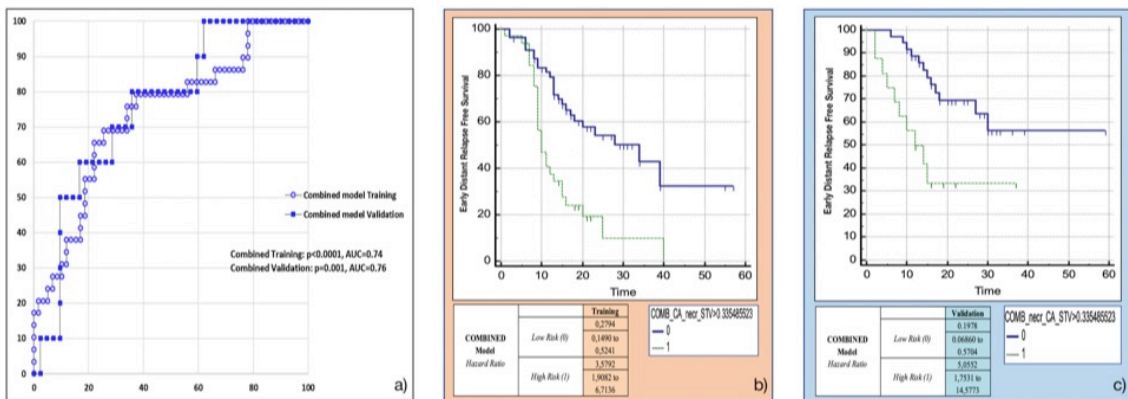
Identification and biological significance of a novel radiological biomarker for IPMN with *high-grade* dysplasia/invasive carcinoma.

Materials and Methods

Patients' cohort

This is a single center prospective observational study conducted at San Raffaele Scientific Institute (Milan, Italy), part of a broader investigation funded by minister of health (*Ricerca Finalizzata Giovani Ricercatori* (2018) GR-2018-12366897 – "Definition of radiological and endocrine/metabolic/inflammatory biomarker(s) to identify high grade dysplasia/invasive carcinoma in patients with IPMN of the pancreas"); data were collected after Ethics Committee approval.

Between July 2020 and November 2021, 48 adult patients (Charlson comorbidity index < 7) scheduled for surgical resection for suspected IPMN (based on current major guidelines) were prospectively enrolled into our study.

Within 60 days before index surgery all patients underwent magnetic resonance cholangiopancreatography (MRCP) with a dedicated protocol (including mDixon Quant Fat FRACTION sequences for liver and pancreatic fat content quantification) and/or contrast-enhanced CT scan of the abdomen.

In order to better characterize the endocrine, metabolic and inflammatory microenvironment of IPMN, samples from peripheral and portal vein blood, as well as from neoplasm cystic fluid have been systematically collected during surgery; in particular: (i) *20 ml of fasting peripheral blood samples*, collected immediately after the induction of general anaesthesia; ii) *20 ml of portal blood*, taken from a collateral vein of portal vein, specifically splenic vein for distal pancreatectomy, superior pancreatic vein for pancreaticoduodenectomy (4 patients didn't undergo portal blood sampling due to portal vein thrombosis; 27 patients had portal blood samples < 20 ml, so they only underwent Luminex®/ELISA analyses); iii) *neoplasm cystic fluid*, collected within an hour after the completion of pancreatic resection from a dedicated pancreatic pathologist (we were unable to obtain samples from 13 patients because of mainly solid lesions). The aforementioned samples were used to evaluate the levels of cytokines/chemokines, apolipoproteins, adipokines, hormones, and other biochemical

parameters related to inflammation and lipid metabolism, in order to assess eventual correlations with pathology.

Finally, after surgery, all these (radiological, biological) data were systematically correlated with the pathological outcome.

MR/CT protocol and imaging findings

When considered of unsatisfactory quality, imaging procedures were repeated at our site upon request. MR examinations were all performed with a 1.5 T scanner (Ingenia, *Philips Medical Systems* – Best, The Netherlands) with an eight-channel phased-array torso coil; imaging protocol has been previously published (Crippa *et al*, 2022). To reduce hyperintense signal from fluid-filled stomach and duodenum, a negative oral contrast agent (high manganese content agents – pineapple or blueberry juice) was administered before acquisition. CT protocol included intravenous administration of iodine contrast medium and consisted of a multiphase acquisition (unenhanced, late arterial and portal venous axial scans of the abdomen).

Two radiologists with a particular commitment for pancreatic imaging reviewed all images; inter reader agreement has been systematically assessed for each one of the selected findings using Fleiss' kappa (categorical variables) and intraclass correlation coefficient (ICC) (continuous variables). Imaging findings were selected for analysis on the basis of clinical experience and existing literature; aside from *worrisome features* (*i)* cyst > 3 cm, *ii)* enhancing mural nodule < 5 mm, *iii)* thickened/enhancing cyst walls, *iv)* main pancreatic duct [MPD] calibre between 5 and 9 mm, *v)* abrupt change in MPD calibre with distal parenchymal atrophy, *vi)* lymphadenopathies) and *high-risk stigmata* (*i*) enhancing mural nodule > 5 mm, *ii)* MPD calibre > 10 mm) (Tanaka *et al*, 2017), the following findings were also considered: *i)* peripancreatic oedema/fat stranding,(evaluated on fat saturated T2-weighted images), *ii*) pancreatic/liver fat content (measured on mDixon Quant Fat FRACTION sequences by sampling and then averaging three standardized regions of interest at the level of V, VI e VIII liver segment [for liver assessment] and three standardized regions of interest at the level of pancreatic head, body and tail [for pancreatic assessment]).

CT imaging post-processing analysis – Two sets of analysis have been systematically carried on for all patients who underwent CT imaging:

- *Body composition assessment* (SliceOmatic version 5.0, Tomovision): an axial CT image at the level of the third lumbar vertebra (L3) was processed for each patient, as per previous literature (Pecorelli *et al*, 2022; Guarneri *et al*, 2022). Total abdominal muscle area (TAMA) (cm^2), then normalized with respect to stature and rationalized for sex (sarcopenia was defined for TAMA values lower than 52.4 and 38.5 cm^2/m^2 in males and females, respectively), visceral fat area (VFA) (cm^2) and subcutaneous fat area (SFA) (cm^2) were systematically calculated. The VFA/TAMA and VFA/SFA ratios were calculated for each patient and analysed independently.

- *Volumetric assessment of IPMN (cystic and solid component) and residual pancreatic parenchyma* (Intellispace version 8.0, Phillips, "tumour tracking" tool).

Statistical analysis [**The following paragraph contains material that has been already published – Palumbo *et al*, DOI: 10.1007/s00330-021-07788-y**]

The reliability of agreement has been assessed for each one of the selected imaging findings using Fleiss' kappa for categorical features and intraclass correlation coefficient (ICC) for continuous variables. Fleiss' kappa values > 0.80 were deemed to be excellent; ICCs > 0.75 were considered to be in high agreement.

The association of any finding with different degrees of dysplasia was studied univariably. Categorical parameters were compared using the χ^2 test, whereas *Student's t* and the *Mann–Whitney U* tests were used to compare groups with and without *high-grade* dysplasia/invasive carcinoma, for parametric and non-parametric continuous variables, respectively. The relationship between continuous variables was assessed by Spearman bivariate correlation.

Statistical analysis was performed using commercially available software (SPSS version 21.0, IBM Corp.). A *p-value* < 0.05 was considered statistically significant.

Results

Patients' characteristics

After exclusions, a total of 48 patients with suspected IPMN candidate for resection (20 females [41.7%], with a median age of 69 years [IQR: 59-74]) have been enrolled into this prospective study.

All patients underwent major surgical procedures: 20 distal pancreatectomies (41.7%), 20 pancreaticoduodenectomies (41.7%), 8 total pancreatectomies (16.6%).

After surgical resection, definitive histological examinations were as follows: 35 IPMN (72.9%) (12 with low dysplasia [34.2%]; 8 with moderate dysplasia [22.8%]; 15 with severe dysplasia/invasive carcinoma [42.8%]); one PDAC without any IPMN component (2%), 12 other (25%).

Imaging findings

Within 60 days before index surgery, 43 patients (89.5%) underwent MRCP (5 patients suffered from claustrophobia and cannot complete scheduled examination), 42 (87.5%) contrast enhanced CT (6 patients had a history of allergic reactions to iodinated contrast medium), 37 (77%) both.

Worrisome features (WF)/High-risk stigmata (HRS) − Data regarding WF/HRS are summarized in Table 4. Presence of an enhancing mural nodule ≥ 5 mm at CT (p = 0.015), and MR evidence of abrupt main pancreatic duct calibre change with distal pancreatic atrophy (p = 0.012) as well as thickened/enhancing cyst walls (p = 0.002) were more common in IPMN patients with *high-grade* dysplasia/invasive carcinoma when compared to those with *low/moderate-grade* dysplasia. Moreover, a trend towards statistical significance exists when considering MR evidence of main pancreatic duct ≥ 10 mm (p = 0.052).

Diagnostic performance of contrast enhanced CT and MRCP for predicting malignant potential of IPMNs − When defining malignant IPMN with a total score of three WF with or without at lest one HRS (such combination represents indeed the best threshold to differentiate, in our cohort, between patients *with low/moderate-grade* dysplasia and those with *high-grade* dysplasia/invasive carcinoma), the sensitivity and specificity were 83.3% [95%CI: 51.6-97.9] (10 of 15) and 78.3% [95%CI: 56.3-92.5] (18 of 20) with CT, respectively, and 85.7% [95%CI: 57.2-98.2] (12 of 15) and 85.7%

[95%CI: 63.6-96.9] (18 of 20) with contrast enhanced MRCP, respectively. Of note, the diagnostic performance of CT and MR were similar in recognizing WF/HRS (area under the curve, 0.79 and 0.85, respectively; p = 0.232).

Interobserver agreement – All HRS/WF had moderate to good interobserver agreement at both CT and MRCP (qualitative variables: k = 0.55–0.82; quantitative variables: ICC = 0.68–0.83) except for the presence of thickened/enhancing cyst walls, which showed poor interobserver agreement for both modalities (k = 0.28 for CT, k = 0.39 and MRCP). Of note, the interobserver agreement for the presence of thickened/enhancing cyst walls at MR significantly improved using subtracted images (k = 0.72).

Intermodality agreement – Between contrast enhanced CT and MRCP, agreement was excellent for both HRS and WF (k = 0.81–0.95) except for cyst dimension > 4 cm, main pancreatic duct size ≥ 10 mm, lymphadenopathies and the presence of thickened/enhancing cyst walls, which showed fair to good agreement (k = 0.55, 0.39, 0.39 and 0.35, respectively).

CT imaging post processing

Body composition assessment – Based upon previously validated thresholds, 33 out of 42 patients (78.6%) who underwent preoperative CT were found to be sarcopenic; no significant differences were found, in terms of sarcopenia distribution, between IPMN patients with different degree of dysplasia (p = 0.348). Differently, VFA/TAMA showed a trend towards statistical significance (p = 0.058) in differentiating patients *with low/moderate-grade* dysplasia and those with *high-grade* dysplasia/invasive carcinoma; specifically, sarcopenic obesity (high VFA/TAMA values) was found to be more frequent in the latter group.

"Volumetric" assessment – No statistically significant differences were found between IPMN patients with different degree of dysplasia when considering median volume of cystic disease (5.29 cc [1.87–24.03]) (p = 0.542), median volume of uninvolved pancreatic parenchyma (33.74 cc [21.99-61.00]) (p = 0.180), nor, when present, median volume of solid disease (2.72 cc [0.77-6.00]) (p = 0.378).

Fat content assessment (measured by mDixon Quant Fat FRACTION sequences)

Mean liver fat content was 4.27% ± 5.27; mean pancreatic fat content was 11.04% ± 10.44. Of note, when dealing with histologically proven IPMN, pancreatic fat content (as measured by mDixon Quant Fat FRACTION sequences) was proven to be significantly higher in those patients with *high-grade* dysplasia/invasive carcinoma (15.65% ± 13.8) when compared to *low/moderate-grade* dysplasia (6.86% ± 4.95) (p = 0.023). Furthermore, pancreatic fat fraction was found to be directly related to MCP (monocyte chemoattractant protein)-1 levels in peripheral blood (R = 0.628, p < 0.001); a trend towards statistical significance also exists when considering MCP-1 levels in cystic fluid (R = 0.418, p = 0.075). Preliminary results from the biological pillar of the same study pointed out a statistically significant association between MCP-1 levels in cystic fluid and IPMN with *high-grade* dysplasia/invasive carcinoma (p = 0.0077), as shown in Figure 7. Finally, differently from liver fat content which was significantly related to apolipoprotein-A1 levels in portal vein blood (R = 0.584, p < 0.001), pancreatic fat fraction showed no correlation with any apolipoprotein (in any biological compartment).

Discussion

Up to 40% of patients undergoing abdominal imaging harbour unsuspected pancreatic cysts, mostly intraductal papillary mucinous neoplasms (IPMN), which are well-known precursors of pancreatic ductal adenocarcinoma (PDAC) but without an established temporal window of progression (Gardner *et al*, 2013; Overbeek *et al*, 2022). Current guidelines identify high-risk stigmata (HRS)/absolute criteria and worrisome features (WF)/relative criteria as indications for surgical resection of IPMN (Tanaka *et al*, 2017). However, no reliable malignancy biomarker for IPMN with *high-grade* dysplasia is known (Crippa *et al*, 2022), and the overtreatment (surgery, with high morbidity/mortality rate (Aleotti *et al*, 2022; Capretti *et al*, 2018)) of benign IPMN remains a critical clinical issue, as well as under treatment. Accordingly, our aim was to identify preoperative radiological biomarker(s) to distinguish *low-* versus *high-risk* IPMN for cancer progression.

As main result of our investigation, we found that the pancreatic fat content (measured by mDixon Quant Fat FRACTION sequences) is significantly higher in IPMN patients with *high-grade* dysplasia/invasive carcinoma when compared to those with *low/moderate-grade* dysplasia.

In this respect, there is literature suggesting a role for severe obesity (BMI [Body Mass Index] $\geq$ 35 kg/m^2) in increasing the frequency of malignancy in patients with branch duct (BD)-IPMN (Mathur *et al*, 2009; Capurso *et al*, 2020). However, a major issue undermines these studies, since BMI cannot provide a reliable depiction of each metabolic compartment; specifically, it cannot differentiate whether fat accrual occurs subcutaneously, in the visceral adipose tissue, or instead within liver/pancreatic parenchyma. On that note, our data highlight the importance of a very specific type of obesity characterized by subtle systemic inflammation (sarcopenic obesity (Pecorelli *et al*, 2022; Guarneri *et al*, 2022)), which seems to be prevalent among IPMN patients with *high-grade* dysplasia/invasive carcinoma. In this setting, pancreatic fat accumulation could be a mere consequence of such metabolic dysregulation; however, two points have to be taken into account.

- Our results demonstrate a strong direct correlation between pancreatic fat content measured by mDixon Quant Fat FRACTION sequences and MCP (monocyte chemoattractant protein)-1 levels in

peripheral blood (p < 0.001); a trend towards statistical significance also exists when considering MCP-1 levels in cystic fluid (p = 0.075). MCP-1 (also known as CCL2 [CC chemokine ligand 2]) is a small chemokine whose expression by pancreatic cancer cells is up regulated by pro inflammatory cytokines (it can be considered a sort of pancreatic cancer marker), resulting in the recruitment of leukocytes at the level of the tumour microenvironment (Gu *et al*, 2021). On that note, Monti *and colleagues* (Monti *et al*, 2003) found that serum MCP-1 levels positively correlate with tumour macrophage infiltration. Of note, macrophages have complex interactions with neoplastic cells, frequently displaying a mixed polarization state. On the one side, M1 macrophages are known to inhibit neoplastic growth; on the other, M2 macrophages produce lymphangiogenic growth factors and proteases, thus promoting digestion of extracellular matrix and, ultimately, neoplastic progression. Moreover, preliminary results from the biological pillar of our study pointed out a statistically significant association between MCP-1 levels in cystic fluid and IPMN with *high-grade* dysplasia/invasive carcinoma.

• Our data also point out a striking, somehow surprising difference between liver steatosis and pancreatic fat storage, two entities brought together by a similar imaging appearance. Differently from liver fat content, which was found to be (as expected) significantly related to apolipoprotein-A1 levels in portal vein blood (p < 0.001) (Karavia *et al*, 2012), pancreatic fat fraction showed no correlation with any protein (in any biological compartment) involved in lipid metabolism.

Taken together, these observations promote the hypothesis of high pancreatic fat fraction being a marker of inflammation at the level of the tumour microenvironment, rather than just the expression of fat over accrual within pancreatic parenchyma, possibly being a target for future therapies. Although not statistically significant, also the high prevalence of sarcopenic obesity among IPMN patients with *high-grade* dysplasia/invasive carcinoma pushes in the same direction of subtle inflammation as a possible primer for pancreatic carcinogenesis (as also supported by existing literature

(Sadot *et al*, 2015)). Previous authors reported pancreatic fat fraction as measured by mDixon Quant Fat FRACTION sequences as being a reliable measure of the actual pathological fat content (Sotozono *et al*, 2022; Lin *et al*, 2022; Wen *et al*, 2022); they also suggest a possible relationship with inflammation, but with no clear data supporting this hypothesis. In this respect, major novelty of the present manuscript is to provide the first *in vivo* biological justification for pancreatic fat fraction as a reliable marker of local inflammation and, ultimately, malignant progression.

The present study has several limitations, the main ones being the small sample size (preliminary findings from the second installment of patients [overall n = 82] corroborate the presented observations) and the lack of precise pathological quantification of pancreatic fat content. On that note, it is worth noting that preliminary observations from pathological specimens of our series confirm the absolute peculiarity of fat accrual within pancreatic parenchyma when compared to liver steatosis: no ballooning degeneration was indeed observed, whereas fatty infiltration is predominant as meaning that pancreatic parenchyma is variably replaced by adipocytes mostly distributed around arterioles. Another limitation to be taken into account refers to possible heterogeneity of fat accrual within the pancreas, resulting in not reliable measurements as obtained by placing the three 2D standardized regions of interest; to avoid/limit this risk, each sample has been taken based on the consensus of the two experienced radiologists. That said, we admit a possible methodological bias related with this issue, and future studies have to prefer volumetric sampling to represent the exact amount of fat accrual (Lin *et al*, 2022).

In conclusion, our preliminary findings suggest that quantification of pancreatic fat content by mDixon Quant Fat FRACTION sequences is a promising tool to provide non-invasive characterization of tumour microenvironment (high pancreatic fat content stands for subtle inflammation within tumour microenvironment, possibly related to the amount of tumour-infiltrating macrophages with pro-oncogenic commitment) and, ultimately, distinguish *low-* versus *high-risk* IPMN for cancer progression, also having regard to the low accuracy of conventional WF/HRS. Therefore, it may complement the current IPMN diagnostic workflow and improve clinical decision-making regarding the need for surgical resection or, alternatively, surveillance.

.

	CT			MR		
	Benign (n = 20)	Malignant (n = 15)	P Value	Benign (n = 20)	Malignant (n = 15)	P Value
Main Pancreatic duct size ≥ 10 mm	4 (20%)	5 (33.3%)	0.812	5 (25%)	7 (46.6%)	<u>0.052</u>
Enhancing mural nodule ≥ 5 mm	0 (0%)	4 (26.6%)	**0.015**	0 (0%)	2 (13.3%)	0.191
Cyst ≥ 3 cm	10 (50%)	8 (53.3%)	0.484	9 (45%)	9 (60%)	0.314
Thickened / Enhancing cyst wall	1 (5%)	2 (13.3%)	0.356	2 (10%)	10 (66.6%)	**0.002**
Main Duct size 5-9 mm	7 (35%)	5 (33.3%)	0.724	6 (30%)	4 (26.6%)	0.745
Enhancing mural nodule < 5 mm	1 (5%)	0 (0%)	0.324	0 (0%)	1 (6.6%)	0.679
Abrupt change in calibre of pancreatic duct with distal pancreatic atrophy	7 (37%)	8 (53.3%)	0.101	6 (30%)	9 (60%)	**0.012**
Lymphadenopathy (short axis > 10 mm)	1 (5%)	1 (6.6%)	0.66	0 (0%)	0 (0%)	N/A
Macroscopic Solid Component (it is NOT the mural nodule)	0 (0%)	4 (26.6%)	**0.029**	1 (5%)	3 (20%)	0.109
Peripancreatic Fat Heterogeneity	2 (10%)	3 (20%)	0.345	4 (20%)	5 (33.3%)	0.151

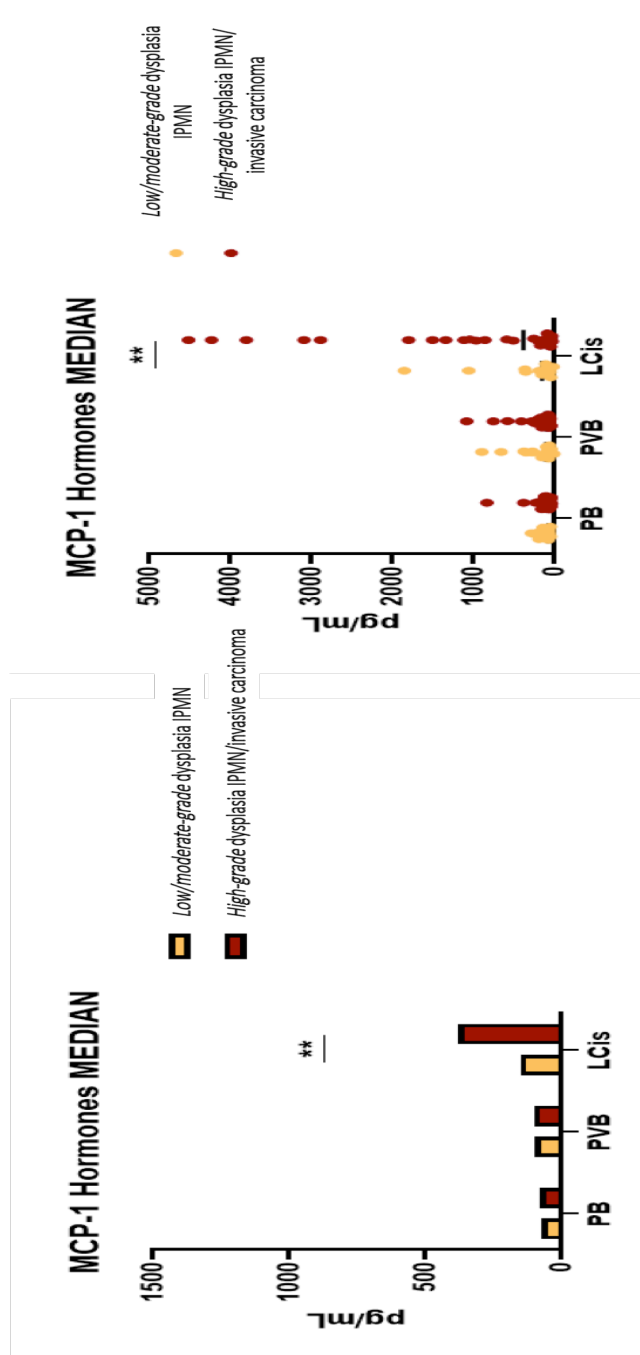**Table #3**: *High Risk Stigmata/Worrisome Features*.

Figure #7: *Correlation between MCP-1 levels in cystic fluid and IPMN with high-grade dysplasia/invasive carcinoma.*

Prediction of the characteristics of aggressiveness of pancreatic neuroendocrine neoplasms based on CT radiomic features.

Materials and Methods

Patients' cohort

This is a single-center, retrospective, observational study including patients who underwent upfront surgery for pancreatic neuroendocrine neoplasms (panNEN) at San Raffaele Scientific Institute (Milan, Italy) from January 2015 and December 2021; data were collected within the context of an Ethics Committee approved study (06/INT/2022) in patients who had signed an institutional procedure specific informed consent. From a prospectively acquired database, adult patients without visible distant metastases who underwent adequate quality abdominal CT imaging within one month before index surgery were enrolled. The resulting population (n = 101) was then randomly split into training (n = 70) and validation (n = 31) cohorts according to the second level of the TRIPOD guidelines for the validation of predictive models in oncology (Moons *et al*, 2015).

The histological endpoints considered, as defined by postoperative histological specimens, were: tumor grade (G) (G1vs.G2/3), the presence of distant metastases (M+), metastatic lymphnodes (N+) and microvascular invasion (VI).

Clinical variables

Demographic variables were retrospectively reviewed from an electronic database.

Radiological variables and radiomic features (RFs)

Methodological details regarding CT protocol have been already published (Palumbo *et al*, 2021b).

The variability of the features between different scanners were tested through Mann-Whitney test, finding no significant inter-scanner variations: this result, together with the careful application of the above mentioned acquisition protocols should limit any potential bias due to features repeatability.

Conventional imaging parameters –The selected CT findings included the

followings: *i)* presence of necrosis, defined as tumoral, non cystic tissue that did not enhance in the arterial and portal venous phases, *ii)* presence of cystic/liquid component, *iii)* pancreatic parenchyma atrophy, defined as a significant reduction in the volume of the gland, *iv)* macroscopic arterial/venous infiltration and *v)* contiguous organs invasion.

Delineation – The robustness of CT RFs against interobserver contouring variability was already assessed by our group in the specific setting of panNENs, showing high intra-correlation coefficient (ICC) for all features (Loi *et al*, 2020; Mori *et al*, 2019). As previously reported, since PanNENs show the greatest conspicuity on arterial phase images, it was established to contour tumors in this phase only. Then, a registration, based upon a region of interest comprising the pancreas, was performed between arterial and basal CT images; finally, the contours were transferred on the co-registered unenhanced CT images, and manually adjusted for minor inconsistencies possibly due to respiration artifacts and organ motion between the two phases.

The choice of unenhanced images for radiomic features extraction was due to the fact that contrast medium administration could modify tissue heterogeneity with respect to the intrinsic inter-patient variability of contrast administration.

RFs extraction [**The following paragraph contains methodology that has been already published – Palumbo *et al*, DOI: 10.3390/cancers13194938**] – All images were resampled at cubic voxels of 0.78x0.78x0.78 mm^3 with bilinear interpolation using an automatic workflow expressly developed in commercially available software (MIM Software Inc., version 6.5.5). This procedure was implemented to reduce directional bias, according to the specific recommendation of the International Biomarker Standardization Initiative (IBSI) (Zwanenburg *et al*, 2020; Palumbo *et al*, 2021b). Image rebinning was also necessary, not only to speed up the process of RFs extraction, but also to limit noise: we chose 64 bins, as reported in literature (Benedetti *et al*, 2021). DICOM files were then imported into MATLAB using the Computational Environment for Radiological Research. 182 RFs of first and higher order were extracted using SPAARC Pipeline for Automated Analysis and Radiomics Computing (IBSI complying).

Statistical analysis [**The following paragraph contains methodology that has been already published – Palumbo *et al*, DOI: 10.3390/cancers13194938**].

As previously stated, the original population was randomly split into training (n = 70) and validation (n = 31) cohorts.

Methodological details regarding RFs redundancy limitation have been already published (Palumbo *et al*, 2021b).

Models development – To assess the best combination of the previously selected clinical, radiological and radiomic variables, a machine learning bootstrap-based method was used. In short, the training set of data was bootstrapped 1000 times and a backward multivariate logistic regression was run for each sample including two (for G and M+) or three (for N+ and VI) variables according to the number of events for each endpoint. Three models were then developed for each endpoint: a "conventional" radiological model, a strictly radiomic model and a combined model considering information from radiomic, conventional radiologic and clinical variables. Finally, a prognostic index was derived according to a previously published formula (Palumbo et al, 2021b).

Models validation – To assess the ability of the prognostic index in stratifying patients according to the histological endpoints, the coefficients of the prognostic index were directly entered into a new univariate logistic regression considering data from validation set. For each model, a $p < 0.05$ was required for considering it validated. Analyses were performed using homemade Matlab codes.

Results

Patients' characteristics

Patients' characteristics are summarized in Table 4. Twenty-five patients (24.7%) had G2/G3 tumor (specifically, three patients only [2.9%] had undifferentiated neoplasm), 37 (36.6%) were shown to have nodal involvement, 14 (13.8%) suffered from distant metastases (mostly in the liver), and 38 (37.6%) had microvascular spread of the disease at pathological specimen. Between training and validation cohorts, no significant differences were found when considering both radiological and clinical variables, nor pathological data. Median tumor volume was 6 cc (range: 1.3-19.9).

Models training and validation

The combination of two variables only (one radiomic and one clinicoradiological feature) resulted in good prediction of the risk of M+ and G with AUC = 0.85 and 0.67, respectively; these results were confirmed in the validation cohort (AUC = 0.77 and 0.72, respectively). The models predicting the risk of IV and N+ (both comprising two radiomic and one clinicoradiological feature) showed AUC = 0.82 and 0.72, respectively, in the training set; these results were confirmed in the validation cohort (0.75 and 0.62, respectively). A pure RF_model could be generated only when considering M+ and G as endpoints, with similar performances of the corresponding COMB_models (AUC=0.81 and 0.68, respectively, in the training cohort, AUC=0.81 and 0.70 in the validation set). A pure "conventional" radiological model failed to be confirmed in the validation set for all endpoints with the only exception of microvascular invasion.

Negative predictive values resulted moderate to high for all validated models for the different endpoints, ranging between 77.8% (G+ COMB_model) and 97 % (M+ RF_model).

The performances of the models (those which successfully passed validation step) for each endpoint are reported in Table 5 and summarized in Figure 8. In Figure 9, ROC curves of the same models are shown.

Discussion

Few studies explored the potential of radiomics in the setting of pancreatic neuroendocrine neoplasms (PanNENs), often showing contrasting results since limited by small sample size issues and potentially biased methodologies (Gu *et al*, 2019; Liang *et al*, 2019; Bian *et al*, 2020a; Bian *et al*, 2020b). In the present study we applied, in a relatively large cohort of patients, a machine learning approach optimized to limit/avoid the risk of overfit; in doing so, we sought to develop and validate preoperative models (including a maximum of three variables) based upon CT images to predict tumor grade (G1vs.G2/3), presence of distant metastasis, metastatic lymphnodes and microvascular invasion at pathological specimen.

The vast majority of the present literature focuses on tumor grade prediction (Gu *et al*, 2019; Liang *et al*, 2019; Bian *et al*, 2020a; Choi *et al*, 2018; Bezzi *et al*, 2021), which is indeed a crucial cornerstone for treatment planning being a surrogate for biological aggressiveness. In this respect, an image-based biomarker able to accurately predict grading could be of great impact, especially when considering patients with small lesions (< 2 cm) which are generally thought to correspond to well-differentiated (G1) tumors to be conservatively managed (Partelli *et al*, 2017a). Nevertheless, these small tumors do sometimes reveal aggressive biological behavior and need a more aggressive approach (Partelli *et al*, 2017b). Importantly, endoscopic ultrasound-guided fine-needle aspiration is not always reliable in determining tumor grading of small tumors (Rebours *et al*, 2015). A possible solution could come from radiomics, since radiomic features (RFs), being derived from the whole volume of the lesion, are paradoxically more representative of the heterogeneity of the entire lesion than histology itself. In the present study we found that coupling one robust RF (*Morphology_areaDensity_aabb*) with one conventional radiological finding (tumor necrosis) resulted in good negative predictive value (77.8%) and area under the curve (0.72) for grade prediction.

These results corroborated our previous findings in a pilot study on 39 patients (Benedetti *et al*, 2021) and highlight the importance of a rigorous radiomic workflow based upon *i)* a strict selection of few robust RFs and *ii)* availability of an independent validation cohort to reduce any risk of overfitting. Other studies tried to avoid this issue by restricting the number of variables (Bian *et al*, 2020a), or splitting the cohort into

training and validation sets (Gu *et al*, 2019; Lian *et al*, 2019); however, with this second approach only, if no attempt is made to achieve optimally robust models in the training group, the performances may significantly reduce in course of validation, and the proposed findings may lack of interpretability. In this respect, our findings are in good agreement with the results obtained by Bian *and colleagues* in a group of 102 PanNEN patients applying LASSO for tumor grade prediction (Bian *et al*, 2020a)**,**.

Apart from G, to our knowledge, few other studies explored the value of radiomics to predict other histological characteristics of panNENs (Mapelli *et al*, 2022, Bezzi *et al*, 2021).

Accurate preoperative N staging represents indeed a major cornerstone in the treatment algorithm of panNENs, since patients with different N stages have different prognosis and may need a different extent of lymphadenectomy or neoadjuvant treatment; specifically, the number of positive lymph nodes is accurate in predicting recurrence for panNENs after surgery (Guarneri *et al*, 2021). A recent report by Mapelli *and colleagues* (Mapelli *et al*, 2022) found that second-order RFs extracted from T2 Magnetic Resonance sequences have good predictive performance (AUC = 0.992) with respect to lymph nodal involvement. Our model has lower accuracy, but with the clear advantage of relying upon RFs derived from CT images, which are, actually, the standard of care when dealing with non-invasive staging of panNENs.

Our models also show good results with respect to prediction of distant metastases and microvascular invasion, providing a further insight into disease biological behaviour; these findings are indeed in agreement with previous literature (Kulali *et al*, 2018).

Interestingly, the degree of clinical/biological interpretability of the RFs finally retained in each validated model is promising, being related to tumor irregular and/or compact shape (as frequently observed in the setting of pancreatic adenocarcinoma (Palumbo et al, 2021b) and/or HU values heterogeneity. With regard to this last point, our group recently found an intriguing explanation connecting, in non-functioning panNENs, microvessel density, radiological appearance in terms of HU values and biological behavior: in short, low microvessel density (assessed by CD34^{+} staining), corresponding to ipoenhancement in arterial phase, has been found to be associated with pathological features of aggressiveness (Battistella *et al*, 2022).

Finally, to objectify the incremental value of radiomics with respect to radiologists' subjective assessment (an issue already mentioned in the existing literature (Partelli *et al*, 2022)) we demonstrated that a model based upon, specifically, conventional radiological parameters failed to be confirmed in the validation set.

The present study has several limitations, the most important ones being its retrospective nature and the relatively small number of events observed. External validation is also warranted and already planned.

In conclusion, despite the above-mentioned limitations, the combination of few radiomic and clinicoradiological features by means of robust methodology that avoid/limit the risk of overfit resulted in robust presurgical prediction of histological characteristics of PanNENs, potentially providing a tool for patients' personalized management, once more extensive external validation will be accomplished.

Clinicoradiological Variables	Training (n=70)	Validation (n=31)
Age*	58 (18-77)	55 (26-84)
Gender (Male)	28 (40.0%)	11 (35.5%)
Presence of necrosis	15 (21.4%)	7 (22.6%)
Cystic morphology	15 (21.4%)	12 (38.7%)
Pancreas atrophy	16 (22.8%)	4 (12.9%)
Arterial invasion	11 (16%)	4 (12.9%)
Venous invasion	21 (30%)	4 (12.9%)
Contiguous organs invasion	7 (10.0%)	6 (19.3%)
Endpoint		
High Grade (G2/G3)	17 (24.3%)	8 (25.8%)
Liver Metastasis (M+)	11 (15.7%)	3 (9.6%)
Microvascular Invasion (VI+)	28 (40.0%)	10 (32.2%)
Metastatic Lymphnodes (N+)	26 (37.1%)	11 (35.4%)
Unless otherwise indicated, data are numbers of patients and data in parentheses are percentages *Data are *medians*, data in parentheses are *interquartile ranges* (IQR)		

Table #4: *Patients' characteristics*.

Table #5: *Overall performance of pure radiomic and clinicoradiological/radiomic models.*

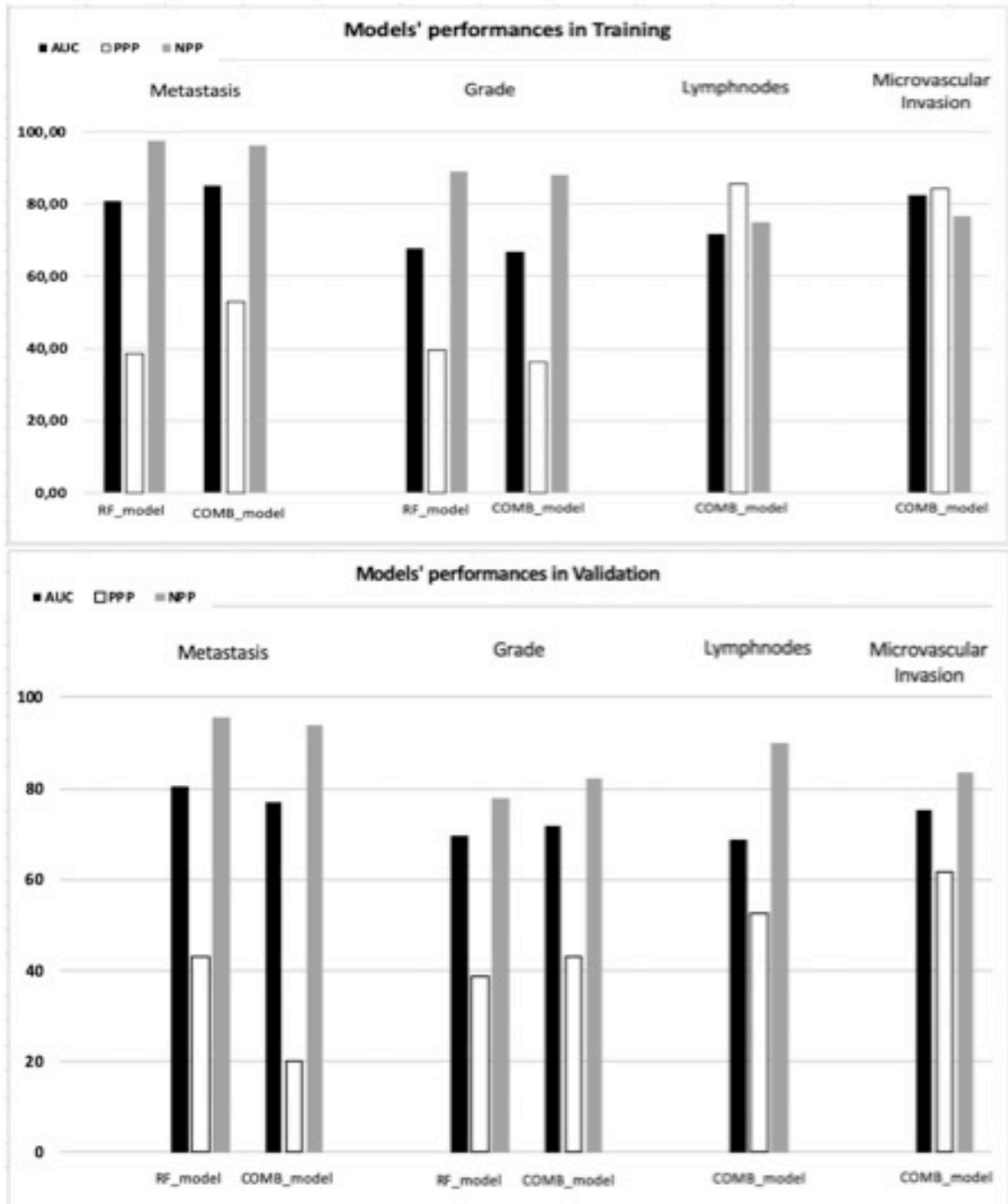|		TRAINING										VALIDATION								
model	COVARIATE	b	Overall fit model p	AUC	p-value	SE	SP	PPV	95% CL	NPV	95% CL	Overall fit model p	AUC	p-value	SE	SP	PPV	95% CL	NPV	95% CL
Metastasis																				
RF_model Training	costant	7.16044																		
	NGTDM_Strenght	-0.39808	0.0057	80.89	<0.001	90.91	72.88	38.50	28.30-49.70	97.70	86.80-99.60	0.0438	80.56	0.007	75.00	85.19	42.90	20.50-68.60	95.80	80.70-99.20
	Morphology_areaDensity_aabb	-12.49944																		
COMB_model Training	costant	-3.85057																		
	ART. INVASION	1.50463	0.0012	85.21	<0.001	81.82	86.44	52.90	35.80-69.80	96.20	87.90-98.90	0.0529	76.85	0.032	75.00	55.56	20.00	11.20-33.60	93.70	72.70-98.90
	Morphology_COMshift	0.04318																		
Grade																				
RF_model Training	costant	-6.62015																		
	Morphology_areaDensity_aabb	8.22075	0.0892	67.70	0.011	76.47	62.26	39.40	29.60-50.10	89.20	77.30-95.20	0.073403	69.70	0.058	55.56	63.64	38.50	21.90-58.30	77.80	61.20-88.60
	Morphology_COMshift	8.61850																		
COMB_model Training	costant	-7.18176																		
	NECROSI	1.13076	0.0426	66.93	0.021	76.47	56.60	36.10	27.40-45.90	88.20	75.5-94.80	0.075041	71.72	0.028	6.67	63.64	42.90	26.70-60.70	82.40	63.70-92.50
	Morphology_areaDensity_aabb	9.22230																		
Lymphnodes																				
COMB_model Training	costant	1.84009																		
	Intensity Histogram_mode	-0.03828																		
	Intensity Histogram_quartileCoefDispersion	-143.86127	0.0196	71.81	0.001	46.15	95.45	85.70	59.30-96.10	75.00	67.60-81.20	0.0934	61.80	0.250	90.91	50.00	47.60	38.30-61.70	90.90	59.50-98.60
	ATROPHY'	1.26017																		
Microvascular Invasion																				
COMB_model Training	costant	1.11448																		
	GLDZM3D_zoneDistanceNonUniformityNorm'	-5.76052																		
	Intensity Histogram_meanAbsoluteDeviation	-0.02732	<0.001	82.40	<0.001	57.14	92.86	84.20	63.10-94.30	76.50	67.80-83.40	0.0181	75.11	0.006	72.73	75.00	61.50	40.80-78.80	83.30	64.80-93.10
	VEN. INVASION'	1.98427																		

Figure #8: *Summary of performances of the models for each endpoint (training and validation).*
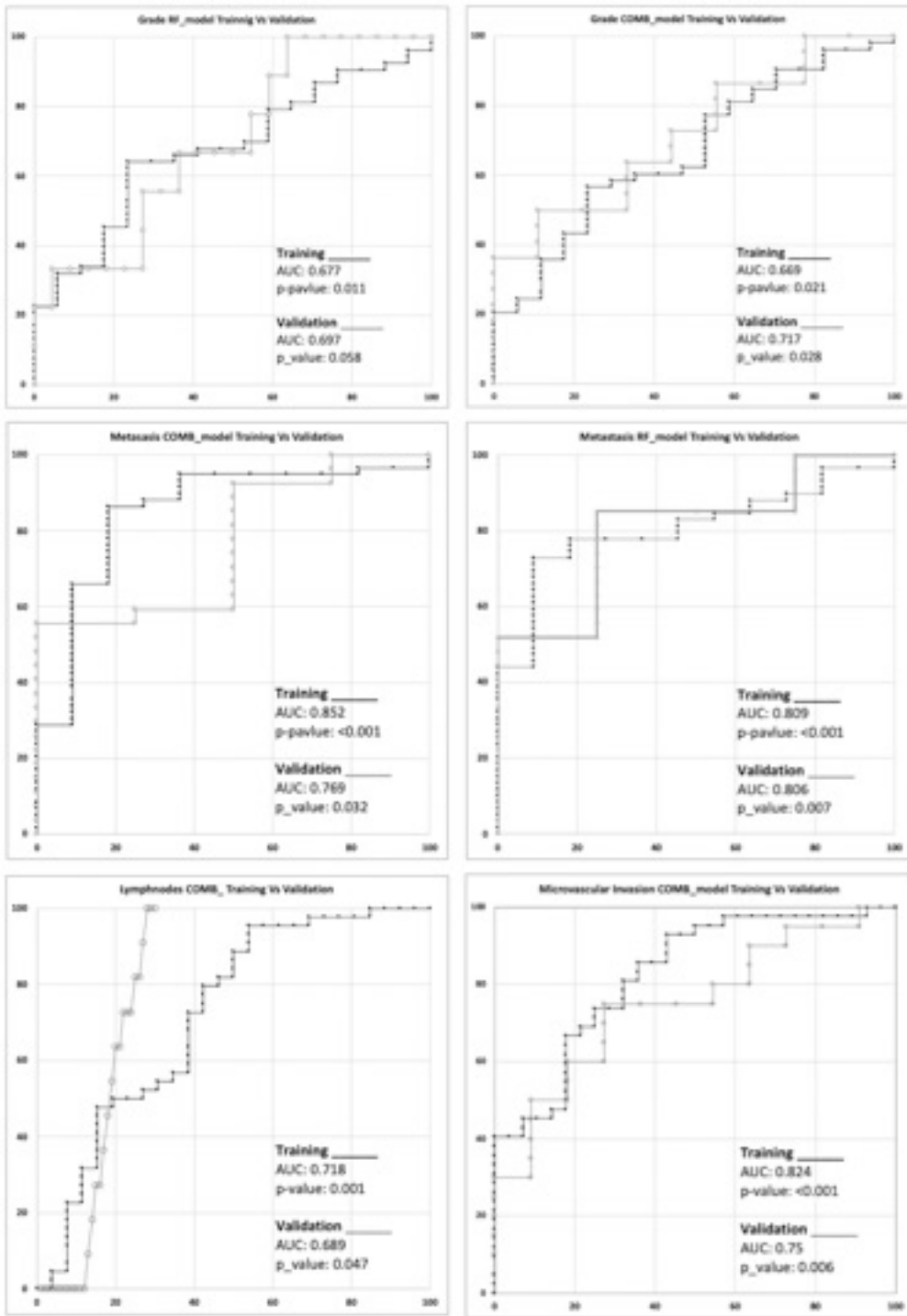
Figure #9: *ROC curves of the models (training and validation).*

Dynamic change of radiomic features within tumour environment predicts pathological response to neoadjuvant chemotherapy and disease relapse of pancreatic adenocarcinoma.

Materials and Methods

Patients' cohort

This is a single-centre retrospective study conducted at San Raffaele Scientific Institute (Milan, Italy); data were collected within the context of an Ethics Committee approved study (06/INT/2022). From a prospectively acquired database, all consecutive adult patients with pancreatic adenocarcinoma (clinical stage I-III, according to TNM 8th edition (van Roessel *et al*, 2018); resectable or borderline resectable, as anatomically defined according to 2021 [*National Comprehensive Cancer Network*] NCCN guidelines (Tempero *et al*, 2021)) who underwent pancreaticoduodenectomy (PD) after neoadjuvant chemotherapy (nCT) between January 2016 and December 2020 were identified (*n* = 330); patients who were evaluated at baseline (before nCT start) and within 30 days before index surgery (after nCT completion) with multiphase, contrast enhanced CT scans at our institution (*n* = 108) were enrolled into our study. Patients who died within 90 days after index surgery (*n* = 3) were excluded from further analysis; moreover, 22 patients had no adequate follow-up and were consequently also excluded.

Endpoints definition

Three endpoints were considered, one clinical (*disease recurrence* after index surgery [any type of], defined at follow-up imaging) and two pathological (*N2 status* [≥ 4 regional lymph-nodes with metastases] and *pathological response to nCT* [Hartman Tumor Regression Grade, TRG (Cacciato Insilla *et al*, 2020) – where "marked response" stands for pathological complete response in terms of no residual cancer cells or small group of cancer cells within a prevalent fibrotic stroma]). With specific regard to this last endpoint, patients' cohort was then divided into good (moderate to marked pathological response) *vs*. poor responders (minimal to poor/absent pathological response).

Multidisciplinary evaluation, neoadjuvant regimens and follow-up data collection

A dedicated multidisciplinary team evaluated radiological, biological (with particular regard to CA 19.9 serum levels after eventual jaundice palliation, where needed) and conditional (performance status and comorbidities) information of the enrolled patients (Petrelli *et al*, 2017) and had deemed all of them as resectable/borderline resectable, according to the 2021 NCCN guidelines (Tempero *et al*, 2021), and potentially eligible for surgical resection after nCT. Neoadjuvant chemotherapy regimens included PAXG, gemcitabine plus abraxane or modified FOLFIRINOX. Resectability status was systematically reassessed after nCT completion; in those patients who were confirmed to have resectable disease after nCT, surgical treatment (PD with lymphadenectomy) has been performed four to six weeks after completion of neoadjuvant treatment, with possible variations depending upon each patient's performance status. Finally, after surgery, all data collected were systematically correlated with the pathological outcome in terms of TRG, which is based on the percentage of viable residual neoplastic cells in comparison with fibrosis/necrosis after treatment, and nodal involvement. Adjuvant treatment (chemoradiotherapy *vs.* chemotherapy alone) was considered depending on pathological information (TNM staging, grading, eventual presence of perineural invasion, …). Follow-up imaging examinations were accurately evaluated looking for disease relapse.

CT protocol and radiomic features extraction

Methodological details regarding CT protocol, radiomic features extraction and redundancy lowering have been already published (Palumbo *et al*, 2021b). 182 radiomic features (RFs) were extracted from late arterial phase of both baseline and post treatment CT scans. For each RF, Δ value was computed as the difference between the absolute value calculated at the two time points, according to the formula: $\Delta\ RF = (RF_{post}) - (RF_{pre})$. In short, the longitudinal net changes in RFs were calculated.

Lesion delineation on CT images – The robustness against interobserver contouring variability was preliminarily assessed focusing on a subgroup of 30 patients whose baseline CT scans were evaluated by three independent readers; only RFs with ICC > 0.80 among these three readers were considered robust (and potentially reproducible)

and therefore retained for subsequent analysis. The contouring has been systematically carried out on late arterial phase, where tumor conspicuity is the most (Zins *et al*, 2018).

Statistical Analysis – After the selection of the most significant RFs (at both timepoints imaging, separetly) and Δ RFs (between timepoints imaging), a machine learning bootstrap-based method was applied to develop radiomic models for *i*) recurrence after index surgery, *ii*) pathological response to neoadjuvant treatment (TRG) and *iii*) pathological N2 status. Specifically, for each endpoint three models (*a*) pre nCT, *b*) post nCT and *c*) delta radiomics) were constructed.

The original population of 83 patients was bootstrapped 1000 times to produce different samples with the same informative content of the original population. A backward multivariate Cox analysis was performed for the prediction of recurrence, and a backward logistic regression for the prediction of the other two endpoints in order to find the best combination of RFs and Δ RFs with the best AUC and p-values (< 0.05).

A prognostic index (P index) was then derived for each one of the three radiomic models. In the case of logistic regression, the PI was given by the risk computation: $PI = B_0 + \sum_{i=1}^{n} B_i x_i$. For the Cox analysis the PI was given by the natural logarithm of the Hazard Ratio: $PI = -\ln \frac{H(t)}{H_0(t)} = \sum_{i=1}^{n} B_i x_i$. In both cases B_i coefficients are given applying the logistic and Cox regressions on the original population, whereas X_i represents the values of variables selected by the bootstrap method.

Model performance assessment [**The following paragraph contains methodology that has been already published – Palumbo *et al*, DOI: 10.3390/cancers13194938**] – To assess the ability of the P index in stratifying patients according to the risk of recurrences, the P index was dichotomized as greater or smaller than the cut-off value (Palumbo *et al*, 2021); finally, the separation of the survival curves of the two groups with high and low risk of recurrence was tested with a Kaplan-Meier test. Regarding TRG and N2, the ability of P index to stratify the risk classes was tested with box-plots. The performances of the three models were quantified using the P index in terms of: area under the ROC curve (AUC), positive and negative predictive values (PPV, NPV), specificity and sensitivity. Analyses were performed using homemade Matlab codes.

Results

Patients' characteristics

After exclusions, 83 patients (31 female [37.3%]) with a median age of 65.5 years [range, 39 – 85]) were enrolled in this study.

The median overall survival and progression-free survival were 21 (IQR: 14 – 31) and 11 (IQR: 5 – 20) months, respectively. 48 patients out of 83 (57.8%) had a disease recurrence during the follow-up time period (median follow-up time: 21 months, (IQR: 15 – 31); of these, 20 (41.6%) experienced an early disease recurrence (occurring < 11 months after index surgery). With respect to pathological endpoints, 14 patients (16.8%) were found to have N2 disease, and 63 (75.9%) demonstrated moderate to marked response.

Pre nCT radiomic models (*Table 6*)

Amongst the most frequent baseline (pre nCT) RFs resulting from the machine learning bootstrap-ranking procedure, three RFs were retained in the model dealing with disease recurrence ($p < 0.001$, AUC = 0.79). Corresponding Kaplan-Meier survival curve is shown in Figure 10, demonstrating that the proposed model have poor performance in stratifying the risk of early disease relapse after surgery (also having regard to a low NPV of 63%), with significant overlap in terms of recurrence chance at 10 months between *high-* and *low-risk* patients. Also pre nCT models for the prediction of N2 disease and pathological regression grade consisted of three RFs only; they demonstrated moderate to high performance (p = 0.0002 and 0.0119, AUC = 0.84 and 0.77, respectively [Figures 11 and 12, respectively]) with considerable NPV (96.7% and 95.2%, respectively).

Post nCT radiomic models (Table 7)

Four post nCT RFs contributed to the realization of the presurgical model for the prediction of disease recurrence after resection (p = 0.003, AUC = 0.78). Corresponding Kaplan-Meier survival curve is shown in Figure 13, demonstrating an overall poor performance in distinguishing the risk of early disease relapse after surgery between *high-risk* (having a 32% recurrence chance at 10 months) and *low-risk* patients (having a 20% recurrence chance at 10 months) (p > 0.05). Both post nCT models for the prediction of N2 disease and pathological regression grade consisted instead of three

RFs only; they demonstrated moderate to high performance (p = 0.079 and 0.0005, AUC = 0.80 and 0.80, respectively [Figures 14 and 15, respectively]) with significant NPV (94.2% and 92.2%, respectively).

Delta radiomics models (Table 8)

Amongst the most frequent longitudinal net changes resulting from the machine learning bootstrap-ranking procedure, four Δ RFs were retained in the final model dealing with disease recurrence (p = 0.0045, AUC = 0.71). Corresponding Kaplan-Meier survival curve is shown in Figure 16. In this respect, the proposed model showed a good performance in stratifying the risk of early disease relapse after surgery by identifying, with a good NPV of 78.3% and a hazard ratio of 2.91, *high-risk* patients with a recurrence chance of 40% at 10 months (*vs. low-risk* patients with a chance of 5% [p < 0.001]). Three highly selected, highly robust Δ RFs took part instead in the construction of the presurgical models for the prediction of the other two endpoints (N2 disease and pathological regression grade), which both demonstrated moderate to high performance (p = 0.0010 and 0.0087, AUC = 0.82 and 0.72, respectively [Figures 17 and 18, respectively]) with considerable NPV (93.9% and 88.7%, respectively).

Discussion

Growing literature suggests that even early stage pancreatic ductal adenocarcinoma (PDAC) (clinical stage I-III, according to TNM 8th edition (van Roessel *et al*, 2018); resectable or borderline resectable, as anatomically defined according to 2021 [*National Comprehensive Cancer Network*] NCCN guidelines (Tempero *et al*, 2021)) could benefit from a neoadjuvant treatment (Lee *et al*, 2019; Reni *et al*, 2019a; Reni *et al*, 2019b); the rationale of such an approach lays in the observation that recurrence occurs in almost 90% of these patients within two years after index surgery, implying that even early stage disease is, actually, a micro metastatic one. It follows that since most patients have (micro) metastatic disease at the time of diagnosis, upfront surgery, only debulking the overall tumour cell burden, is less effective in providing disease control when compared to chemotherapy, which instead potentially prevents both local and distant tumour progression by reducing the number of neoplastic cells in their exponential growth stage (Haeno *et al*, 2012). Both AIOM (*Associazione Italiana Oncologia Medica*) and NCCN guidelines now include neoadjuvant chemotherapy (nCT) as a possible treatment option for resectable disease. However, a major radiological issue is currently undermining such an approach; several authors have recently claimed that neoadjuvant chemotherapy/radiotherapy might impair the diagnostic accuracy of CT imaging in the evaluation of tumour response, mostly due to abundant desmoplastic reaction mimicking cancer growth (Cassinotto *et al*, 2014; Park *et al*, 2021). There are indeed no reliable criteria to assess treatment response of PDACs after nCT, ultimately hampering the chance for an optimal therapeutic path.

Most authors have sought possible solutions to this diagnostic dilemma focusing on tumour morphological (*i.e.* RECIST [*Response Evaluation Criteria In Solid Tumours*]) (Park *et al*, 2021) and/or densitometric changes (Marchegiani *et al*, 2018; Wagner *et al*, 2016) occurring in course of nCT, but with inconsistent (generally poor, mostly in terms of low specificity) results, ultimately contributing to overstage issues. Other authors (Jeon *et al*, 2022; Noda *et al*, 2022; Jang *et al*, 2021) shifted their focus on the extent of tumour contact with adjacent vessels: in this respect, they found even a minimal regression (also not significant according to RECIST criteria and/or 2019 NCCN guidelines) being a reliable predictor of R0 resection. However, *Jeon and colleagues* (Jeon *et al*, 2022) found that this assumption can be deemed correct only when referring

to pancreatic neoplasms classified (after nCT) as resectable or locally advanced; borderline resectable tumours, which represent indeed one of the major surgical issues, escape this criterion. On that note, it is worth noting that there are still few studies claiming a role for radiomics to unsolve this matter, almost all relying on baseline imaging (before nCT start) for features extraction (Elsherif *et al*, 2022; Borhani *et al*, 2020); in this respect, a possible ethical concern exists: baseline stratification tools rarely rule out patients from undergoing *standard of care* treatment, unless demonstrating almost perfect accuracy in identifying responder/non responder status.

To address this challenge, we applied a machine learning approach to develop three clusters of models based upon, respectively, *a*) pre-/ *b*) post-nCT and *c*) delta radiomics. To the best of our knowledge, this is the first manuscript PDAC-orientated exploring, at once, the contribution of radiomic features (RFs) derived at different time points together with their dynamical change in course of nCT. Of note, our data support the predominant role of Δ RFs over "raw" ones in stratifying the risk of early disease relapse after surgery by identifying *high-risk* patients with a recurrence chance of 40% at 10 months (*vs. low-risk* patients with a chance of 5%). On the opposite, when considering pathological endpoints (namely, *pathological response to nCT* and *N2 status*) no significant differences were found among these three approaches in terms of accuracy (overall good NPVs, together with poor PPVs).

Taken together, these results could have huge clinical consequences:

- After scheduled nCT completion, knowing whether the index lesion had a good (*vs.* poor) response in terms of viable residual neoplastic cells could be a crucial information to be taken into account during multidisciplinary team discussion for surgical candidacy *vs.* switching to another chemotherapy regimen. Similar considerations can be made regarding nodal status, which is thought to be a major determinant for disease recurrence after surgery (Crippa *et al*, 2021): having available a sort of radiomic presurgical "advance" of the final pathological report shortly after nCT completion (by applying post-nCT/Δ radiomics) could be indeed a game changer in the decision-making algorithm of PDAC patients.

- On the opposite, thinking of a possible role for pre-nCT radiomics in the PDAC real clinical setting is far more challenging. As stated before, baseline

stratification tools rarely rule out patients from undergoing *standard of care* treatment, unless demonstrating almost perfect accuracy in identifying responder/non responder status; on that note, there is a long way off for our pre-nCT models from being 100% reliable. One may speculate that baseline information could be potentially useful for selecting the most appropriate nCT regimen for each patient (an issue currently explored at our Institution within the CASSANDRA trial [PACT-21 protocol]), also having regard to possible germline pathogenic variants (Macchini *et al*, 2021); however, our data do not support this fascinating hypothesis and further studies are needed to prove/refute it.

• That said, what comes up from our results is the predominant prognostic role of Δ radiomics over "raw" pre-/post RFs; the clinical significance of the neoplastic environment dynamical change in course of nCT appears indeed significantly superior when compared to its stationary consideration, a concept already used in the evaluation of biological markers (Reni *et al*, 2020). This assumption highlights the inherent limitations of the present restaging strategies, mostly relying on "single frame" analysis.

Significant strengths of the present study are *i*) its methodological robustness based upon a previously validated machine learning approach optimized to limit/avoid the risk of overfit when selecting RFs for model development (Palumbo *et al*, 2021b), *ii*) the effort for including each radiomic model into a real clinical setting emphasizing its possible contribution, and *iii*) the choice of biologically/clinically significant endpoints. With regard to this last point, it is worth noting that the vast majority of the existing (mostly radiological) PDAC-orientated literature dealing with the identification of novel markers of treatment response after nCT selects as major endpoint the status of resection margins (Jeon *et al*, 2022; Noda *et al*, 2022; Jang *et al*, 2021), which is indeed a surgical variable highly dependant on surgeons' experience, rather than more objective parameters of treatment response (as, for instance, TRG), resulting in methodological biases possibly undermining presented results. Finally, we would like to highlight the biological significance of the RFs finally retained in the Δ models, mostly belonging to the morphological family of features. Specifically, the assumption that the

longitudinal positive net change of *Surface to Volume ratio* strongly predicts ≥ 4 regional lymphnodes with metastases is somehow striking, as meaning that it implies that tumour complexity (in terms of spiculatedness) is a major determinant of prognosis, even superior to tumour volume; in other words, a reduction of tumour size in course of nCT should be considered cautiously, since it could hide a modification of tumour contours towards spiculatedness.

Limitations also exist, the main ones being the relative small sample size (although highly selected) and the absence of a validation cohort, which is indeed already planned. Another more subtle, yet clinically crucial limitation has to be taken into account: overall, the three clusters of radiomic models referring to both pathological endpoints are clearly hampered by low PPVs (as low as 35.5%), which counteract very good NPVs (up to 96.7%). It follows that the presented models are very good at identifying patients with good pathological response and no nodal metastases (*low-risk* patients) but often fail in identifying poor responders to nCT (*high-risk* patients). Such an imbalance could ultimately lead to incorrect exclusion of *low-risk* patients from surgical chance; to limit/avoid this risk our models have to be embedded in the *standard of care* multidisciplinary assessment of the patient.

In conclusion, our results point out the crucial significance of Δ radiomics as a marker of treatment response for PDAC undergoing nCT, providing substantial aid towards personalized therapeutic path and, hopefully, prognosis improvement.

PRE		Coefficient	P-value	OR	95% CI	Overall Fit Model	AUC	95% CI	Sensivity	Specifity	PPP	NPP
RECURRENCE	Morphology_COMshift	8.5077	0,1445	4952.5558	0.0540 to 454440882.1346	<0.0001	0.799	0.696 to 0.879	64.6	82.9	83.8	63.0
	Intensity_Histogram_minGradientGreyLevel	-0.1016	0,0002	0.9034	0.8569 to 0.9523							
	GLSZM3D_largeZoneLowGLEmphasis	-0.007185	0,0379	0.9928	0.9861 to 0.9996							
	GLDZM3D_smallDistanceLowGreyLEmphasis	-592.8695	0,0239	3.312E-258	0.0000 to 9.2431E-35							
TRG	Intensity_Histogram_skewness	-2.03172	0,0213	0.1311	0.0233 to 0.7387	0.0119	0.772	0.666 to 0.858	88.9	63.5	41.0	95.2
	GLCM3D_avg_infoCorrelation1	19.77308	0,0412	3.87E+15	2.2043 to 67.8273E+15							
	GLDZM3D_largeDistancehighGreyLEmphasis	-0.00018468	0,0108	0.9998	0.9997 to 1.0000							
N2	Morphology_elongation	-4.63098	0,1684	0.0097	0.0000 to 7.0925	0.0002	0.842	0.745 to 0.913	85.7	84.1	52.2	96.7
	GLCM3D_comb_angularSecondMoment	272.57295	0,0132	2.38E+118	6.1741E+24 to 91.8926E+210							
	GLSZM3D_zoneSizeEntropy	7.64596	0,0015	2092.1809	18.5262 to 236271.8503							

Table #6: *Overall performance of pre nCT radiomic models..*

POST		Coefficient	P-value	OR	95% CI	Overall Fit Model	AUC	95% CI	Sensivity	Specifity	PPP	NPP
RECURRENCE	Statistical_skewness	0.8447	0.0165	2.3273	1.1665 to 4.6432	0.0003	0.782	0.677 to 0.865	81.3	68.6	78.0	72.7
	GLCM3D_avg_differenceAverage	0.6537	0.0195	1.9226	1.1107 to 3.3278							
	GLSZM3D_zoneSizeNonUniformityNorm	-17.3811	0.0411	0	1.6132E-15 to 0.4958							
	GLSZM3D_zoneSizeEntropy	1.2137	0.1412	3.3661	0.6683 to 16.9553							
TRG	Morphology_COMshift	-38.2703	0.0469	2.3956E-17	9.6142E-34 to 0.5969	0.0005	0.805	0.702 to 0.885	77.8	74.6	46.7	92.2
	Morphology_areaDensity_aabb	-10.6096	0.0137	0	5.3474E-9 to 0.1139							
	GLDZM3D_largeDistancelowGreyLEmphasis	-92.0242	0.1291	1.0824E-40	2.6259E-92 to 446E+009							
N2	GLCM3D_avg_correlation	-15.5953	0.0074	0	1.8513E-12 to 0.0154	0.0079	0.801	0.699 to 0.881	73.5	71.0	35.5	94.2
	GLRL3D_avg_runEntropy	10.54151	0.0028	37854.5487	37.3718 to 38343579.5774							
	GLRL3D_comb_shortRunEmp	-16.8653	0.0024	9.2384E-73	2.7092E-119 to 3.1502E-26							

Table #7: *Overall performance of post nCT radiomic models.*

DELTA		Coefficient	P-value	OR	95% CI	Overall Fit Model	AUC	95% CI	Sensivity	Specifity	PPP	NPP
RECURRENCE	Morphology_minorAxisLength	0.1091	0.0174	1.1152	1.0193 to 1.2202	0.0045	0.707	0.596 to 0.801	89.6	51.4	71.7	78.3
	Intensity_Histogram_kurtosis	-0.2967	0.0052	0.7432	0.6035 to 0.9153							
	GLSZM3D_zoneSizeEntropy	-2.2843	0.0011	0.1018	0.0259 to 0.4012							
	GLDZM3D_zoneDistanceVariance	-0.1317	0.011	0.8766	0.7920 to 0.9703							
TRG	Morphology_areaDensity_aabb	7.20016	0.0696	1339.6388	0.5618 to 3194705.2763	0.0087	0.723	0.612 to 0.817	66.7	74.6	42.9	88.7
	GLSZM3D_smallZoneHighGLEmphasis	-0.002541	0.0287	0.9975	0.9952 to 0.9997							
	GLDZM3D_smallDistanceLowGreyLEmphasis	-358.881	0.035	1.38E-156	0.0000 to 1.027E-11							
N2	Morphology_surfAreaToVolumeRatio	4.68784	0.0859	108.6183	0.5160 to 22864.3232	0.0010	0.823	0.724 to 0.898	71.4	89.9	58.8	93.9
	Morphology_COMshift	22.79572	0.0487	7940000000	1.1338 to 55.6642E+18							
	Morphology_max3Ddiameter	0.1022	0.0201	1.1076	1.0162 to 1.2073							

Table #8: *Overall performance of Δ radiomic models.*

Figure #10: *Kaplan-Meier survival curve stratifying the chance of disease relapse after surgery based upon the prognostic index resulting from pre nCT radiomic model.*
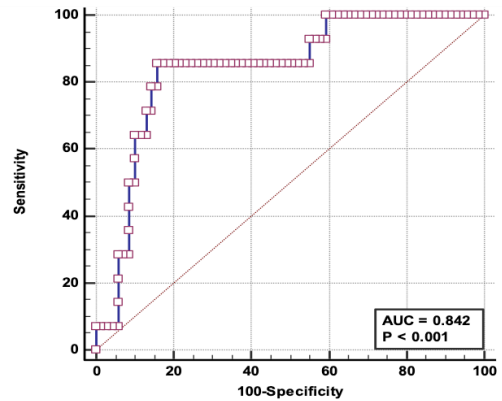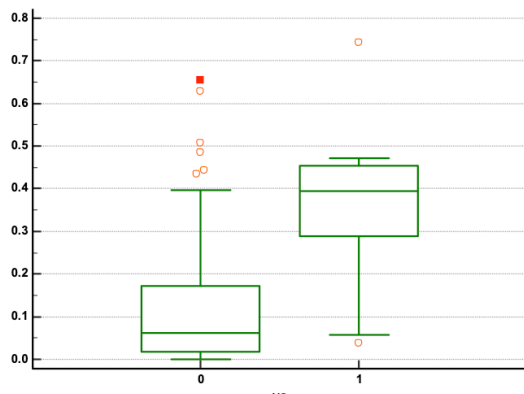

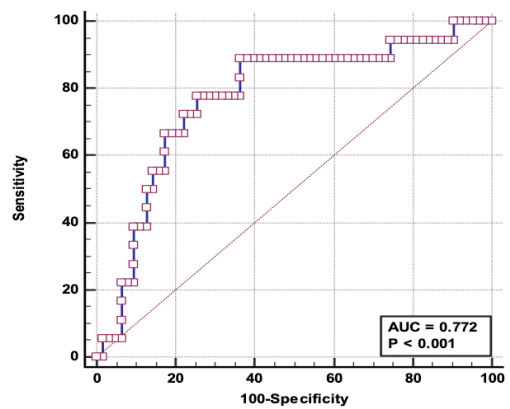**Figure #11**: *Box plot illustrating N2 prediction model (pre nCT radiomic model), and corresponding ROC curve.*
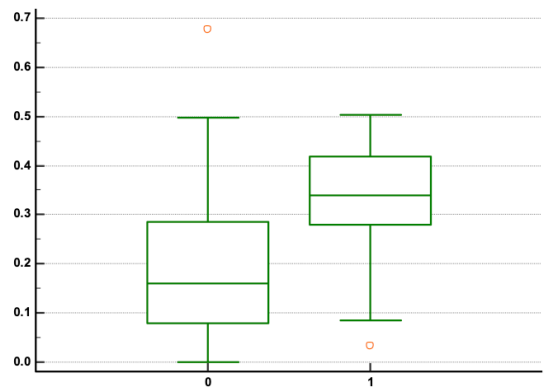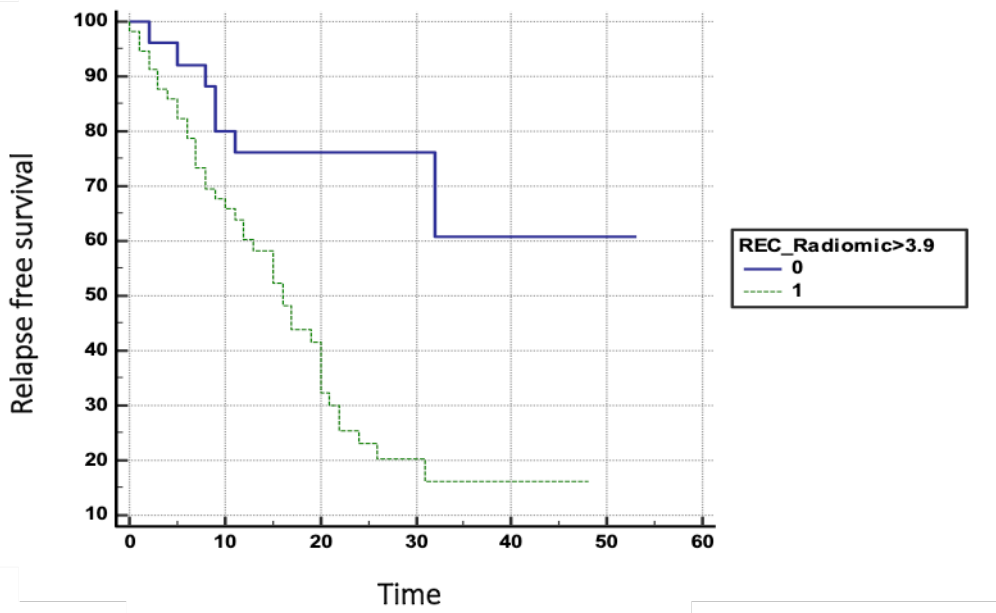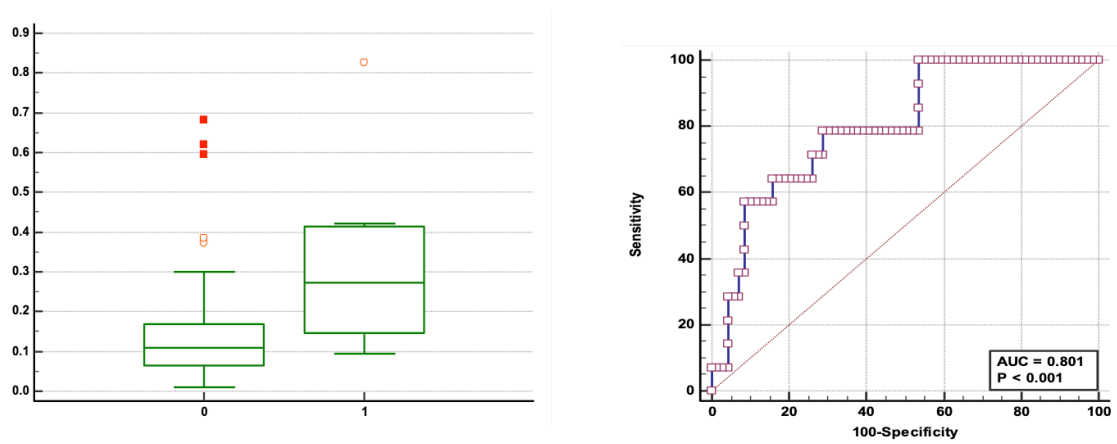

Figure #12: *Box plot illustrating TRG prediction model (pre nCT radiomic model), and corresponding ROC curve.*

Figure #13: *Kaplan-Meier survival curve stratifying the chance of disease relapse after surgery based upon the prognostic index resulting from post nCT radiomic model.*

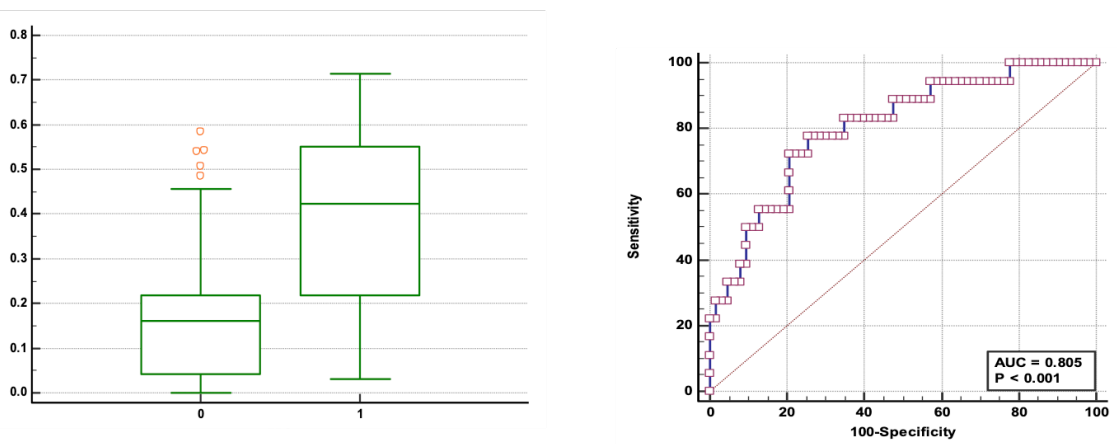**Figure #14**: *Box plot illustrating N2 prediction model (post nCT radiomic model), and corresponding ROC curve.*

Figure #15: *Box plot illustrating TRG prediction model (post nCT radiomic model), and corresponding ROC curve.*
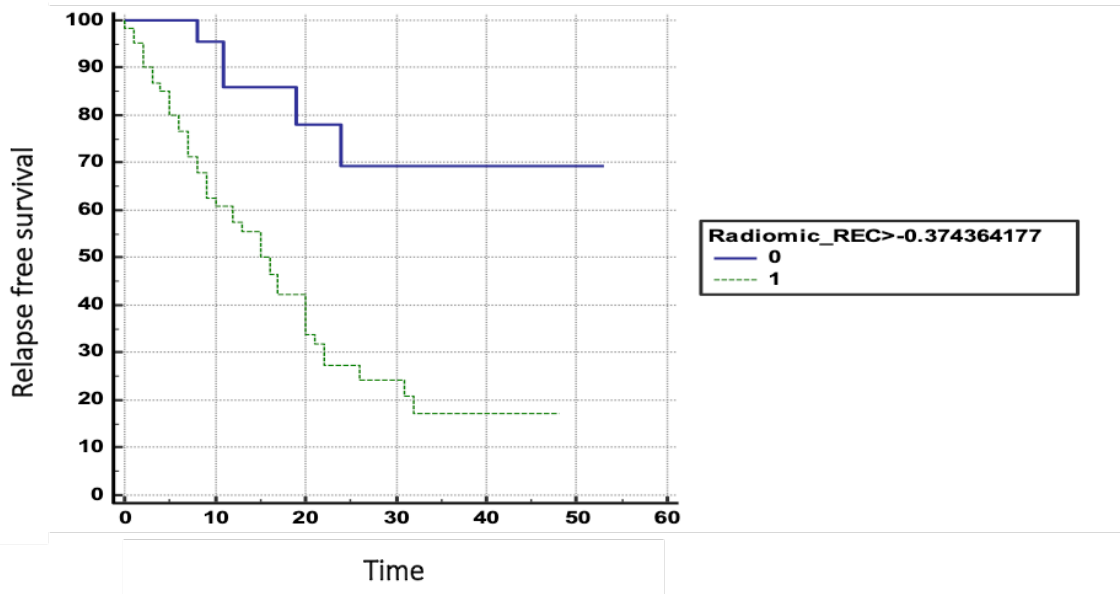

Figure #16: *Kaplan-Meier survival curve stratifying the chance of disease relapse after surgery based upon the prognostic index resulting from Δ radiomic model.*
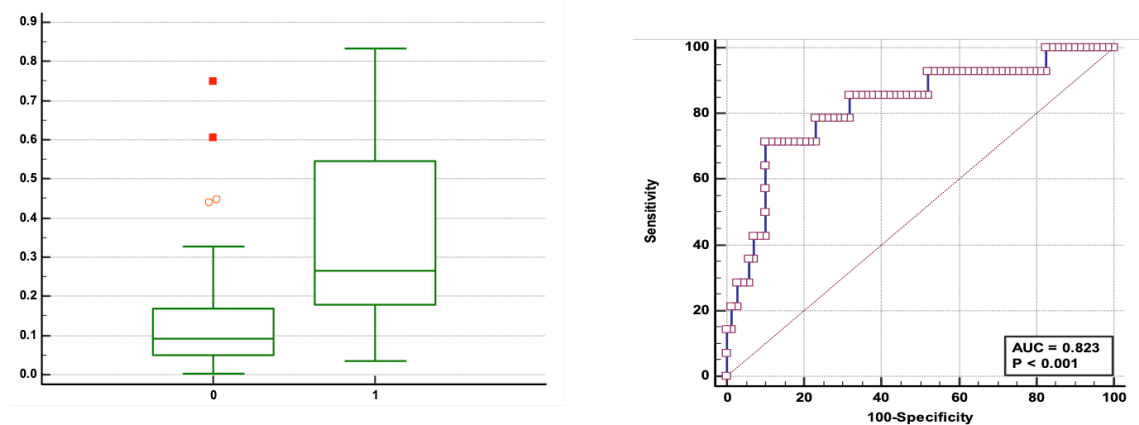

Figure #17: *Box plot illustrating N2 prediction model (Δ radiomic model), and corresponding ROC curve.*
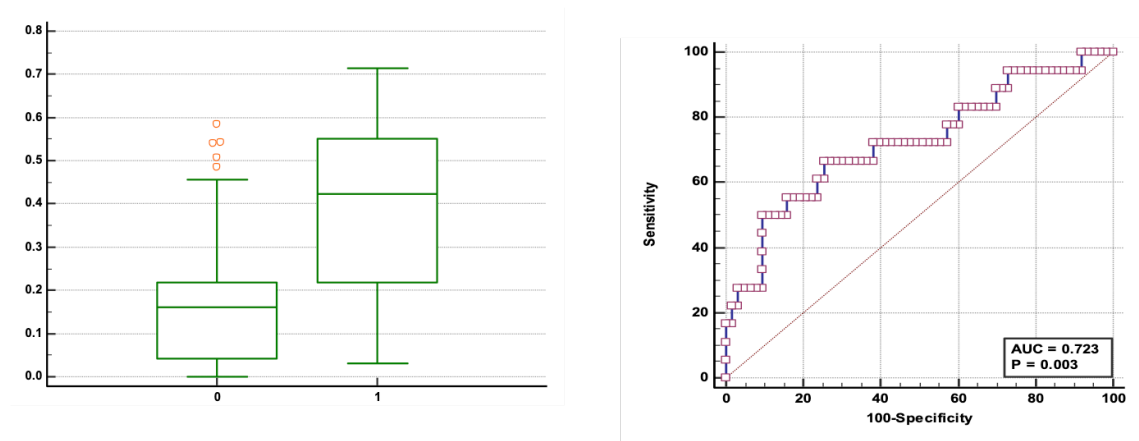
Figure #18: *Box plot illustrating TRG prediction model (Δ radiomic model), and corresponding ROC curve.*

Early assessment of pathological response to neoadjuvant chemoradiotherapy for oesophageal cancer using fully hybrid PET/MR: results from interim analysis.

Materials and Methods

Patients' cohort

This is a single center prospective observational study conducted at San Raffaele Scientific Institute (Milan, Italy), part of a broader investigation funded by AIRC (IG 23105 – "ESCAPE" protocol); data were collected after Ethics Committee approval (07/INT/2020).

From January 2020 to August 2022, all consecutive patients with biopsy-proven potentially resectable oesophageal cancer (adenocarcinoma or squamous cell carcinoma) scheduled to receive neoadjuvant chemoradiationtherapy (nCRT) were prospectively enrolled. In addition to standard of care imaging (De Cobelli *et al*, 2020), a fully integrated hybrid Positron Emission Tomography (PET)/Magnetic Resonance (MR) was performed at three time points. Specifically, in all cases, evaluation has been performed *i)* within two weeks before the beginning of nCRT and then repeated *ii)* after the beginning of the third chemotherapy weekly cycle (that means, between radiotherapy fractions 11-14) and *iii)* at the end of nCRT, within one week before surgery. Study flowchart is summarized in Figure 19.

Neoadjuvant treatment regimen consists of fractioned radiotherapy (41.4 Gy in 23 fractions, delivered 5 days per week) plus concurrent administrations of carboplatin and paclitaxel. In those patients who were confirmed to have resectable disease after nCRT, surgical treatment (Ivor Lewis esophagectomy with lymphadenectomy (Low *et al*, 2019)) has been performed eight to twelve weeks after completion of neoadjuvant treatment, with possible variations depending upon each patient's performance status and response to nCRT. Finally, after surgery, all data collected were systematically correlated with the pathological outcome in terms of Mandard *Tumour Regression Grade* (TRG, 1 to 5) (Mandard *et al*, 1994), which is based on the percentage of viable residual neoplastic cells in comparison with fibrosis/necrosis after treatment (TRG = 1 refers to pathological complete response status) and represents indeed the primary

70

endpoint of this study. Accordingly, patients' cohort was then divided into complete responders (TRG = 1) and partial/non-responders (TRG ≥ 2).

Patients with unexpected distant metastases after nCRT (n = 2), withdrawal from study participation (n = 1) or failure to complete scheduled neoadjuvant treatment (n = 3) were excluded from final study population.

PET/MR protocol

PET/MR images were acquired using a fully hybrid 3T Signa PET/MR system (General Electrics Healthcare – US) and applying a combined protocol already validated in clinical practice at our Institution in other clinical settings (Mapelli *et al*, 2022a; Mapelli *et al*, 2022b; Ironi *et al*, 2022).

The acquisition protocol starts with MR localizers to determine the exact number of "table positions" needed to cover the whole body [18]FDG-PET study. LAVA (LAVA = Liver Acquisition with Volume Acceleration)-Flex sequences are also acquired over the whole body scan for anatomical match of the PET signal. At the end of the whole body scan, a complete MR examination, spatially corresponding to a single PET "table position" (the one including the target lesion), is acquired. MR sequences included: i) axial T2-weighted FSE (Fast Spin Echo) PROPELLER (Periodically Rotated Overlapping ParallEL Lines with Enhanced Reconstruction), ii) axial and coronal/sagittal T2-weighted FSE black blood, iii) axial DWI (b=0, 400, 800 s/mm^2), iv) Dynamic contrast enhanced (DCE) MR T1-weighted SPGR (Spoiled Gradient Recoiled Echo) DIfferential Sub sampling with Cartesian Ordering (DISCO), v) axial post contrast fat saturated T1-weighted. For DCE MR examination, a contrast medium (gadobutrol 0,1 mmol/kg, Gadovist, Bayer) is routinely administered.

Importantly, data from the first five patients enrolled have been used to fine tune the protocol based on a consensus agreement between the involved physicians and confirmed by two expert readers. Two nuclear medicine physicians and two radiologists evaluated all data sets, for PET and MR images, respectively. Inter reader agreement was systematically assessed.

Imaging findings

Analysis of PET and MR images has been performed by using the Advantage Workstation (General Electrics Healthcare – US).

Baseline MR imaging has been compared to those collected at the other two defined time points in terms of percentage change in tumor volume (contoured on axial high-resolution T2 weighted images with a semi automated delineation method allowing for manual editing), apparent diffusion coefficient (ADC) (obtained from ADC map generated from diffusion-weighted images [b=0, 400, 800 s/mm^2]) and perfusion parameters (both semi-quantitative [area under the concentration time curve] and quantitative [volume transfer coefficient reflecting vascular permeability = K_{trans}; flux rate constant = K_{ep}; extracellular volume ratio = V_e]).

On the subject of tumour volume percentage change (measured on both DWI and axial high-resolution T2 weighted images) between scheduled time points, for each patient, an *early regression index* (ERI), originally developed and already externally validated by our own group in the setting of rectal cancer (Cusumano *et al*, 2020; Broggi *et al*, 2020; Fiorino *et al*, 2019; Fiorino *et al*, 2018), was computed according to the following formula, which accounts for the assumed fraction of survival neoplastic cells after the delivery of a certain treatment dose: ERI= ln [1–(tumour volume$_{mid}$/tumour volume$_{pre}$)$^{\text{tumour volume pre}}$]. In short, with this formula each patient is given a probability of having a complete pathological response after completion of neoadjuvant treatment.

PET-derived quantitative parameters obtained at the different time points have been also evaluated after contouring of the entire primary lesion. Specifically, SUV (standardized uptake value)$_{max}$, MTV (metabolic tumor volume) and TLG (total lesion glycolysis) have been systematically calculated.

Statistical analysis

The association of any finding with different types of treatment response was studied univariably. Categorical parameters were compared using the χ^2 test, whereas *Student's t* and the *Mann–Whitney U* tests were used to compare groups with complete *vs.*, partial/non-responders for parametric and non-parametric continuous variables, respectively.

72

Statistical analysis was performed using commercially available software (SPSS version 21.0, IBM Corp.). A *p-value* < 0.05 was considered statistically significant.

Results

Patients' characteristics

After exclusions, 25 patients (9 female [36%]) with a median age of 59.1 years [range, 48-79]) were prospectively enrolled in this study; adenocarcinoma accounted for the most frequent histotype (16/25, 64%). A pathological complete response (TRG = 1) was found in eight patients (32%). No significant differences were found between responder and partial/non-responder patients in terms of baseline demographic data, T stage and pathological subtype.

Imaging findings

The median tumour volume contoured on baseline axial high-resolution T2 weighted images was 19.64 cc (IQR 12.17 –52.24 cc), not significantly different between complete and partial/non-responders (25.19 cc [1.6 – 31.08 cc] *vs*. 39.78 cc [13.96 – 45.72 cc], respectively, p = 0.782). Interestingly, patients with a pathological complete response demonstrated significantly lower median ERI values (measured on axial high-resolution T2 weighted images) when compared with those patients with TRG $\geq$ 2 (6.05 [range, 0.18–14.89] *vs*.19.83 [range, 8.53–37.07], p <0.001). Differently, ERI values calculated on axial DWI images did not obtain any statistical significance (p = 0.388).

The percentage modification in tumour ADC values from baseline to intermediate scans (Δ ADC$_{mid}$) was found to be significantly different between complete and partial/non-responder patients (median increase: 32% vs. 5%, p = 0.002). On the opposite, no differences (nor in terms of increase or decrease of tumour ADC) were observed when looking at the third imaging time point (Δ ADC$_{post}$) (p = 0.185).

When considering the percentage modification in tumour K$_{trans}$ from baseline to intermediate scans (Δ K$_{trans}$) a trend toward statistical significance exists between complete and partial/non-responder patients; specifically, the former have an early drop in K$_{trans}$ values (median Δ = – 30.9%) when compared to the latter (median Δ = – 15.8%) (p = 0.075).

No PET parameter was shown to significantly correlate to TRG; however, patients with TRG = 1 had usually lower values of mean SUV$_{max}$ at baseline imaging (13.2 *vs*. 19.8 in partial/non-responder patients, p = 0.180). This difference almost completely

disappears when looking at the other two imaging time points, possibly because of a sort of treatment-related inflammatory superimposition, which flattens any biological-driven metabolic difference.

Discussion

Primary chemoradiotherapy with (or without) subsequent surgery is currently regarded as the treatment of choice for locally advanced oesophageal cancer. The CROSS trial (van Hagen *et al*, 2012; Shapiro *et al*, 2015) demonstrated indeed that preoperative administration of neoadjuvant chemoradiotherapy (nCRT) doubles the median overall survival of this subset of patients in comparison to surgery alone, with a huge 29% of pathological complete response; many authors suggest that these patients do not benefit from additional surgery, also having regard to the fact that oesophageal resections are associated with substantial morbidity and postoperative mortality rates even in highly experienced centres (Low *et al*, 2019). On the opposite, 18% of patients who underwent nCRT were deemed as being non-responders; these patients do not benefit from nCRT and only suffer from its side effects.

It follows that early identification of tumour response to nCRT is a crucial issue, since it could enable tailored therapeutic plans, avoiding unnecessary treatment efforts and related adverse effects, with major impact on patients' quality of life as well as health care costs.

Currently, response to nCRT is evaluated by using morphologic imaging (De Cobelli *et al*, 2020; Palumbo *et al*, 2020), such as Computed Tomography (CT) and endoscopic ultrasonography (EUS); limitations exist, the main ones being the difficulty in distinguishing viable tumor from necrotic scar tissue and the delay between cell death and macroscopic tumor shrinkage. Conversely, ^{18}F-fluorodeoxyglucose (FDG) Positron Emission Tomography (PET) detects alterations in tissue metabolism that generally precede anatomic changes (Beukinga *et al*, 2018); however, it must be taken into account that false positive results can occur during nCRT as a result of FDG uptake from inflamed tissue. False negative results can also occur because of substantial tumor shrinkage.

On that note, the recently available fully hybrid PET/Magnetic Resonance (MR) imaging allows combination of metabolic, functional and anatomic information. There are some obvious advantages on using PET/MRI over PET/CT that cannot be unnoticed, such as a substantial radiation exposure reduction explained by the fact that MR involves no high energy radiation and that the tracer dose is generally lower than that used in case of PET/CT; moreover, MR, with its excellent soft tissue resolution (De

76

Cobelli *et al*, 2020), can overcome some of the abovementioned limitations of the other staging modalities.

As main results of our investigation inquiring the role of a fully hybrid PET/MR multimodal protocol in the characterization of locally advanced esophageal neoplasms undergoing nCRT, we found that patients with a pathological complete response demonstrated *i*) significantly lower median ERI (*Early Regression Index*) values and *ii*) significantly different percentage modification in tumour ADC (in terms of median increase) from baseline to intermediate scans when compared to patients with TRG ≥ 2.

- To the best of our knowledge, this is the first study reporting association between an *early regression index* (ERI), originally developed and already externally validated by our own group in the setting of rectal cancer, and the pathological response after neoadjuvant treatment for locally advanced oesophageal neoplasms. In short, each patient is given a probability of having a complete pathological response after completion of nCRT based upon the amount of tumour shrinkage at the first imaging assessment, which accounts for the assumed fraction of dead neoplastic cells after the delivery of a certain treatment dose; in this respect, ERI could be considered a sort of a measurement surrogate for the sensitivity of neoplastic clones to the treatment. Specifically, our data demonstrate that a significant neoplastic volume drop within the third chemotherapy weekly cycle is associated with high probability of pathological complete response (TRG = 1).

- Previous studies specifically inquiring the role of DW imaging for the characterization of oesophageal cancer patients undergoing neoadjuvant treatment (Wang *et al*, 2016; Li *et al*, 2018; Xie *et al*, 2015; Giganti *et al*, 2016; De Cobelli *et al*, 2013; Borggreve *et al*, 2020a; Borggreve *et al*, 2020b) have already identified Δ ADC as a robust marker of tumour response. The vast majority of these studies (Wang *et al*, 2016; Li *et al*, 2018; Xie *et al*, 2015; Giganti *et al*, 2016; De Cobelli *et al*, 2013) reported significant modifications in tumour ADC after treatment completion when compared to baseline values; other authors (Borggreve *et al*, 2020a; Borggreve *et al*, 2020b) focused instead on possible differences occurring during nCRT. Of note, our results corroborate

the findings of these latter authors, highlighting the importance of functional modifications possibly occurring earlier than morphological ones; on that note, it is worth noting that in our patients' cohort ERI values and Δ ADC$_{mid}$ were always concordant in predicting pathological response, except for two case in which Δ ADC$_{mid}$ was later found (after surgery) to be the one correct.

Taken together, these findings highlight the importance of the intermediate evaluation (after the beginning of the third chemotherapy weekly cycle) in predicting pathological response; in this respect, our findings are consistent with those recently published by Borggreve *and colleagues*, which identify MR imaging obtained during the second week of neoadjuvant treatment for locally advanced oesophageal carcinoma as the most informative one for predicting pathological complete response (Borggreve *et al*, 2020a). This is of crucial significance, since only a very early assessment of treatment response could actually change, in a positive way, the therapeutic path of a given patient, thus providing a real benefit in terms of overall survival and, possibly, quality of life. On that note, our results could lay firm basis for the development of a novel algorithm founded on two extreme situations: on the one hand, those patients thought to be complete responders could benefit from "organ sparing"/surgery-free approaches (e.g. enrolment into endoscopic follow-up programs with *bite to bite* biopsies); on the other, poorly responder patients could undergo early surgery.

One might argue that our findings account for local disease response only, and cannot be confidently assumed as markers of overall response. Here comes the importance of a fully hybrid imaging comprising PET, which has the ability of accurately identifying eventual distant metastases. That said, in our case series no quantitative PET parameter was shown to significantly correlate to TRG; this finding is consistent with previous literature, which suggests that [18]FDG-PET does not achieve sufficient reliability to drive clinical decision.

Significant strengths of the current study include the use of the same nCRT regimen for all patients (CROSS regimen), the presence of a pathological reference standard in all patients and the consistent delineation by semi automated contouring.

Limitations also exist, the main one being the small sample size, which might have lead to false negative results with particular regard to PET quantitative parameters. Moreover, patients were enrolled during a long study period of 32 months, since

COVID-19 pandemics established a temporary downsizing of the normal oncological/surgical activity at our Institution; the same issue also undermined a homogeneous collection of data regarding longitudinal follow-up, which will be the central matter for a second study. Finally, although PET/MR is generally well tolerated by patients, shorter acquisition times could further improve their experience.

In conclusion, our preliminary results suggest the existence of a novel, easily computable MR-based biomarker (ERI), able to early predict pathological tumour response during nCRT; the percentage modification in tumour ADC values from baseline to intermediate scans (Δ ADC$_{mid}$) was also found to be significantly different between responder and partial/non-responder patients. Embedded in a multidisciplinary algorithm, these findings have the potential to provide an actual aid for patients' personalized therapeutic path.
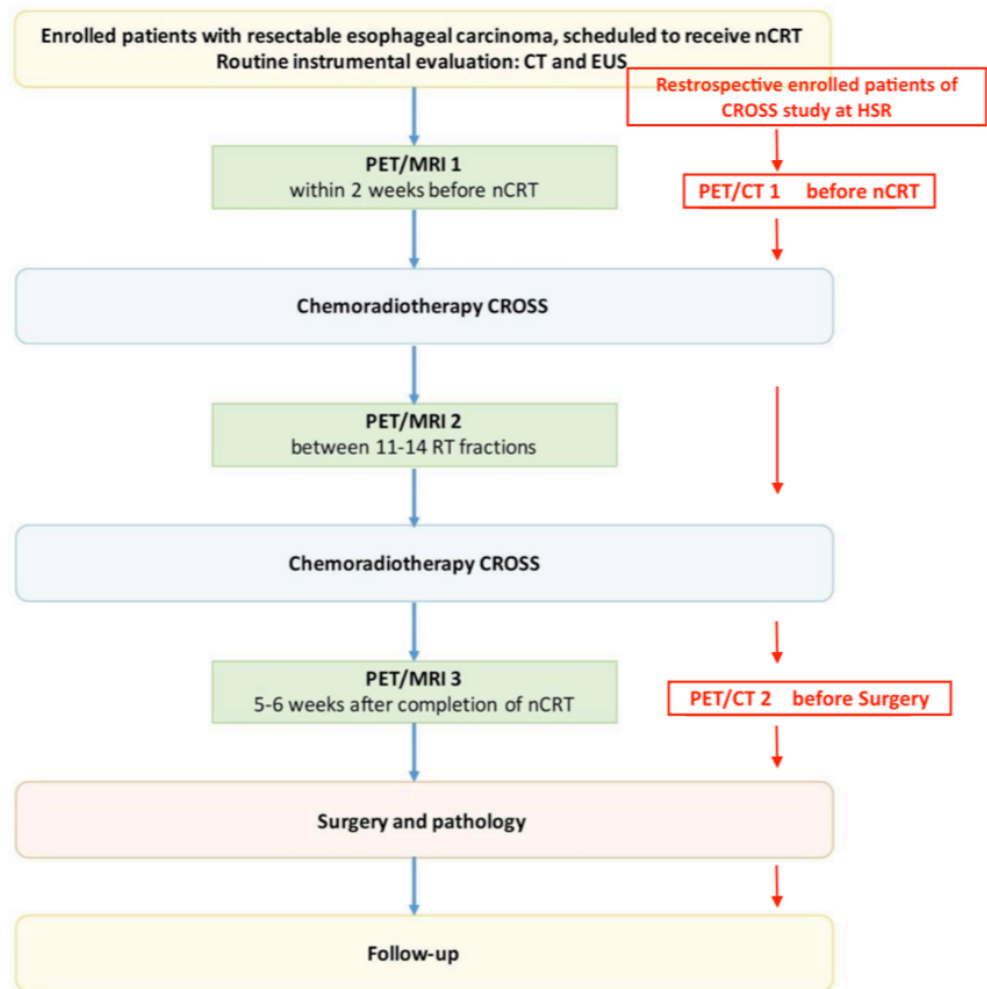
Figure #19: *Study flowchart.*

CONCLUSION

As major findings of this thesis, we identified novel imaging biomarkers for *i*) prognosis stratification and *ii*) treatment response assessment, not intended for a stand alone deployment but to be embedded into more comprehensive multidisciplinary algorithms. Now, regardless of any methodological difference possibly distinguishing each single study that composes the entire thesis, a joint substrate actually exists, being a firm intention of providing an effective biological justification for any finding resulting from our researches. Such intent has been translated into various ways, ranging from straight correlation with pathological endpoints (studies #3 and #4) and fascinating speculations exploring cellular perspectives (studies #1 and #3) to more subtle characterization of tumour microenvironment intended as both an extraordinary reflection of systemic/local metabolic dysregulations (study #2) as well as a dynamic motif to be considered as a longitudinal net change rather than a stationary frame (studies #4 and #5). In this respect, it is worth noting that the same finding could have dichotomous biological meaning depending on its specific clinical setting. As an instance, when referring to pancreatic adenocarcinoma, the same radiomic feature finally retained in the predictive models of study #1 (upfront surgery cohort) and #4 (neoadjuvant treatment cohort), *Surface Area to Volume ratio*, was found to have, in the two groups, completely different significance in terms of biological aggressiveness. A low *ratio* implies indeed a significant chance of early distant disease recurrence after primary resection, whereas its positive longitudinal net change in course of neoadjuvant chemotherapy strongly predicts ≥ 4 regional lymphnodes with metastases and, ultimately, a worse prognosis. How to solve such incongruence? Our findings support the assumption that *Surface Area to Volume ratio* conveys a balance between volume and complexity of the index lesion, with the scale tipping towards tumour volume as major determinant of prognosis (high *ratio*) in the upfront surgery cohort, and towards tumour complexity (low *ratio*) in the neoadjuvant treatment cohort. It follows that a reduction of tumour size in course of neoadjuvant chemotherapy should be considered cautiously, since it could hide a modification of tumour contours towards spiculatedness, which is a message with huge clinical implications.

Obviously, our findings do not claim to fully address all issues implied in the original research questions we presented, much less to unveil the underlying whole

biological complexity. That said, what we hope to have managed to disclose in the previous chapters is the possibility for a strict, *biologically orientated* research methodology based upon medical images. It should be a stepwise approach, *i*) starting from clinically relevant research questions, then moving down *ii*) looking for the most appropriate imaging tool to address that issue, and finally *iii*) providing a suitable biological explanation of any finding. On the opposite, a substantial part of the most recent radiological literature is built around the intriguing promise of radiomics (and, by extension, of any artificial intelligence based approach) to easily provide a large amount of data, so much that it justifies by itself its broad application escaping clinical/biological explanation. "Unweighted" deployment of radiomics accounts indeed for one of the major grounds currently hampering its actual clinical translation, together with the eventual lack of a methodologically robust radiomic workflow, as previously pointed out (Introduction). On that note, another issue that has to be considered is the complex implementation of radiomics into *real world* clinical practice; with regard to this last point, our group is currently developing *clinician-friendly* interfaces to promptly extract radiomic features ready to be embedded into *easy to use* multidisciplinary algorithms, but there is still a long way to go. Taken together, these observations corroborate our methodological approach not solely dependent on radiomics, relying instead also on other more "conventional" imaging analysis tools (a sort of "as simple as reasonably achievable" precept to adapt, time after time, to each research question), ranging, for example, from body composition assessment, diffusion weighted imaging sampling and definition of qualitative features of the index lesion (such as, tumour necrosis).

In conclusion, this thesis (and, broadly speaking, this PhD journey) allowed me to construct far more than just five studies trying addressing, at least partially, two research questions, but it laid methodological basis for other projects dealing with different subjects.

References

Abunahel BM, Pontre B, Kumar H, Petrov MS. Pancreas image mining: a systematic review of radiomics. *Eur Radiol.* 2021 May;31(5):3447-3467.

Aleotti F, Crippa S, Belfiori G, Tamburrino D, Partelli S, Longo E, Palumbo D, Pecorelli N, Lena MS, Capurso G *et al.* Pancreatic resections for benign intraductal papillary mucinous neoplasms: Collateral damages from friendly fire. *Surgery.* 2022 Oct;172(4):1202-1209.

Al-Hawary MM, Francis IR, Chari ST, Fishman EK, Hough DM, Lu DS, Macari M, Megibow AJ, Miller FH, Mortele KJ *et al.* Pancreatic ductal adenocarcinoma radiology reporting template: consensus statement of the society of abdominal radiology and the american pancreatic association. *Gastroenterology.* 2014 Jan;146(1):291-304.e1.

Arnold M, Rutherford MJ, Bardot A, Ferlay J, Andersson TM, Myklebust TÅ, Tervonen H, Thursfield V, Ransom D, Shack L, *et al.* Progress in cancer survival, mortality, and incidence in seven high-income countries 1995-2014 (ICBP SURVMARK-2): a population-based study. *Lancet Oncol.* 2019 Nov;20(11):1493-1505.

Ballehaninna UK, Chamberlain RS. The clinical utility of serum CA 19-9 in the diagnosis, prognosis and management of pancreatic adenocarcinoma: An evidence based appraisal. *J Gastrointest Oncol.* 2012 Jun;3(2):105-19.

Barugola G, Partelli S, Marcucci S, Sartori N, Capelli P, Bassi C, Pederzoli P, Falconi M. Resectable pancreatic cancer: who really benefits from resection? *Ann Surg Oncol.* 2009 Dec;16(12):3316-22.

Battistella A, Partelli S, Andreasi V, Marinoni I, Palumbo D, Tacelli M, Lena MS, Muffatti F, Mushtaq J, Capurso G *et al.* Preoperative assessment of microvessel density in nonfunctioning pancreatic neuroendocrine tumors (NF-PanNETs). *Surgery.* 2022 Oct;172(4):1236-1244.

Benedetti G, Mori M, Panzeri MM, Barbera M, Palumbo D, Sini C, Muffatti F, Andreasi V, Steidler S, Doglioni C *et al.* CT-derived radiomic features to discriminate histologic characteristics of pancreatic neuroendocrine tumors. *Radiol Med.* 2021 Jun;126(6):745-760.

Bezzi C, Mapelli P, Presotto L, Neri I, Scifo P, Savi A, Bettinardi V, Partelli S, Gianolli L, Falconi M *et al.* Radiomics in pancreatic neuroendocrine tumors: methodological issues and clinical significance. *Eur J Nucl Med Mol Imaging.* 2021 Nov;48(12):4002-4015.

Beukinga RJ, Hulshoff JB, Mul VEM, Noordzij W, Kats-Ugurlu G, Slart RHJA, Plukker JTM. Prediction of Response to Neoadjuvant Chemotherapy and Radiation Therapy with Baseline and Restaging ^{18}F-FDG PET Imaging Biomarkers in Patients with Esophageal Cancer. *Radiology.* 2018 Jun;287(3):983-992.

Bian Y, Zhao Z, Jiang H, Fang X, Li J, Cao K, Ma C, Guo S, Wang L, Jin G *et al.* Noncontrast Radiomics Approach for Predicting Grades of Nonfunctional Pancreatic Neuroendocrine Tumors. *J Magn Reson Imaging.* 2020 Oct;52(4):1124-1136.

Bian Y, Jiang H, Ma C, Wang L, Zheng J, Jin G, Lu J. CT-Based Radiomics Score for Distinguishing Between Grade 1 and Grade 2 Nonfunctioning Pancreatic Neuroendocrine Tumors. *AJR Am J Roentgenol.* 2020 Oct;215(4):852-863.

Borhani AA, Dewan R, Furlan A, Seiser N, Zureikat AH, Singhi AD, Boone B, Bahary N, Hogg ME, Lotze M *et al.* Assessment of Response to Neoadjuvant Therapy Using CT Texture Analysis in Patients With Resectable and Borderline Resectable Pancreatic Ductal Adenocarcinoma. *AJR Am J Roentgenol.* 2020 Feb;214(2):362-369.

Borggreve AS, Heethuis SE, Boekhoff MR, Goense L, van Rossum PSN, Brosens LAA, van Lier ALHMW, van Hillegersberg R, Lagendijk JJW, Mook S *et al.* Optimal timing for prediction of pathologic complete response to neoadjuvant chemoradiotherapy with diffusion-weighted MRI in patients with esophageal cancer. *Eur Radiol.* 2020 Apr;30(4):1896-1907.

Borggreve AS, Goense L, van Rossum PSN, Heethuis SE, van Hillegersberg R, Lagendijk JJW, Lam MGEH, van Lier ALHMW, Mook S, Ruurda JP *et al.* Preoperative Prediction of Pathologic Response to Neoadjuvant Chemoradiotherapy in Patients With Esophageal Cancer Using ^{18}F-FDG PET/CT and DW-MRI: A Prospective Multicenter Study. *Int J Radiat Oncol Biol Phys.* 2020 Apr 1;106(5):998-1009.

Braga M, Pecorelli N, Ariotti R, Capretti G, Greco M, Balzano G, Castoldi R, Beretta L. Enhanced recovery after surgery pathway in patients undergoing pancreaticoduodenectomy. *World J Surg*. 2014 Nov;38(11):2960-6.

Bretthauer M, Løberg M, Wieszczy P, Kalager M, Emilsson L, Garborg K, Rupinski M, Dekker E, Spaander M, Bugajski M, *et al.* Effect of Colonoscopy Screening on Risks of Colorectal Cancer and Related Death. *N Engl J Med.* 2022 Oct 27;387(17):1547-1556.

Bribiesca, E. Measure of compactness for 3D shapes. *Comput. Math. with Appl.* 40, 2000 1275-1284.

Broggi S, Passoni P, Gumina C, Palmisano A, Bresolin A, Burgio V, Di Chiara A, Elmore U, Mori M, Slim N *et al.* Predicting pathological response after radio-chemotherapy for rectal cancer: Impact of late oxaliplatin administration. *Radiother Oncol.* 2020 Aug;149:174-180.

Cacciato Insilla A, Vivaldi C, Giordano M, Vasile E, Cappelli C, Kauffmann E, Napoli N, Falcone A, Boggi U, Campani D. Tumor Regression Grading Assessment in Locally Advanced Pancreatic Cancer After Neoadjuvant FOLFIRINOX: Interobserver Agreement and Prognostic Implications. *Front Oncol.* 2020 Feb 7;10:64.

Capretti G, Balzano G, Gianotti L, Stella M, Ferrari G, Baccari P, Zuliani W, Braga M, Zerbi A. Management and Outcomes of Pancreatic Resections Performed in High-Volume Referral and Low-Volume Community Hospitals Lead by Surgeons Who Shared the Same Mentor: The Importance of Training. *Dig Surg.* 2018;35(1):42-48.

Capurso G, Crippa S, Vanella G, Traini M, Zerboni G, Zaccari P, Belfiori G, Gentiluomo M, Pessarelli T, Petrone MC *et al.* Factors Associated With the Risk of Progression of Low-Risk Branch-Duct Intraductal Papillary Mucinous Neoplasms. *JAMA Netw Open.* 2020 Nov 2;3(11):e2022933.

Cassinotto C, Mouries A, Lafourcade JP, Terrebonne E, Belleannée G, Blanc JF, Lapuyade B, Vendrely V, Laurent C, Chiche L *et al.* Locally advanced pancreatic adenocarcinoma: reassessment of response with CT after neoadjuvant chemotherapy and radiation therapy. *Radiology.* 2014 Oct;273(1):108-16.

Choi TW, Kim JH, Yu MH, Park SJ, Han JK. Pancreatic neuroendocrine tumor: prediction of the tumor grade using CT findings and computerized texture analysis. *Acta Radiol*. 2018 Apr;59(4):383-392.

Conroy T, Hammel P, Hebbar M, Ben Abdelghani M, Wei AC, Raoul JL, Choné L, Francois E, Artru P, Biagi JJ *et al*. FOLFIRINOX or Gemcitabine as Adjuvant Therapy for Pancreatic Cancer. *N Engl J Med*. 2018 Dec 20;379(25):2395-2406.

Crippa S, Pergolini I, Rubini C, Castelli P, Partelli S, Zardini C, Marchesini G, Zamboni G, Falconi M. Risk of misdiagnosis and overtreatment in patients with main pancreatic duct dilatation and suspected combined/main-duct intraductal papillary mucinous neoplasms. *Surgery*. 2016 Apr;159(4):1041-9.

Crippa S, Belfiori G, Bissolati M, Partelli S, Pagnanelli M, Tamburrino D, Gasparini G, Rubini C, Zamboni G, Falconi M. Recurrence after surgical resection of pancreatic cancer: the importance of postoperative complications beyond tumor biology. *HPB (Oxford)*. 2021 Nov;23(11):1666-1673.

Crippa S, Aleotti F, Longo E, Belfiori G, Partelli S, Tamburrino D, Di Salvo F, De Cobelli F, Romano L, Arcidiacono PG *et al*. Main Duct Thresholds for Malignancy Are Different in Intraductal Papillary Mucinous Neoplasms of the Pancreatic Head and Body-Tail. *Clin Gastroenterol Hepatol*. 2022 Feb;20(2):390-399.e7.

Cusumano D, Boldrini L, Yadav P, Yu G, Musurunu B, Chiloiro G, Piras A, Lenkowicz J, Placidi L, Broggi S *et al*. External Validation of Early Regression Index (ERI$_{TCP}$) as Predictor of Pathologic Complete Response in Rectal Cancer Using Magnetic Resonance-Guided Radiation Therapy. *Int J Radiat Oncol Biol Phys*. 2020 Dec 1;108(5):1347-1356.

De Cobelli F, Giganti F, Orsenigo E, Cellina M, Esposito A, Agostini G, Albarello L, Mazza E, Ambrosi A, Socci C *et al*. Apparent diffusion coefficient modifications in assessing gastro-oesophageal cancer response to neoadjuvant treatment: comparison with tumour regression grade at histology. *Eur Radiol*. 2013 Aug;23(8):2165-74.

De Cobelli F, Palumbo D, Albarello L, Rosati R, Giganti F. Esophagus and Stomach: Is There a Role for MR Imaging? *Magn Reson Imaging Clin N Am*. 2020 Feb;28(1):1-15.

D'Journo XB, Boulate D, Fourdrain A, Loundou A, van Berge Henegouwen MI, Gisbertz SS, O'Neill JR, Hoelscher A, Piessen G, van Lanschot J *et al.* Risk Prediction Model of 90-Day Mortality After Esophagectomy for Cancer. *JAMA Surg.* 2021 Sep 1;156(9):836-845.

Eilaghi A, Baig S, Zhang Y, Zhang J, Karanicolas P, Gallinger S, Khalvati F, Haider MA. CT texture features are associated with overall survival in pancreatic ductal adenocarcinoma - a quantitative analysis. *BMC Med Imaging.* 2017 Jun 19;17(1):38.

Eklund M, Jäderling F, Discacciati A, Bergman M, Annerstedt M, Aly M, Glaessgen A, Carlsson S, Grönberg H, Nordström T; STHLM3 consortium. MRI-Targeted or Standard Biopsy in Prostate Cancer Screening. *N Engl J Med.* 2021 Sep 2;385(10):908-920.

Elsherif SB, Javadi S, Le O, Lamba N, Katz MHG, Tamm EP, Bhosale PR. Baseline CT-based Radiomic Features Aid Prediction of Nodal Positivity after Neoadjuvant Therapy in Pancreatic Cancer. *Radiol Imaging Cancer.* 2022 Mar;4(2):e210068.

Fiorino C, Gumina C, Passoni P, Palmisano A, Broggi S, Cattaneo GM, Di Chiara A, Esposito A, Mori M, Raso R *et al.* A TCP-based early regression index predicts the pathological response in neo-adjuvant radio-chemotherapy of rectal cancer. *Radiother Oncol.* 2018 Sep;128(3):564-568.

Fiorino C, Passoni P, Palmisano A, Gumina C, Cattaneo GM, Broggi S, Di Chiara A, Esposito A, Mori M, Ronzoni M *et al.* Accurate outcome prediction after neo-adjuvant radio-chemotherapy for rectal cancer based on a TCP-based early regression index. *Clin Transl Radiat Oncol.* 2019 Jul 3;19:12-16.

Gardner TB, Glass LM, Smith KD, Ripple GH, Barth RJ, Klibansky DA, Colacchio TA, Tsapakos MJ, Suriawinata AA, Tsongalis GJ *et al.* Pancreatic cyst prevalence and the risk of mucin-producing adenocarcinoma in US adults. *Am J Gastroenterol.* 2013 Oct;108(10):1546-50.

Gillies RJ, Kinahan PE, Hricak H. Radiomics: Images Are More than Pictures, They Are Data. *Radiology.* 2016 Feb;278(2):563-77.

Giganti F, Salerno A, Ambrosi A, Chiari D, Orsenigo E, Esposito A, Albarello L, Mazza E, Staudacher C, Del Maschio A *et al.* Prognostic utility of diffusion-

weighted MRI in oesophageal cancer: is apparent diffusion coefficient a potential marker of tumour aggressiveness? Radiol Med. 2016 Mar;121(3):173-80.

Groot VP, Rezaee N, Wu W, Cameron JL, Fishman EK, Hruban RH, Weiss MJ, Zheng L, Wolfgang CL, He J. Patterns, Timing, and Predictors of Recurrence Following Pancreatectomy for Pancreatic Ductal Adenocarcinoma. Ann Surg. 2018 May;267(5):936-945.

Gu D, Hu Y, Ding H, Wei J, Chen K, Liu H, Zeng M, Tian J. CT radiomics may predict the grade of pancreatic neuroendocrine tumors: a multicenter study. *Eur Radiol. 2019* Dec;29(12):6880-6890.

Gu H, Deng W, Zheng Z, Wu K, Sun F. CCL2 produced by pancreatic ductal adenocarcinoma is essential for the accumulation and activation of monocytic myeloid-derived suppressor cells. *Immun Inflamm Dis.* 2021 Dec;9(4):1686-1695.

Guarneri G, de Mestier L, Landoni L, Partelli S, Gaujoux S, Andreasi V, Nessi C, Dokmak S, Fontana M, Dousset B et al. A. Prognostic Role of Examined and Positive Lymph Nodes after Distal Pancreatectomy for Non-Functioning Neuroendocrine Neoplasms. *Neuroendocrinology.* 2021;111(8):728-738.

Guarneri G, Palumbo D, Pecorelli N, Prato F, Gritti C, Cerchione R, Tamburrino D, Partelli S, Crippa S, Reni M *et al.* The Impact of CT-Assessed Liver Steatosis on Postoperative Complications After Pancreaticoduodenectomy for Cancer. *Ann Surg Oncol.* 2022 Oct;29(11):7063-7073.

Haeno H, Gonen M, Davis MB, Herman JM, Iacobuzio-Donahue CA, Michor F. Computational modeling of pancreatic cancer reveals kinetics of metastasis suggesting optimum treatment strategies. *Cell*. 2012 Jan 20;148(1-2):362-75.

van Hagen P, Hulshof MC, van Lanschot JJ, Steyerberg EW, van Berge Henegouwen MI, Wijnhoven BP, Richel DJ, Nieuwenhuijzen GA, Hospers GA, Bonenkamp JJ *et al.* Preoperative chemoradiotherapy for esophageal or junctional cancer. *N Engl J Med.* 2012 May 31;366(22):2074-84.

Harris LK, Theriot JA. Surface Area to Volume Ratio: A Natural Variable for Bacterial Morphogenesis. *Trends Microbiol*. 2018 Oct;26(10):815-832.

Ho WJ, Jaffee EM, Zheng L. The tumour microenvironment in pancreatic cancer - clinical challenges and opportunities. *Nat Rev Clin Oncol.* 2020 Sep;17(9):527-540.

Ironi G, Mapelli P, Bergamini A, Fallanca F, Candotti G, Gnasso C, Taccagni GL, Sant'Angelo M, Scifo P, Bezzi C *et al.* Hybrid PET/MRI in Staging Endometrial Cancer: Diagnostic and Predictive Value in a Prospective Cohort. *Clin Nucl Med.* 2022 Mar 1;47(3):e221-e229.

Jang JK, Byun JH, Kang JH, Son JH, Kim JH, Lee SS, Kim HJ, Yoo C, Kim KP, Hong SM *et al.* CT-determined resectability of borderline resectable and unresectable pancreatic adenocarcinoma following FOLFIRINOX therapy. *Eur Radiol.* 2021 Feb;31(2):813-823.

Jeon SK, Lee JM, Lee ES, Yu MH, Joo I, Yoon JH, Jang JY, Lee KB, Lee SH. How to approach pancreatic cancer after neoadjuvant treatment: assessment of resectability using multidetector CT and tumor markers. *Eur Radiol.* 2022 Jan;32(1):56-66.

Karavia EA, Papachristou DJ, Liopeta K, Triantaphyllidou IE, Dimitrakopoulos O, Kypreos KE. Apolipoprotein A-I modulates processes associated with diet-induced nonalcoholic fatty liver disease in mice. *Mol Med.* 2012 Sep 7;18(1):901-12.

Kim DW, Lee SS, Kim SO, Kim JH, Kim HJ, Byun JH, Yoo C, Kim KP, Song KB, Kim SC. Estimating Recurrence after Upfront Surgery in Patients with Resectable Pancreatic Ductal Adenocarcinoma by Using Pancreatic CT: Development and Validation of a Risk Score. *Radiology.* 2020 Sep;296(3):541-551.

Klimstra DS. Pathologic Classification of Neuroendocrine Neoplasms. *Hematol Oncol Clin North Am.* 2016 Feb;30(1):1-19.

Koo MM, Swann R, McPhail S, Abel GA, Elliss-Brookes L, Rubin GP, Lyratzopoulos G. Presenting symptoms of cancer and stage at diagnosis: evidence from a cross-sectional, population-based study. *Lancet Oncol.* 2020 Jan;21(1):73-79.

Kudo M, Kobayashi T, Gotohda N, Konishi M, Takahashi S, Kobayashi S, Sugimoto M, Okubo S, Martin J, Cabral H *et al.* M. Clinical Utility of Histological and

Radiological Evaluations of Tumor Necrosis for Predicting Prognosis in Pancreatic Cancer. Pancreas. 2020 May/Jun;49(5):634-641.

Kulali F, Semiz-Oysu A, Demir M, Segmen-Yilmaz M, Bukte Y. Role of diffusion-weighted MR imaging in predicting the grade of nonfunctional pancreatic neuroendocrine tumors. *Diagn Interv Imaging.* 2018 May;99(5):301-309.

Lambin P, Leijenaar RTH, Deist TM, Peerlings J, de Jong EEC, van Timmeren J, Sanduleanu S, Larue RTHM, Even AJG, Jochems A et al. Radiomics: the bridge between medical imaging and personalized medicine. *Nat Rev Clin Oncol.* 2017 Dec;14(12):749-762.

Lee YS, Lee JC, Yang SY, Kim J, Hwang JH. Neoadjuvant therapy versus upfront surgery in resectable pancreatic cancer according to intention-to-treat and per-protocol analysis: A systematic review and meta-analysis. *Sci Rep.* 2019 Oct 30;9(1):15662.

Li QW, Qiu B, Wang B, Wang DL, Yin SH, Yang H, Liu JL, Fu JH, Liu MZ, Xie CM *et al.* Prediction of pathologic responders to neoadjuvant chemoradiotherapy by diffusion-weighted magnetic resonance imaging in locally advanced esophageal squamous cell carcinoma: a prospective study. *Dis Esophagus.* 2018 Feb 1;31(2).

Liang W, Yang P, Huang R, Xu L, Wang J, Liu W, Zhang L, Wan D, Huang Q, Lu Y *et al.* A Combined Nomogram Model to Preoperatively Predict Histologic Grade in Pancreatic Neuroendocrine Tumors. *Clin Cancer Res.* 2019 Jan 15;25(2):584-594.

Limkin EJ, Reuzé S, Carré A, Sun R, Schernberg A, Alexis A, Deutsch E, Ferté C, Robert C. The complexity of tumor shape, spiculatedness, correlates with tumor radiomic shape features. *Sci Rep.* 2019 Mar 13;9(1):4329.

Lin D, Wang Z, Li H, Zhang H, Deng L, Ren H, Sun S, Zheng F, Zhou J, Wang M. Automated Measurement of Pancreatic Fat Deposition on Dixon MRI Using nnU-Net. *J Magn Reson Imaging.* 2022 May 30.

Loi S, Mori M, Benedetti G, Partelli S, Broggi S, Cattaneo GM, Palumbo D, Muffatti F, Falconi M, De Cobelli F *et al.* Robustness of CT radiomic features against image discretization and interpolation in characterizing pancreatic neuroendocrine neoplasms. *Phys Med.* 2020 Aug;76:125-133.

Low DE, Kuppusamy MK, Alderson D, Cecconello I, Chang AC, Darling G, Davies A, D'Journo XB, Gisbertz SS, Griffin SM *et al.* Benchmarking Complications Associated with Esophagectomy. *Ann Surg.* 2019 Feb;269(2):291-298.

Macchini M, Centonze F, Peretti U, Orsi G, Militello AM, Valente MM, Cascinu S, Reni M. Treatment opportunities and future perspectives for pancreatic cancer patients with germline BRCA1-2 pathogenic variants. *Cancer Treat Rev.* 2021 Nov;100:102262.

Makohon-Moore AP, Matsukuma K, Zhang M, Reiter JG, Gerold JM, Jiao Y, Sikkema L, Attiyeh MA, Yachida S, Sandone C *et al.* Precancerous neoplastic cells can move through the pancreatic ductal system. *Nature.* 2018 Sep;561(7722):201-205.

Mandard AM, Dalibard F, Mandard JC, Marnay J, Henry-Amar M, Petiot JF, Roussel A, Jacob JH, Segol P, Samama G, *et al.* Pathologic assessment of tumor regression after preoperative chemoradiotherapy of esophageal carcinoma. Clinicopathologic correlations. *Cancer.* 1994 Jun 1;73(11):2680-6.

Mapelli P, Ghezzo S, Samanes Gajate AM, Preza E, Palmisano A, Cucchiara V, Brembilla G, Bezzi C, Rigamonti R, Magnani P *et al.* ^{68}Ga-PSMA and ^{68}Ga-DOTA-RM2 PET/MRI in Recurrent Prostate Cancer: Diagnostic Performance and Association with Clinical and Histopathological Data. *Cancers (Basel).* 2022 Jan 11;14(2):334.

Mapelli P, Bezzi C, Palumbo D, Canevari C, Ghezzo S, Samanes Gajate AM, Catalfamo B, Messina A, Presotto L, Guarnaccia A et al. ^{68}Ga-DOTATOC PET/MR imaging and radiomic parameters in predicting histopathological prognostic factors in patients with pancreatic neuroendocrine well-differentiated tumours. *Eur J Nucl Med Mol Imaging.* 2022 Jun;49(7):2352-2363.

Marchegiani G, Todaro V, Boninsegna E, Negrelli R, Sureka B, Bonamini D, Salvia R, Manfredi R, Pozzi Mucelli R, Bassi C. Surgery after FOLFIRINOX treatment for locally advanced and borderline resectable pancreatic cancer: increase in tumour attenuation on CT correlates with R0 resection. *Eur Radiol.* 2018 Oct;28(10):4265-4273.

Mathur A, Zyromski NJ, Pitt HA, Al-Azzawi H, Walker JJ, Saxena R, Lillemoe KD. Pancreatic steatosis promotes dissemination and lethality of pancreatic cancer. *J Am Coll Surg.* 2009 May;208(5):989-94; discussion 994-6.

Matsumoto I, Murakami Y, Shinzeki M, Asari S, Goto T, Tani M, Motoi F, Uemura K, Sho M, Satoi S *et al.* Proposed preoperative risk factors for early recurrence in patients with resectable pancreatic ductal adenocarcinoma after surgical resection: A multi-center retrospective study. *Pancreatology.* 2015 Nov-Dec;15(6):674-80.

Mizrahi JD, Surana R, Valle JW, Shroff RT. Pancreatic cancer. *Lancet.* 2020 Jun 27;395(10242):2008-2020.

Monti P, Leone BE, Marchesi F, Balzano G, Zerbi A, Scaltrini F, Pasquali C, Calori G, Pessi F, Sperti C *et al.* The CC chemokine MCP-1/CCL2 in pancreatic cancer progression: regulation of expression and potential mechanisms of antimalignant activity. *Cancer Res.* 2003 Nov 1;63(21):7451-61.

Moons KG, Altman DG, Reitsma JB, Ioannidis JP, Macaskill P, Steyerberg EW, Vickers AJ, Ransohoff DF, Collins GS. Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): explanation and elaboration. *Ann Intern Med.* 2015 Jan 6;162(1):W1-73.

Mori M, Benedetti G, Partelli S, Sini C, Andreasi V, Broggi S, Barbera M, Cattaneo GM, Muffatti F, Panzeri M *et al.* Ct radiomic features of pancreatic neuroendocrine neoplasms (panNEN) are robust against delineation uncertainty. *Phys Med.* 2019 Jan;57:41-46.

Mori M, Passoni P, Incerti E, Bettinardi V, Broggi S, Reni M, Whybra P, Spezi E, Vanoli EG, Gianolli L *et al.* Training and validation of a robust PET radiomic-based index to predict distant-relapse-free-survival after radio-chemotherapy for locally advanced pancreatic cancer. *Radiother Oncol.* 2020 Dec;153:258-264.

Mori M, Palumbo D, De Cobelli F, Fiorino C. Does radiomics play a role in the diagnosis, staging and re-staging of gastroesophageal junction adenocarcinoma? *Updates Surg.* 2022 Sep 17.

Mukherjee S, Patra A, Khasawneh H, Korfiatis P, Rajamohan N, Suman G, Majumder S, Panda A, Johnson MP, Larson NB *et al.* Radiomics-based Machine-learning

Models Can Detect Pancreatic Cancer on Prediagnostic Computed Tomography Scans at a Substantial Lead Time Before Clinical Diagnosis. *Gastroenterology*. 2022 Nov;163(5):1435-1446.e3.

National Lung Screening Trial Research Team, Aberle DR, Adams AM, Berg CD, Black WC, Clapp JD, Fagerstrom RM, Gareen IF, Gatsonis C, Marcus PM *et al*. Reduced lung-cancer mortality with low-dose computed tomographic screening. *N Engl J Med.* 2011 Aug 4;365(5):395-409.

Neoptolemos JP, Palmer DH, Ghaneh P, Psarelli EE, Valle JW, Halloran CM, Faluyi O, O'Reilly DA, Cunningham D, Wadsley J *et al.* Comparison of adjuvant gemcitabine and capecitabine with gemcitabine monotherapy in patients with resected pancreatic cancer (ESPAC-4): a multicentre, open-label, randomised, phase 3 trial. *Lancet.* 2017 Mar 11;389(10073):1011-1024.

Noda Y, Pisuchpen N, Mercaldo ND, Sekigami Y, Michelakos T, Parakh A, Wo JY, Qadan M, Ferrone C, Lillemoe KD *et al*. Arterial involvement and resectability scoring system to predict R0 resection in patients with pancreatic ductal adenocarcinoma treated with neoadjuvant chemoradiation therapy. *Eur Radiol*. 2022 Apr;32(4):2470-2480.

Overbeek KA, Goggins MG, Dbouk M, Levink IJM, Koopmann BDM, Chuidian M, Konings ICAW, Paiella S, Earl J, Fockens P, *et al.* Timeline of Development of Pancreatic Cancer and Implications for Successful Early Detection in High-Risk Individuals. *Gastroenterology*. 2022 Mar;162(3):772-785.e4.

Palumbo D, Mapelli P, Nicoletti V, Steidler S, Picchio M, De Cobelli F. Imaging in evaluation of response to neo-adjuvant treatment. *Ann Esophagus.* 2020 Dec

Palumbo D, Tamburrino D, Partelli S, Gusmini S, Guazzarotti G, Cao R, Crippa S, Falconi M, De Cobelli F. Before sentinel bleeding: early prediction of postpancreatectomy hemorrhage (PPH) with a CT-based scoring system. *Eur Radiol.* 2021 Sep;31(9):6879-6888.

Palumbo D, Mori M, Prato F, Crippa S, Belfiori G, Reni M, Mushtaq J, Aleotti F, Guazzarotti G, Cao R et al. Prediction of Early Distant Recurrence in Upfront Resectable Pancreatic Adenocarcinoma: A Multidisciplinary, Machine Learning-Based Approach. *Cancers (Basel)*. 2021 Sep 30;13(19):4938.

Park S, Jang JK, Byun JH, Kim JH, Lee SS, Kim HJ, Hong SB, Park SH. CT in the prediction of margin-negative resection in pancreatic cancer following neoadjuvant treatment: a systematic review and meta-analysis. *Eur Radiol.* 2021 May;31(5):3383-3393.

Partelli S, Cirocchi R, Crippa S, Cardinali L, Fendrich V, Bartsch DK, Falconi M. Systematic review of active surveillance versus surgical management of asymptomatic small non-functioning pancreatic neuroendocrine neoplasms. *Br J Surg.* 2017 Jan;104(1):34-41.

Partelli S, Bartsch DK, Capdevila J, Chen J, Knigge U, Niederle B, Nieveen van Dijkum EJM, Pape UF, Pascher A, Ramage J *et al.*ENETS Consensus Guidelines for Standard of Care in Neuroendocrine Tumours: Surgery for Small Intestinal and Pancreatic Neuroendocrine Tumours. *Neuroendocrinology.* 2017;105(3):255-265.

Partelli S, Muffatti F, Andreasi V, Giannone F, Rossi G, Palumbo D, Mapelli P, Schiavo Lena M, Arcidiacono PG *et al.* A Single-center Prospective Observational Study Investigating the Accuracy of Preoperative Diagnostic Procedures in the Assessment of Lymph Node Metastases in Nonfunctioning Pancreatic Neuroendocrine Tumors. *Ann Surg.* 2022 Nov 1;276(5):921-928.

Pasaoglu E, Dursun N, Ozyalvacli G, Hacihasanoglu E, Behzatoglu K, Calay O. Comparison of World Health Organization 2000/2004 and World Health Organization 2010 classifications for gastrointestinal and pancreatic neuroendocrine tumors. *Ann Diagn Pathol.* 2015 Apr;19(2):81-7.

Pecorelli N, Palumbo D, Guarneri G, Gritti C, Prato F, Schiavo Lena M, Vallorani A, Partelli S, Crippa S, Doglioni C *et al.* Preoperative CT image analysis to improve risk stratification for clinically relevant pancreatic fistula after distal pancreatectomy. *Br J Surg.* 2022 Oct 29:znac348.

Peduzzi P, Concato J, Kemper E, Holford TR, Feinstein AR. A simulation study of the number of events per variable in logistic regression analysis. *J Clin Epidemiol.* 1996 Dec;49(12):1373-9.

Pereira SP, Oldfield L, Ney A, Hart PA, Keane MG, Pandol SJ, Li D, Greenhalf W, Jeon CY, Koay EJ, *et al.* Early detection of pancreatic cancer. *Lancet Gastroenterol Hepatol.* 2020 Jul;5(7):698-710.

Petrelli F, Inno A, Barni S, Ghidini A, Labianca R, Falconi M, Reni M, Cascinu S; GISCAD (Gruppo Italiano per lo Studio dei Carcinomi dell'Apparato Digerente) and San Raffaele Hospital. Borderline resectable pancreatic cancer: More than an anatomical concept. *Dig Liver Dis.* 2017 Feb;49(2):223-226.

Pinto Dos Santos D, Dietzel M, Baessler B. A decade of radiomics research: are images really data or just patterns in the noise? *Eur Radiol.* 2021 Jan;31(1):1-4.

Rebours V, Cordova J, Couvelard A, Fabre M, Palazzo L, Vullierme MP, Hentic O, Sauvanet A, Aubert A, Bedossa P *et al.* Can pancreatic neuroendocrine tumour biopsy accurately determine pathological characteristics? *Dig Liver Dis.* 2015 Nov;47(11):973-7.

Reni M, Balzano G, Zanon S, Zerbi A, Rimassa L, Castoldi R, Pinelli D, Mosconi S, Doglioni C, Chiaravalli M et al. Safety and efficacy of preoperative or postoperative chemotherapy for resectable pancreatic adenocarcinoma (PACT-15): a randomised, open-label, phase 2-3 trial. *Lancet Gastroenterol Hepatol.* 2018 Jun;3(6):413-423.

Reni M, Zanon S, Balzano G, Passoni P, Pircher C, Chiaravalli M, Fugazza C, Ceraulo D, Nicoletti R, Arcidiacono PG *et al.* A randomised phase 2 trial of nab-paclitaxel plus gemcitabine with or without capecitabine and cisplatin in locally advanced or borderline resectable pancreatic adenocarcinoma. *Eur J Cancer.* 2018 Oct;102:95-102.

Reni M, Peretti U, Zanon S, Macchini M, Balzano G, Mazza E, Tamburrino D, Orsi G, Arcidiacono PG, Falconi M *et al.* Time to CA19-9 nadir: a clue for defining optimal treatment duration in patients with resectable pancreatic ductal adenocarcinoma. *Cancer Chemother Pharmacol.* 2020 Apr;85(4):641-650.

van Roessel S, Kasumova GG, Verheij J, Najarian RM, Maggino L, de Pastena M, Malleo G, Marchegiani G, Salvia R, Ng SC *et al.* International Validation of the Eighth Edition of the American Joint Committee on Cancer (AJCC) TNM Staging System in Patients With Resected Pancreatic Cancer. *JAMA Surg.* 2018 Dec 1;153(12):e183617.

Sadot E, Basturk O, Klimstra DS, Gönen M, Lokshin A, Do RK, D'Angelica MI, DeMatteo RP, Kingham TP, Jarnagin WR *et al.* Tumor-associated Neutrophils and Malignant Progression in Intraductal Papillary Mucinous Neoplasms: An

Opportunity for Identification of High-risk Disease. *Ann Surg.* 2015 Dec;262(6):1102-7.

Sawaya GF, Smith-McCune K, Kuppermann M. Cervical Cancer Screening: More Choices in 2019. *JAMA*. 2019 May 28;321(20):2018-2019.

Shapiro J, van Lanschot JJB, Hulshof MCCM, van Hagen P, van Berge Henegouwen MI, Wijnhoven BPL, van Laarhoven HWM, Nieuwenhuijzen GAP, Hospers GAP, Bonenkamp JJ *et al.* Neoadjuvant chemoradiotherapy plus surgery versus surgery alone for oesophageal or junctional cancer (CROSS): long-term results of a randomised controlled trial. Lancet Oncol. 2015 Sep;16(9):1090-1098.

Sotozono H, Kanki A, Yasokawa K, Yamamoto A, Sanai H, Moriya K, Tamada T. Value of 3-T MR imaging in intraductal papillary mucinous neoplasm with a concomitant invasive carcinoma. *Eur Radiol.* 2022 Jun 3.

Spadarella G, Stanzione A, Akinci D'Antonoli T, Andreychenko A, Fanni SC, Ugga L, Kotter E, Cuocolo R. Systematic review of the radiomics quality score applications: an EuSoMII Radiomics Auditing Group Initiative. *Eur Radiol.* 2022 Oct 25.

Strijker M, Chen JW, Mungroop TH, Jamieson NB, van Eijck CH, Steyerberg EW, Wilmink JW, Groot Koerkamp B, van Laarhoven HW, Besselink MG. Systematic review of clinical prediction models for survival after surgery for resectable pancreatic cancer. *Br J Surg.* 2019 Mar;106(4):342-354.

Tanaka M, Fernández-Del Castillo C, Kamisawa T, Jang JY, Levy P, Ohtsuka T, Salvia R, Shimizu Y, Tada M, Wolfgang CL. Revisions of international consensus Fukuoka guidelines for the management of IPMN of the pancreas. *Pancreatology*. 2017 Sep-Oct;17(5):738-753.

Tempero MA, Malafa MP, Al-Hawary M, Behrman SW, Benson AB, Cardin DB, Chiorean EG, Chung V, Czito B, Del Chiaro M *et al*. Pancreatic Adenocarcinoma, Version 2.2021, NCCN Clinical Practice Guidelines in Oncology. *J Natl Compr Canc Netw.* 2021 Apr 1;19(4):439-457.

The Lancet. Cancer care: beyond survival. *Lancet*. 2022 Apr 16;399(10334):1441.

Wagner M, Antunes C, Pietrasz D, Cassinotto C, Zappa M, Sa Cunha A, Lucidarme O, Bachet JB. CT evaluation after neoadjuvant FOLFIRINOX chemotherapy for

borderline and locally advanced pancreatic adenocarcinoma. *Eur Radiol.* 2017 Jul;27(7):3104-3116.

Wang L, Liu L, Han C, Liu S, Tian H, Li Z, Ren X, Shi G, Wang Q, Wang G. The diffusion-weighted magnetic resonance imaging (DWI) predicts the early response of esophageal squamous cell carcinoma to concurrent chemoradiotherapy. *Radiother Oncol.* 2016 Nov;121(2):246-251.

Welch HG, Prorok PC, O'Malley AJ, Kramer BS. Breast-Cancer Tumor Size, Overdiagnosis, and Mammography Screening Effectiveness. *N Engl J Med.* 2016 Oct 13;375(15):1438-1447.

Wen Y, Chen C, Kong X, Xia Z, Kong W, Si K, Han P, Vivian Liu W, Li X. Pancreatic fat infiltration, β-cell function and insulin resistance: A study of the young patients with obesity. *Diabetes Res Clin Pract.* 2022 May;187:109860.

Whitaker K. Earlier diagnosis: the importance of cancer symptoms. *Lancet Oncol.* 2020 Jan;21(1):6-8.

Xie H, Sun T, Chen M, Wang H, Zhou X, Zhang Y, Zeng H, Wang J, Fu W. Effectiveness of the apparent diffusion coefficient for predicting the response to chemoradiation therapy in locally advanced rectal cancer: a systematic review and meta-analysis. *Medicine (Baltimore).* 2015 Feb;94(6):e517.

Xie T, Wang X, Li M, Tong T, Yu X, Zhou Z. Pancreatic ductal adenocarcinoma: a radiomics nomogram outperforms clinical model and TNM staging for survival estimation after curative resection. *Eur Radiol.* 2020 May;30(5):2513-2524.

Yamamoto Y, Ikoma H, Morimura R, Konishi H, Murayama Y, Komatsu S, Shiozaki A, Kuriu Y, Kubota T, Nakanishi M *et al.* Optimal duration of the early and late recurrence of pancreatic cancer after pancreatectomy based on the difference in the prognosis. *Pancreatology.* 2014 Nov-Dec;14(6):524-9.

Yano M, Misra S, Carpenter DH, Salter A, Hildebolt CF. Pancreatic Neuroendocrine Tumors: Computed Tomography Enhancement, But Not Histological Grade, Correlates With Tumor Aggression. *Pancreas.* 2017 Nov/Dec;46(10):1366-1372.

Zins M, Matos C, Cassinotto C. Pancreatic Adenocarcinoma Staging in the Era of Preoperative Chemotherapy and Radiation Therapy. *Radiology.* 2018 May;287(2):374-390.

Zwanenburg A, Vallières M, Abdalah MA, Aerts HJWL, Andrearczyk V, Apte A, Ashrafinia S, Bakas S, Beukinga RJ, Boellaard R *et al.* The Image Biomarker Standardization Initiative: Standardized Quantitative Radiomics for High-Throughput Image-based Phenotyping. *Radiology*. 2020 May;295(2):328-338.