Resectability of Pancreatic Cancer Is in the Eye of the Observer

A Multicenter, Blinded, Prospective Assessment of Interobserver Agreement on NCCN Resectability Status Criteria

Fabio Giannone, MD,* Giovanni Capretti, MD,†‡ Mohammed Abu Hilal, MD, PhD,§‖ Ugo Boggi, MD,¶ Donata Campra, MD,# Carla Cappelli, MD,** Riccardo Casadei, MD,†† Raffaele De Luca, MD,‡‡ Massimo Falconi, MD,*§§ Gabriele Giannotti, MD,‖‖ Luca Gianotti, MD,¶¶ Roberto Girelli, MD,## Paola Gollini, MD,*** Davide Ippolito, MD,††† Giorgio Limerutti, MD,‡‡‡ Lorenzo Maganuco, MD,§§§ Valeria Malagnino, MD,‖‖‖‖ Giuseppe Malleo, MD, PhD,¶¶¶ Mario Morone, MD,### Cristina Mosconi, MD,**** Federica Mrakic, MD,†††† Diego Palumbo, MD,‡‡‡‡ Roberto Salvia, MD, PhD,¶¶¶ Salvatore Sgroi, MD,§§§§ Alessandro Zerbi, MD,†‡ and Gianpaolo Balzano, MD*

Objectives: To determine the reproducibility of the National Comprehensive Cancer Network (NCCN) resectability status classification for pancreatic cancer.

Background: The NCCN classification defines 3 resectability classes (resectable, borderline resectable, locally advanced), according to vascular invasion. It is used to recommend different approaches and stratify patients during clinical trials.

Methods: Prospective, multicenter, observational study (trial ID: NCT03673423). Main outcome measure was the interobserver agreement of tumor assignment to different resectability classes and quantification of vascular invasion degrees. Agreement was measured by Fleiss' k (k = 1 perfect agreement; k = 0 agreement by chance). Sixty-nine computed tomography (CT) scans of pathologically confirmed pancreatic adenocarcinoma were independently reviewed in a blinded fashion by 22 observers from 11 hospitals (11 surgeons and 11 radiologists). Rating differences between surgeons or radiologists and between hospitals with different volumes (≥60 or <60 resections/year) were assessed.

Results: Complete agreement among 22 observers was recorded in 5 CT scans (7.2%), whereas 25 CT scans (36.2%) were variously assigned to all 3 resectability classes. Interobserver agreement varied from fair to moderate (Fleiss' k range: 0.282–0.555), with the lowest agreement for borderline resectable tumors. Assessing vascular contact ≤180° had the lowest agreement for all vessels (k range: 0.196–0.362). The highest concordance was recorded for venous invasion >180° (k range: 0.619–0.756). Neither reviewers' specialty nor hospital volume influenced the agreement.

Conclusions: There is high variability in the assignment to resectability categories, which may compromise the reliability of treatments recommendations and the evidence of trials stratifying patients in resectability classes. Criteria should be revised to allow a reproducible classification.

Keywords: interobserver agreement, NCCN classification, pancreatic cancer, resectability, surgery

From the *Division of Pancreatic Surgery, Pancreas Translational and Clinical Research Center, IRCCS San Raffaele Scientific Institute, Milan, Italy; †Pancreatic Surgery Unit, Humanitas Clinical and Research Center-IRCCS, Rozzano, Milan, Italy; ‡Department of Biomedical Sciences, Humanitas University, Pieve Emanuele, Milan, Italy; §Department of Surgery, Hepato-pancreato-biliary Unit, Fondazione Poliambulanza, Brescia, Italy; ‖Department of Surgery, HPB Unit, Southampton University Hospital, Southampton, United Kingdom; ¶Department of Surgery, University of Pisa and Azienda Ospedaliero Universitaria Pisana, Pisa, Italy; #General Surgery and Emergency Department, AOU Città della Salute e della Scienza di Torino, Turin, Italy; **Department of Radiology, University of Pisa and Azienda Ospedaliero Universitaria Pisana, Pisa, Italy; ††Division of Pancreatic Surgery, IRCCS, Azienda Ospedaliero Universitaria di Bologna and Department of Internal Medicine and Surgery, Alma Mater Studiorum, University of Bologna, Bologna, Italy; ‡‡Department of Surgical Oncology, IRCCS Istituto Tumori "Giovanni Paolo II," Bari, Italy; §§Vita-Salute San Raffaele University, Milan, Italy; ‖‖Department of Radiology, University of Verona Hospital Trust, Verona, Italy; ¶¶School of Medicine and Surgery, Milano-Bicocca University and Department of Surgery, San Gerardo Hospital, Monza, Italy; ##Pancreatic Surgery Unit, "P. Pederzoli" Hospital, Peschiera del Garda, Verona, Italy; ***Radiology and Neuroradiology Unit, San Giovanni Bosco Hospital, ASL Città di Torino, Turin, Italy; †††Department of Radiology, San Gerardo Hospital, Monza, Italy; ‡‡‡Department of Radiology 2, AOU Città della Salute e della Scienza di Torino, Turin, Italy; §§§General Surgery Unit 2 and Pancreas Group, San Giovanni Bosco Hospital, ASL Città di Torino, Turin, Italy; ‖‖‖‖Department of Radiology, IRCCS Istituto Tumori "Giovanni Paolo II," Bari, Italy; ¶¶¶The Pancreas Institute, Department of General and Pancreatic Surgery, University of Verona Hospital Trust, Verona, Italy; ###Department of Radiology, Fondazione Poliambulanza, Brescia, Italy; ****Department of Radiology, Policlinico "S. Orsola Malpighi," Bologna, Italy; ††††Department of Radiology, Humanitas Clinical and Research Center-IRCCS, Rozzano, Milan, Italy; ‡‡‡‡Department of Radiology, IRCCS San Raffaele Scientific Institute, Milan, Italy; and §§§§Department of Radiology, "P.Pederzoli" Hospital, Peschiera del Garda, Verona, Italy.

INTRODUCTION

Only 15%–20% of pancreatic cancer patients undergo resection,[1] because cancer early spreads to other organs and/or peripancreatic vessels, limiting the role of surgery. At diagnosis, distant metastases are evident in more than 50% of patients,[2] while vascular invasion prevents resection in an additional 20%–30% of cases. To grade vascular infiltration, several classifications have been suggested.[3–6] The most widely adopted is the National Comprehensive Cancer Network (NCCN) resectability status classification, which considers 3 categories: resectable (RES), borderline resectable (BR), and locally advanced (LA). In the attempt to tailor a stage-specific treatment, the attribution to a specific resectability class has taken on a key role to drive therapeutic choices, as indicated by international and national guidelines.[6–8] Further, all surgical studies have adopted the resectability categories to stratify patients, aiming to yield homogeneous classes predicting prognosis and technical difficulties. Studies from time to time include RES,[9] or both RES and borderline,[10] or exclusively LA cancers.[11] According to NCCN, the preferred method to define the resectability status is contrast-enhanced computed tomography (CT) scan with a dedicated pancreas protocol.

The attribution to a resectability class is, however, affected by a certain degree of subjectivity. Clinicians should assess the presence/absence of tumor contact with arteries and veins, contour regularity, circumferential involvement, presence of arterial variant, and possibility of safe venous reconstruction. To reduce the risk of variability, the Radiological Society of North America and the American Pancreatic Association defined a pancreatic cancer reporting template, allowing to itemize the characteristics and the infiltration degree of each vessel;[12] the NCCN adapted and incorporated such template.[6] Nevertheless, the interpretation of vascular invasion remains challenging.[13,14] Variability could lead to incorrect therapeutic choices and could compromise clinical studies' reliability in which patients are stratified by resectability status, hampering comparison among researches.

The purpose of this multicenter study was to assess the reproducibility of the NCCN resectability status classification. Secondary outcomes were to assess if a high hospital volume may reduce the variability and to compare evaluations performed by surgeons or radiologists.

METHODS

Study Design and Data Collection

Prospective, multicenter, observational study. The trial was registered on ClinicalTrials.gov (ID: NCT03673423) and endorsed by the Italian Society of Surgical Oncology (SICO).

Participation in the study was proposed to surgeons belonging to the SICO *Pancreas Oncoteam*. Eleven Italian surgical centers with variable pancreatic surgery volume accepted to participate. For each center, 1 expert surgeon and 1 expert radiologist, with at least 5 years of clinical experience in pancreatic cancer, separately evaluated the imaging, filling the NCCN pancreatic cancer radiology reporting template (version 1.2020, section PANC-A).[6] Each rater independently assessed CT scans, blinded to CT scan report, medical history, and other ratings. Eligible patients had pathologically confirmed pancreatic adenocarcinoma, without distant metastases, and were naive from any pancreatic cancer treatment. Patients were enrolled, after signing an informed consent, in 3 very high-volume centers (San Raffaele Hospital, Milan; Humanitas Research Hospital, Rozzano; Hospital Trust, Verona) from January to May 2019. CT scans were performed either at the 3 institutions or other hospitals, providing that they meet the following requirements:

multidetector CT scans, including pancreatic parenchymal phase and portal venous phase axial scans, with a section thickness <3 mm as suggested by NCCN.[6] A pool of 138 CT scans was collected, 69 of which were randomly selected for evaluation. Imaging files were anonymized, randomly labeled with an identification number and provided to the centers on a flash drive. All completed templates were reported in excel files and sent for the analysis to the coordinator center (San Raffaele Hospital). The STrengthening the Reporting of Observational studies in Epidemiology flow chart is available (see Figure, Supplemental Digital Content 1, http://links.lww.com/AOSO/A46, showing patients' recruitment data).[15] The study protocol was approved by the institutional review board of 3 hospitals where CT scans were collected.

Study Outcomes

Primary outcome was to assess the interobserver agreement in applying the NCCN resectability status classification (version 1.2020).[6] Besides the allocation of a patient to a resectability class (RES, BR, and LA), the infiltration rating of every single vessel was assessed to investigate which vessels could be more responsible for the variability. The tumor-vessel contact was graded as absent, present ≤180°, or present >180°. As indicated by NCCN, either direct tumor-vessel contact or contact of the vessel with increased hazy attenuation or stranding was considered. The arterial evaluation included superior mesenteric artery, celiac axis (CA), and common hepatic artery (CHA); the venous assessment included main portal vein (MPV) and superior mesenteric vein (SMV).

Secondary aims were to assess if the agreement could be influenced by hospital volume or by the reviewers' specialty (surgeons or radiologists). Hospitals were divided into 2 categories by applying a cutoff of 60 resections/year, according to the findings of a recent nationwide analysis in Italy.[16] In such study, hospitals performing ≥60 resections/year belonged to the high- or very high-volume categories, whereas those with <60 resections/year to medium- or low-volume categories. Six centers performed ≥60 resections/year (range 61–458) and 5 centers <60 resections/year (range 20–54). We defined as complete agreement the assignment of a CT scan to the same resectability class or the same vessel invasion grade by all 22 observers, or observers of the same subgroup when subgroups were analyzed. We defined as major disagreement the CT scan assignment to all 3 resectability classes or vessel contact grades. The ratings of subgroups of observers were compared to assess differences and explore the possibility that a more or less advanced rating could be influenced by the volume classes or by the appraisers' specialty.

Statistical Analysis

The sample size was calculated a priori using the interobserver agreement as the primary endpoint, evaluated by the Fleiss' k statistic.[17] With a relative error of 20% and an expected difference between the overall agreement probability and the chance-agreement probability of 0.6, the required sample size for 2 observers would be 69 cases. Starting from this assumption, a greater number of observers could only increase the study power. A descriptive data analysis was performed and compared for the cluster of observation stratified by volume and observer type. Categorical data were expressed as absolute and relative frequencies, with distribution between groups assessed using the χ^2 test, or Fisher exact test. An inter-rater agreement analysis using the k statistic was performed to determine consistency among observers on resectability status and degree of vascular involvement for every vessel. Interobserver agreement was graded using suggested benchmarks: 0–0.2, poor agreement; 0.21–0.4, fair agreement; 0.41–0.6, moderate

agreement; 0.61–0.8, substantial agreement; 0.81–1, almost perfect agreement.[18] Associated 95% confidence intervals were shown. Besides overall analysis including all raters, k values were calculated according to the different subgroups: surgeons or radiologists and hospitals with ≥60 or <60 resections/year. To compare the ratings of different subgroups, we assigned a score (0–1–2) to the 3 grades of the resectability status (RES—BR—LA, respectively) and the vessel contact (absent: ≤180° to >180°, respectively). Means were calculated for each subgroup and were compared by the z test. Statistical analysis was performed using SPSS (Statistics for Macintosh, Version 26.0, IBM Corp).

RESULTS

Sixty-nine CT scans were randomly selected from the initial pool. Patients' median age was 66 years (range 41–97 years), 39 were men (56.5%).

Resectability Status

When considering all 22 observers, a complete agreement was recorded in 5 cases (7.2%) and a major disagreement in 25 cases (36.2%). Among the 5 cases with complete agreement, 3 were classified as RES and 2 as LA cancers; no patient was rated as BR by all reviewers, nor by any subgroup. No significant difference

between subgroups on complete agreement and major disagreement was found (Table 1). Figure 1 shows the variability in the assignment to resectability classes of each CT scan evaluation. Interobserver agreement on resectability status, measured by Fleiss' k, ranged from fair to moderate (range k = 0.282–0.555), with the lowest agreement for the BR category. Minimal variations of k values were observed for different subgroups (Table 2). Figure 2 summarizes through a heatmap the agreement regarding resectability status and tumor-vessel contact degrees, among all 22 observers and subgroups. Comparison by z test showed no significant difference between subgroups to rate the resectability status (see Figure, Supplemental Digital Content 2, http://links.lww.com/AOSO/A47, showing similarity of the resectability status ratings assigned by observers' subgroups).

Venous Evaluation

In assessing SMV involvement, a complete agreement among all observers was found in 10 cases (14.5%), 8 of them describing absence of contact; major disagreement was found in 23 cases (33.3%) (Table 1). No significant difference between subgroups was detected. Fleiss' k to grade SMV invasion was variable, ranging from fair to substantial. The lower agreement was recorded to describe SMV contact ≤180° (k = 0.294, fair agreement), and the higher to assess contact >180° (k = 0.638,
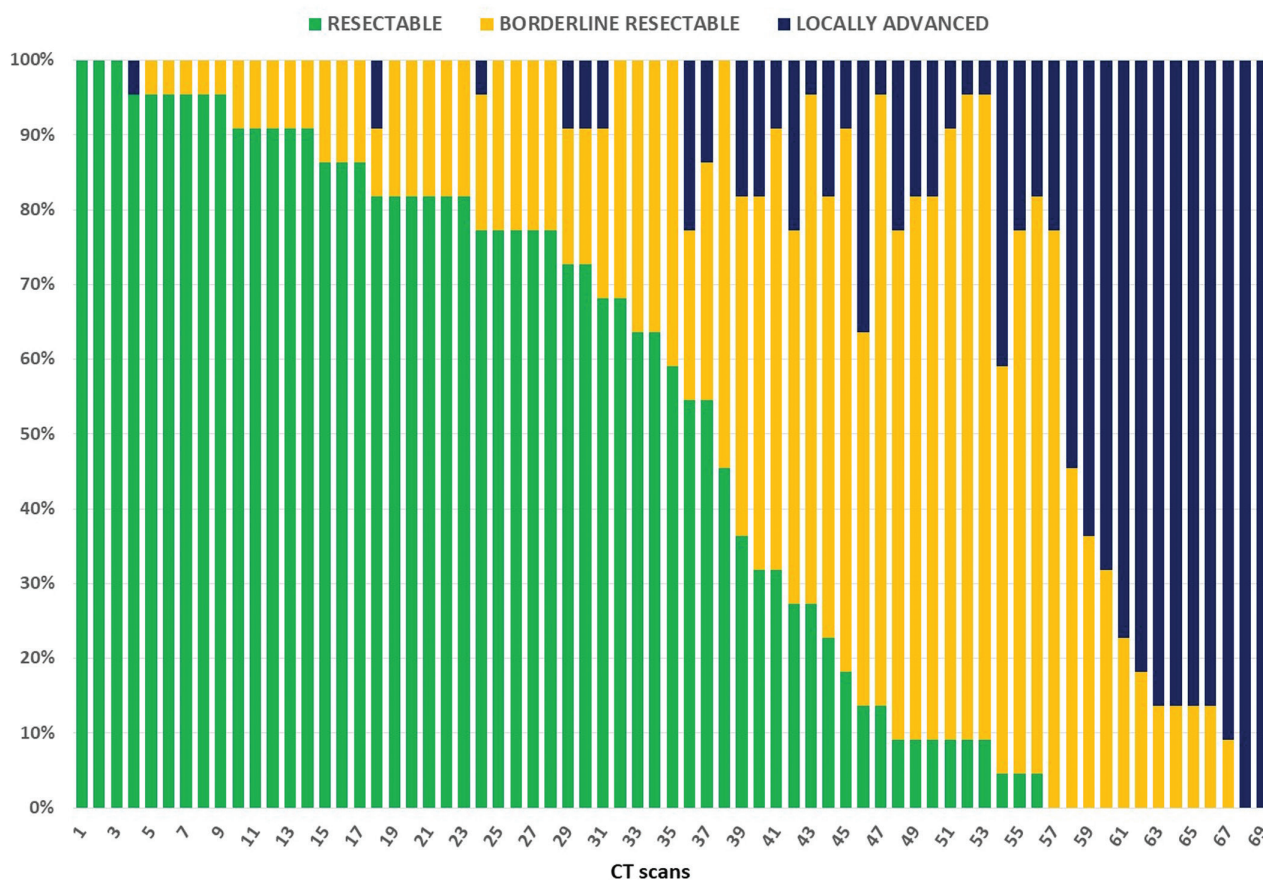
TABLE 1.

Complete Agreement and Major Disagreement in Defining Resectability Status and Tumor-Vessel Contact Among All Observers or Subgroups

Outcome Measures	All Observers	Volume ≥60*	Volume <60*	P	Surgeons	Radiologists	P
Resectability status							
Complete agreement, n (%)	5 (7.2)	9 (13)	16 (23.2)	0.184	13 (18.8)	5 (7.2)	0.074
Resectable	3	7	14		10	3	
Borderline resectable	0	0	0		0	0	
Locally advanced	2	2	2		3	2	
Major disagreement, n (%)	25 (36.2)	10 (14.5)	14 (20.3)	0.501	20 (30)	16 (23.2)	0.561
Tumor-venous contact							
Superior mesenteric vein							
Complete agreement, n (%)	10 (14.5)	10 (14.5)	20 (29)	0.062	16 (23.2)	10 (14.5)	0.276
Absent	8	8	13		11	8	
≤180°	0	0	1		0	0	
>180°	2	2	6		5	2	
Major disagreement, n (%)	23 (33.3)	15 (21.7)	13 (18.8)	0.832	18 (26.1)	17 (24.6)	1
Main portal vein							
Complete agreement, n (%)	21 (30.4)	28 (40.6)	32 (46.4)	0.606	25 (36.2)	28 (40.6)	0.726
Absent	20	26	27		22	27	
≤180°	0	0	0		0	0	
>180°	1	2	5		3	1	
Major disagreement, n (%)	17 (24.6)	14 (20.3)	5 (7.2)	**0.046**	15 (21.7)	10 (14.5)	0.377
Tumor-arterial contact							
Superior mesenteric artery							
Complete agreement, n (%)	24 (34.8)	30 (43.5)	41 (59.4)	0.088	36 (52.2)	26 (37.7)	0.123
Absent	24	29	38		35	25	
≤180°	0	0	1		0	0	
>180°	0	1	2		1	1	
Major disagreement, n (%)	16 (23.2)	10 (14.5)	5 (7.2)	0.274	10 (14.5)	12 (17.4)	0.817
Celiac axis							
Complete agreement, n (%)	29 (42)	31 (44.9)	52 (75.4)	**<0.001**	44 (63.8)	31 (44.9)	**0.039**
Absent	29	30	51		44	31	
≤180°	0	0	0		0	0	
>180°	0	1	1		0	0	
Major disagreement, n (%)	13 (18.8)	9 (13)	7 (10.1)	0.791	9 (13)	10 (14.5)	1
Common hepatic artery							
Complete agreement, n (%)	15 (21.7)	17 (24.6)	48 (69.6)	**<0.001**	41 (59.4)	16 (23.2)	**<0.001**
Absent	15	17	47		41	16	
≤180°	0	0	0		0	0	
>180°	0	0	1		0	0	
Major disagreement, n (%)	23 (33.3)	14 (20.3)	7 (10.1)	0.154	14 (20.3)	15 (21.7)	1

In case of complete agreement, the categories to which CT scans were assigned are indicated. P refers to subgroups' comparison. Bold mean that the value is significant (P < 0.05).
*Resections/year.

FIGURE 1. Agreement and disagreement among 22 reviewers about resectability status evaluation of 69 CT scans. On *x*-axis, CT scans are listed from higher to lower frequency of assignment to the resectable status. Frequency of assignment to different resectability status is shown on *y*-axis.

substantial agreement). Similar k values were observed in different subgroups (Table 2 and Fig. 2). Regarding MPV assessment, complete agreement and major disagreement were recorded in 21 (30.4%) and 17 cases (24.6%), respectively. Twenty of 21 concordant cases described no venous contact. Major disagreement was higher in reviewers belonging to hospitals with ≥60 resections/year (Table 1). Fleiss' k agreement was moderate among all observers. Similarly to SMV assessment, the lower k value was recorded in defining MPV contact ≤180° (k = 0.258, fair agreement), whereas the agreement was substantial when defining an invasion >180° (k = 0.659; Table 2 and Fig. 2). The variability in the evaluation of each CT scan regarding degrees of venous infiltration is reported (see Figure, Supplemental Digital Content 3, http://links.lww.com/AOSO/A48, showing agreement and disagreement about the assessment of tumor-vessel contact). Hospital volume and observers' specialty did not influence a more or less advanced rating regarding venous invasion (see Figure, Supplemental Digital Content 4, http://links.lww.com/AOSO/A49, showing ratings assigned by observers' subgroups to vascular invasion).

Arterial Evaluation

Complete agreement on superior mesenteric artery involvement among all reviewers was recorded in 24 cases (34.8%), all without tumor contact. Major disagreement was found in 16 cases (23.2%). No difference in such evaluation was found between subgroups (Table 1). Fleiss' k agreement was moderate in all categories, except for contact ≤180° (k = 0.297, fair agreement) (Table 2 and Fig. 2).

CA assessment showed a complete agreement in 29 cases (42%); the raters indicated no contact in all of them. In 13 CT

scans (18.8%), a major disagreement was recorded. Complete agreement was higher for the subgroups of surgeons and of reviewers belonging to hospitals performing <60 resections/year (Table 1). Fleiss' k indicated a moderate or substantial agreement in CA evaluation, except in the assessment of contact ≤180° (k = 0.249, fair agreement) (Table 2 and Fig. 2).

Evaluation of CHA invasion showed complete agreement in 15 CT scans (21.7%), all describing absence of vascular contact, and major disagreement in 23 cases (33.3%). Similarly to CA assessment, complete agreement in CHA evaluation was higher for both surgeons and reviewers of hospitals with <60 resections/year (Table 1). Fleiss' k for CHA involvement was generally moderate, but it was fair when assessing contact ≤180° (k = 0.254), or even poor among radiologists (k = 0.196) (Table 2 and Fig. 2). The variability in the evaluation of each CT scan regarding arterial involvement is reported (see Figure, Supplemental Digital Content 3, http://links.lww.com/AOSO/A48, showing agreement and disagreement about the assessment of tumor-vessel contact). In assessing arterial invasion, hospital volume and specialty did not influence a more or less advanced rating, as proved by z test (see Figure, Supplemental Digital Content 4, http://links.lww.com/AOSO/A49, showing ratings assigned by observers' subgroups to vascular invasion).

DISCUSSION

"Beauty is in the eye of the beholder," is an aphorism suggesting that beauty cannot be standardized. Evidence-based medicine differs from beauty, as it requires objective parameters, without which the results of a trial could not be reproducible. The present study showed that the NCCN resectability status of pancreatic

TABLE 2.

Interobserver Agreement Measured by Fleiss k (95% CI) in the Assessment of NCCN Resectability Status and Tumor-Vessel Contact, Among All Observers or Subgroups

Outcome Measures	All Observers	Volume ≥60 Resections/Year	Volume <60 Resections/Year	Surgeons	Radiologists
Resectability status					
Resectability status	0.460 (0.459–0.460)	0.495 (0.494–0.496)	0.452 (0.451–0.452)	0.486 (0.486–0.487)	0.411 (0.410–0.412)
Resectable	0.555 (0.554–0.555)	0.602 (0.601–0.602)	0.540 (0.539–0.541)	0.579 (0.578–0.580)	0.506 (0.505–0.507)
Borderline resectable	0.282 (0.281–0.282)	0.314 (0.313–0.315)	0.281 (0.280–0.283)	0.302 (0.300–0.303)	0.232 (0.231–0.233)
Locally advanced	0.548 (0.548–0.549)	0.573 (0.572–0.574)	0.539 (0.538–0.541)	0.578 (0.577–0.579)	0.511 (0.510–0.512)
Tumor-venous contact					
Superior mesenteric vein	0.427 (0.427–0.428)	0.428 (0.428–0.429)	0.458 (0.457–0.458)	0.436 (0.436–0.437)	0.392 (0.391–0.393)
Absent	0.420 (0.420–0.421)	0.414 (0.413–0.414)	0.452 (0.451–0.453)	0.424 (0.423–0.425)	0.383 (0.382–0.384)
≤180°	0.294 (0.294–0.295)	0.321 (0.321–0.322)	0.307 (0.306–0.308)	0.306 (0.305–0.307)	0.253 (0.252–0.254)
>180°	0.638 (0.637–0.638)	0.627 (0.626–0.628)	0.669 (0.668–0.670)	0.647 (0.646–0.648)	0.619 (0.618–0.620)
Main portal vein	0.505 (0.504–0.505)	0.499 (0.498–0.499)	0.551 (0.550–0.551)	0.495 (0.494–0.496)	0.491 (0.490–0.491)
Absent	0.568 (0.568–0.569)	0.569 (0.568–0.570)	0.585 (0.584–0.586)	0.559 (0.558–0.560)	0.556 (0.555–0.557)
≤180°	0.258 (0.258–0.259)	0.270 (0.269–0.271)	0.270 (0.269–0.271)	0.223 (0.222–0.224)	0.260 (0.259–0.261)
>180°	0.659 (0.659–0.660)	0.647 (0.646–0.648)	0.756 (0.755–0.757)	0.657 (0.656–0.658)	0.643 (0.632–0.644)
Tumor-arterial contact					
Superior mesenteric artery	0.450 (0.449–0.450)	0.450 (0.449–0.451)	0.508 (0.507–0.509)	0.497 (0.497–0.498)	0.387 (0.386–0.388)
Absent	0.523 (0.523–0.524)	0.508 (0.507–0.509)	0.579 (0.578–0.581)	0.588 (0.587–0.589)	0.443 (0.442–0.444)
≤180°	0.297 (0.296–0.297)	0.318 (0.318–0.319)	0.352 (0.351–0.353)	0.311 (0.310–0.312)	0.261 (0.260–0.262)
>180°	0.528 (0.528–0.529)	0.560 (0.559–0.561)	0.556 (0.555–0.557)	0.558 (0.557–0.559)	0.494 (0.493–0.495)
Celiac axis	0.528 (0.528–0.529)	0.495 (0.494–0.496)	0.600 (0.599–0.601)	0.560 (0.559–0.561)	0.477 (0.477–0.478)
Absent	0.638 (0.637–0.638)	0.617 (0.616–0.618)	0.699 (0.698–0.700)	0.668 (0.667–0.669)	0.594 (0.593–0.595)
≤180°	0.249 (0.249–0.250)	0.255 (0.254–0.256)	0.277 (0.276–0.278)	0.265 (0.264–0.266)	0.197 (0.196–0.198)
>180°	0.592 (0.592–0.593)	0.527 (0.526–0.528)	0.663 (0.662–0.664)	0.629 (0.628–0.630)	0.539 (0.538–0.540)
Common hepatic artery	0.468 (0.467–0.468)	0.429 (0.428–0.430)	0.552 (0.551–0.553)	0.504 (0.503–0.505)	0.414 (0.414–0.415)
Absent	0.554 (0.554–0.555)	0.532 (0.531–0.533)	0.618 (0.617–0.619)	0.580 (0.579–0.581)	0.507 (0.506–0.508)
≤180°	0.254 (0.254–0.255)	0.228 (0.227–0.229)	0.336 (0.335–0.337)	0.309 (0.308–0.310)	0.196 (0.195–0.197)
>180°	0.508 (0.508–0.509)	0.462 (0.461–0.463)	0.578 (0.577–0.580)	0.522 (0.521–0.523)	0.476 (0.475–0.477)

CI indicates confidence interval.

cancer is scarcely reproducible among different observers, regardless of hospital volume or raters' specialty. This finding is relevant in both clinical practice and research. The resectability status drives therapeutic decisions: according to NCCN guidelines, RES cancer may proceed to primary surgery, whereas neoadjuvant treatment is suggested in case of borderline cancer. In addition, patients' stratification according to resectability status is applied in all studies dealing with pancreatic cancer surgery. Owing to the subjectivity to assess vascular invasion, patients assigned to the same resectability class may be inhomogeneous, compromising the reliability of the study results and hampering the comparison among studies. In the present study, only 5 CT scans out of 69 were assigned to the same resectability class by all observers. Conversely, the same CT scan was assigned to any of the 3 resectability classes in more than one-third of cases (36.2%). This finding undermines the reliability of treatment recommendations based on resectability, as well as the evidence of studies stratifying patients in resectability classes.

NCCN criteria mainly rely on the absence/presence of a contact between tumor and vessels and the quantification of circumferential involvement. NCCN resectability status considers either direct tumor-vessel contact or contact of the vessel with increased hazy attenuation or stranding, whose presence may be more challenging to define. Further factors are the vessel wall irregularity, the possibility of a safe vein reconstruction, or CA resection, the involvement of aorta or vena cava and the contact with an arterial variant, if present. The assignment to a resectability class results from combinations of all these factors; a low reproducibility of the classification could be expected as reported by previous experiences.[13,14] A recent article from Germany evaluated pre-therapeutic CT scans or Magnetic Resonance to define the consensus among 5 experienced surgeons in assessing pancreatic cancer's resectability.[14] In such study, only BR and LA cancers were included, criteria for resectability were slightly different from NCCN, and authors did not select CT scans by

minimum quality requirements. Authors found high variability in the assessment of BR cancers, especially for evaluating the venous-tumor contact. The agreement was strong only for identifying arterial contact, but quantification of contact degrees (≤180° or >180°) reduced the agreement. A further retrospective research from Korea evaluated the agreement in assigning pancreatic cancer patients to resectability categories, among 8 radiologists with different abdominal imaging experience (6–10 years or 1–2 years of experience).[13] In such study, NCCN criteria for resectability were partially modified, and in one-third of CT scans, the section thickness was greater than recommended by NCCN (4–5 mm). They found a moderate agreement for resectability classification (k = 0.48), with slightly higher agreement among more experienced reviewers.

Differently from previous studies, our research was prospectively planned to address specifically the reproducibility of the NCCN resectability status classification, adopting quality standards to select CT scans, and using the pancreatic cancer reporting template suggested by NCCN. Overall interobserver agreement on resectability status was graded as moderate (k = 0.460), in a scale where a k of 1 indicates perfect agreement, and a k of 0 indicates agreement equivalent to chance. As previously noted,[14] the lowest agreement was found for the BR category (k = 0.282), for which a complete agreement was never recorded, both considering all reviewers and considering subgroups of reviewers. The present study also evaluated the concordance on any single vessel involvement, to investigate if some vessels could be more responsible for the disagreement than others. There was no vessel to which the disagreement can be mainly attributed: agreement was moderate for all vessels, with k values ranging from 0.427 (SMV) to 0.528 (CA). Regarding the quantification of tumor-vessel contact, concordance was always lower (fair agreement) when assessing a tumor-vessel contact ≤180°, for all vessels and all subgroups. K values for this category ranged from 0.196 to 0.362. On the other side, higher agreement was

Agreement color-scale

| 0 | poor | 0,2 | fair | 0,4 | moderate | 0,6 | substantial | 0,8 | almost perfect | 1 |

		All observers	Volume>60 resections/year	Volume<60 resections/year	Surgeons	Radiologists
		Agreement on resectability status				
Resectability status		0,46	0,495	0,452	0,486	0,411
	RES	0,555	0,602	0,54	0,579	0,506
	BR	0,282	0,314	0,281	0,302	0,232
	LA	0,548	0,573	0,539	0,578	0,511
		Agreement on vessels involvement				
Veins	SMV	0,427	0,428	0,458	0,436	0,392
	MPV	0,505	0,499	0,551	0,495	0,491
Arteries	SMA	0,45	0,45	0,508	0,497	0,387
	CA	0,528	0,495	0,6	0,56	0,477
	CHA	0,468	0,429	0,552	0,504	0,414
		Agreement on contact circumferential degrees				
Absence	SMV	0,42	0,414	0,452	0,424	0,383
	MPV	0,568	0,569	0,585	0,559	0,556
	SMA	0,523	0,508	0,579	0,588	0,443
	CA	0,638	0,617	0,699	0,668	0,594
	CHA	0,554	0,532	0,618	0,58	0,507
Contact <180°	SMV	0,294	0,321	0,307	0,306	0,253
	MPV	0,258	0,27	0,27	0,223	0,26
	SMA	0,297	0,318	0,352	0,311	0,261
	CA	0,249	0,255	0,277	0,265	0,197
	CHA	0,254	0,228	0,336	0,309	0,196
Contact >180°	SMV	0,638	0,627	0,669	0,647	0,619
	MPV	0,659	0,647	0,756	0,657	0,643
	SMA	0,528	0,56	0,556	0,558	0,494
	CA	0,592	0,527	0,663	0,629	0,539
	CHA	0,508	0,462	0,578	0,522	0,476

FIGURE 2. Heatmap showing the interobserver agreement on resectability status and degrees of tumor-vessel contact for all observers and subgroups. Agreement is shown by k values and corresponding colors. SMA indicates superior mesenteric artery.

recorded in the quantification of venous invasion >180°, for both SMV and MPV: in this class, the agreement was substantial, with k values ranging from 0.619 to 0.756. In assessing arteries, the agreement was generally moderate; only the definition of no contact between tumor and CA reached a substantial agreement for almost all reviewers groups, with k values ranging from 0.594 to 0.699. The heatmap of Figure 2 facilitates the understanding of how disagreement mainly relies on BR status and the quantification of tumor-vessel contact ≤180°.

A further aim of the present study was to test the hypothesis that the agreement could be higher for raters with more experience in treating pancreatic cancer, as previously suggested.[13] The threshold of 60 resections/year per hospital was adopted, since a recent analysis in Italy, identified hospitals performing

≥60 resections/year as high-or very high-volume hospitals.[16] Results showed that the agreement was similar among raters belonging to different volume subgroups; in some items, complete agreement was even slightly higher among lower-volume hospital reviewers. The lack of difference regarding hospital volume could probably be attributed to the relatively good experience and specific interest in pancreatic cancer of all participating centers, even in lower-volume subgroups. All surgeons belonged to the SICO Pancreas Oncoteam, and true low-volume hospitals were not included in this study.

The agreement was also similar when considering the reviewers' specialty: minimal variations were recorded in radiologists and surgeons' judgments. The possibility that specific subgroups could tend a more optimistic or more pessimistic assessment of

vascular invasion was also investigated, but no significant differences were found.

The present study did not consider the ability to predict tumor resection based on the resectability status assignment, because contraindications to resection due to vascular involvement may vary according to local policies and surgical skills. Thanks to neoadjuvant therapy and skills improvements, simultaneous venous and arterial resections with complex vascular reconstructions may be presently indicated in referral centers.[19,20] Some years ago, NCCN called unresectable a tumor with arterial encasement (>180°). At present, such tumors are called LA, and series reporting successful resection of LA cancers through arterial resection or arterial divestment have been published.[21,22] The surgical judgment on resectability of pancreatic cancer invading peripancreatic vessels seems no more related to classification, but it is in the eye (and one should also say in the hands) of the surgeon. Future studies will better define the indications and prognostic benefit of such an aggressive surgical strategy.

The present study has some limitations. As recommended by NCCN, only CT scans with section thickness <3 mm were selected, but CT scans were performed in either referral centers or peripheral centers. A higher agreement among observers could be expected if all CT scans were performed in referral centers, with thinner section thickness (0.5–1 mm), and multiplanar reconstructions. Each appraiser individually evaluated CT scans, without a multidisciplinary assessment, that could improve the agreement. Nevertheless, the images assessment applied in this study reflects real life, in which the therapeutic decision-making often occurs in an outpatient setting, on a CT scan performed elsewhere.

In conclusion, this study highlights the need to test the reproducibility of any classification influenced by subjectivity before its introduction in clinical practice. In the last decade, surgeons stratified pancreatic cancer patients without the resectability status classification being validated, potentially comparing inhomogeneous patients' populations or choosing inappropriate treatments. Criteria to define the resectability status should be urgently revised, but any change of the present classification should undergo a validation process before its introduction in the clinical practice. At present, the only practical suggestion is a caution to use the NCCN resectability status classification for surgical decisions or in clinical researches.

ACKNOWLEDGMENTS

REFERENCES

1. Huang L, Jansen L, Balavarca Y, et al. Resection of pancreatic cancer in Europe and USA: an international large-scale study highlighting large variations. Gut. 2019;68:130–139.
2. van der Geest LGM, Lemmens VEPP, de Hingh IHJT, et al; Dutch Pancreatic Cancer Group. Nationwide outcomes in patients undergoing surgical exploration without resection for pancreatic cancer. Br J Surg. 2017;104:1568–1577.
3. Vauthey JN, Dixon E. AHPBA/SSO/SSAT consensus conference on resectable and borderline resectable pancreatic cancer: rationale and overview of the conference. Ann Surg Oncol. 2009;16:1725–1726.
4. Katz MH, Marsh R, Herman JM, et al. Borderline resectable pancreatic cancer: need for standardization and methods for optimal clinical trial design. Ann Surg Oncol. 2013;20:2787–2795.
5. Varadhachary GR, Tamm EP, Abbruzzese JL, et al. Borderline resectable pancreatic cancer: definitions, management, and role of preoperative therapy. Ann Surg Oncol. 2006;13:1035–1046.
6. National Comprehensive Cancer Network. Pancreatic Adenocarcinoma (Version 1.2020). 2019. Available at: https://www.nccn.org/professionals/physician_gls/pdf/pancreatic.pdf. Accessed January 18, 2020.
7. Ducreux M, Cuhna AS, Caramella C, et al; ESMO Guidelines Committee. Cancer of the pancreas: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Ann Oncol. 2015;26(suppl 5):v56–v68.
8. Silvestris N, Brunetti O, Bittoni A, et al. Clinical practice guidelines for diagnosis, treatment and follow-up of exocrine pancreatic ductal adenocarcinoma: evidence evaluation and recommendations by the Italian Association of Medical Oncology (AIOM). Cancers (Basel). 2020;12:E1681.
9. Reni M, Balzano G, Zanon S, 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;3:413–423.
10. Versteijne E, Suker M, Groothuis K, et al; Dutch Pancreatic Cancer Group. Preoperative chemoradiotherapy versus immediate surgery for resectable and borderline resectable pancreatic cancer: results of the Dutch randomized phase III PREOPANC trial. J Clin Oncol. 2020;38:1763–1773.
11. Murphy JE, Wo JY, Ryan DP, et al. Total neoadjuvant therapy with FOLFIRINOX in combination with losartan followed by chemoradiotherapy for locally advanced pancreatic cancer: a phase 2 clinical trial. JAMA Oncol. 2019;5:1020–1027.
12. Al-Hawary MM, Francis IR, Chari ST, et al. Pancreatic ductal adenocarcinoma radiology reporting template: consensus statement of the Society of Abdominal Radiology and the American Pancreatic Association. Radiology. 2014;270:248–260.
13. Joo I, Lee JM, Lee ES, et al. Preoperative CT classification of the resectability of pancreatic cancer: interobserver agreement. Radiology. 2019;293:343–349.
14. Wittel UA, Lubgan D, Ghadimi M, et al. Consensus in determining the resectability of locally progressed pancreatic ductal adenocarcinoma - results of the Conko-007 multicenter trial. BMC Cancer. 2019;19:979.
15. von Elm E, Altman DG, Egger M, et al; STROBE Initiative. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. J Clin Epidemiol. 2008;61:344–349.
16. Balzano G, Guarneri G, Pecorelli N, et al. Modelling centralization of pancreatic surgery in a nationwide analysis. Br J Surg. 2020;107:1510–1519.
17. Fleiss JL. Measuring nominal scale agreement among many raters. Psychol Bull. 1971;76:378–382.
18. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics. 1977;33:159–174.
19. Del Chiaro M, Rangelova E, Halimi A, et al. Pancreatectomy with arterial resection is superior to palliation in patients with borderline resectable or locally advanced pancreatic cancer. HPB (Oxford). 2019;21:219–225.
20. Loos M, Kester T, Klaiber U, et al. Arterial resection in pancreatic cancer surgery: effective after a learning curve [published online ahead of print June 12, 2020]. Ann Surg. doi: 10.1097/SLA.0000000000004054.
21. Napoli N, Kauffmann E, Cacace C, et al. Factors predicting survival in patients with locally advanced pancreatic cancer undergoing pancreatectomy with arterial resection. Updates Surg. 2021;73:233–249.
22. Habib JR, Kinny-Köster B, van Oosten F, et al. Periadventitial dissection of the superior mesenteric artery for locally advanced pancreatic cancer: surgical planning with the "halo sign" and "string sign." Surgery. 2021;169:1026–1031.