Multicentre study of multidisciplinary team assessment of pancreatic cancer resectability and treatment allocation

Kirkegård, J; Aahlin, E K; Al-Saiddi, M; Bratlie, S O; Coolsen, M; de Haas, R J; den Dulk, M; Fristrup, C; Harrison, E M; Mortensen, M B; Nijkamp, M W; Persson, J; Søreide, J A; Wigmore, S J; Wik, T; Mortensen, F V

Published in:
The British journal of surgery

DOI:
10.1002/bjs.11093

Publication date:
2019

Document version
Accepted manuscript

Citation for published version (APA):

Terms of use
This work is brought to you by the University of Southern Denmark through the SDU Research Portal. Unless otherwise specified it has been shared according to the terms for self-archiving. If no other license is stated, these terms apply:
• You may download this work for personal use only.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying this open access version

If you believe that this document breaches copyright please contact us providing details and we will investigate your claim. Please direct all enquiries to puresupport@bib.sdu.dk

Download date: 08. Dec. 2020
Multidisciplinary Team Assessments of Pancreatic Cancer Resectability and Treatment Allocation Between Centers: A Multicenter Study

Authors: Jakob Kirkegård (MD)¹, Eirik Kjus Aahlin (MD, PhD)², Mohammed Al-Saiddi (MD)³, Svein Olav Bratlie (MD, PhD)⁴, Marielle Coolsen (MD, PhD)⁵, Robbert J. de Haas (MD, PhD)⁶, Marcel den Dulk (MD, PhD)⁵,¹³, Claus Fristrup (MD, PhD)⁷, Ewen M. Harrison (MD, PhD)⁸, Michael Bau Mortensen (MD, PhD, DMSc)⁷, Maarten W. Nijkamp (MD, PhD)⁹, Jan Persson (MD, PhD)⁴, Jon Arne Søreide (MD, PhD)¹⁰,¹¹, Stephen J. Wigmore (MD)⁸, Trude Wik (MD)¹², Frank Viborg Mortensen (MD, DMSc)¹

Affiliations:

¹ Department of Surgery, HPB Research Unit, Aarhus University Hospital, Palle Juul-Jensens Boulevard 99, 8200 Aarhus N, Denmark
² Department of Gastrointestinal and HPB Surgery, University Hospital of Northern Norway, 9038 Breivika, Norway
³ Department of Radiology, Stavanger University Hospital, Armauer Hansens Vei 30 N4011, Norway
⁴ Department of Surgery, Sahlgrenska University Hospital, SE-41345 Gothenburg, Sweden
⁵ Department of Surgery, Maastricht University Medical Center, Maastricht, The Netherlands
⁶ Department of Radiology, University Medical Center Groningen, PO Box 30001, 9700 RB Groningen, The Netherlands
⁷ Odense Pancreas Center, Department of Surgical Gastroenterology, Odense University Hospital, J.B. Winsløws Vej 4, 5000 Odense C, Denmark
Correspondence to:
Jakob Kirkegård (M.D.), Department of Surgery, HPB Research Unit, Aarhus University Hospital, Palle Juul-Jensens Boulevard 99, 8200 Aarhus N, Denmark
E-mail: jakob.kirkegaard@auh.rm.dk / Phone: +45 22 90 06 04

Funding:
No specific funding was obtained for this study. Jakob Kirkegård is funded by the Danish Cancer Society (R124-A7521). The funder had no role in conduction of this study.

Category:
Original research article
ABSTRACT

Background:
Multidisciplinary team meetings (MDTs) are widely adopted to ensure optimal treatment for cancer patients. Agreements in tumor staging, resectability assessments, and treatment allocation between different MDTs were assessed.

Methods:
A multicenter study on similarity of decisions made in seven different MDTs across Northern Europe. Among patients referred to the MDT at one hospital, 20 patients considered to have non-metastatic pancreatic cancer for evaluation were randomly sampled. One patient with obvious metastatic lesions was excluded. Anonymized clinical information and radiological images were disseminated to the MDTs. All patients were reviewed by the MDTs for radiological tumor-node-metastasis (TNM) stage, resectability assessment, and treatment allocation. Each MDT was blinded to decisions at the other MDTs. Agreements were expressed as raw-% and Krippendorff’s alpha with 95% confidence intervals (CIs).

Results:
A total of 132 evaluations in 19 patients were conducted at the seven MDTs (one evaluation excluded due to technical problems). There were high agreements for TNM stage. There were substantial variations in the resectability assessments with seven patients considered to be resectable at one MDT but unresectable at another MDT. In less than half of the patients (9/19), the MDTs fully agreed on either a curative or palliative strategy. Only a fair agreement in treatment allocation (Krippendorff’s alpha: 0.31; 95% CI: 0.16-0.45) was observed. Among evaluations with the same resectability assessment, there were high agreements in treatment allocation.
Conclusions:

Considerable disparities in MDT evaluations of pancreatic cancer patients exist, including substantial variations in resectability assessments and treatment allocations.
INTRODUCTION

Pancreatic cancer is a major cause of cancer-related death worldwide and associated with a dismal prognosis.\textsuperscript{1} Curative-intent surgery offers the only chance of survival, but less than 20% of patients are eligible for resection at the time of diagnosis because of locally advanced or metastatic disease.\textsuperscript{2} Furthermore, resection for pancreatic cancer is a major surgical procedure with substantial morbidity and mortality.\textsuperscript{3} It is therefore essential to identify patients expected to benefit from surgery while avoiding futile resections in unresectable patients. However, a substantial risk of both under- and overtreatment may occur as a result of variations in management.

Multidisciplinary team meetings (MDTs) are widely adopted to ensure that cancer patients are allocated to the optimal treatment and to mitigate variations in treatment.\textsuperscript{4} A proper staging of the patient’s tumor is a prerequisite for optimal treatment decision. Allocation of patients to the optimal treatment modality relies on the tumor staging and resectability assessment performed at the MDTs and assessment of performance status. While staging is often straightforward in patients with localized or metastatic tumors, patients with borderline resectable or locally advanced tumors can pose a substantial challenge to the MDTs. As these patients constitute up to 25%,\textsuperscript{5} an accurate and uniform resectability assessment is important to select the right patients for surgery. Several different criteria to classify resectability have been proposed,\textsuperscript{6-9} but lack of consensus reflects the challenges in assessing these patients and implies a room for variations in resectability assessment and treatment allocations at the local MDT.

Previous work has shown that MDTs changed the proposed treatment strategy in up to 25% of pancreatic cancers.\textsuperscript{10, 11} Furthermore, other studies reported a considerable variation in judgement of resectability in for example liver resection for colorectal liver metastases.\textsuperscript{12, 13} Yet, a comparison of outcomes across different MDTs has not yet been conducted. The aim of the present study was to examine agreements in tumor staging, resectability assessment, and treatment
allocation in pancreatic cancer patients between the MDTs at pancreatic surgical centers in Northern Europe.

METHODS

Setting and participating centers

A multicenter study with participation of seven pancreatic surgical centers in addition to a reference center (Aarhus University Hospital, Denmark) was performed. Participating sites included Odense University Hospital (Denmark), Stavanger University Hospital and University Hospital of Northern Norway, Tromsø (Norway), Sahlgrenska University Hospital (Sweden), Maastricht University Medical Center and University Medical Center Groningen (The Netherlands), and Royal Infirmary of Edinburgh (United Kingdom). All centers are located within a universal healthcare setting with no competing private institutions for management of pancreatic cancer patients. This study including the dissemination of information on the anonymized patients was approved by the Danish Data Protection Agency (j.nr. 1-16-02-831-17). According to Danish law, ethical approval was not required.

Study design and population

The study was conducted among patients referred to the MDT for tumors in the liver, biliary tract, and pancreas at Aarhus University Hospital, Denmark, between 2013-2017. From these patients, 20 patients with T1-T4 pancreatic cancer without signs of distant metastases were randomly selected for evaluation at the MDTs at the participating centers. One patient was subsequently identified to have widespread metastatic lesions and was excluded from the study. The 20 original evaluations from Aarhus University Hospital were also excluded, as these were performed in an authentic
clinical setting and therefore not entirely comparable with the evaluations at the other hospitals. Furthermore, one evaluation was excluded because of a malfunctioning digital video disc (DVD), leaving 132 evaluations for analyses. Findings in the present study are reported according to the Strengthening the Reporting of Observational studies in Epidemiology (STROBE) guidelines for cross-sectional studies.\(^\text{14}\)

**Information on patients and dissemination of material**

We reviewed each patient’s medical record for information on age, sex, medical history, signs and symptoms, tobacco smoking and alcohol consumption, performance status, and laboratory results available at the time of referral. Each patient was assigned a unique study identification number and thereafter completely de-identified. The information obtained from the medical records was entered into a patient information chart (see example in the Supplementary Material; Appendix 1). For each patient, imaging files from the computed tomography (CT) scan used at the initial MDT evaluation were obtained. The multidetector CT scans were performed according to a standardized biphasic pancreas protocol with the parenchymal/late arterial phase approximately 35 seconds and the portal venous phase approximately 70 seconds after contrast injection. Cross-sectional imaging with slices of 0.625 mm was generated. The imaging files, following de-identification, were copied to a DVD. In a bundle including the patient information charts and detailed instructions, each participating center received a copy of DVDs for each patient using registered mail. No patients received any chemotherapy prior to evaluation.

**Evaluation at the MDTs**

All 19 patients’ imaging files and information chart were reviewed and evaluated at each participating center’s respective MDT. In order to simulate an actual clinical setting, it was required
that no more than two patients were evaluated at one session at a time. Accordingly, each center evaluated the patients during a course of at least 10 MDT meetings. At the MDT, the attending physicians were requested to evaluate the tumor-node-metastasis (TNM) stage (7th edition), assess resectability (resectable; borderline resectable; not resectable [locally advanced disease]; not resectable [metastatic disease]; unable to decide), and suggest a treatment allocation (resection; further investigation [e.g. endoscopic ultrasound, laparoscopy, magnetic resonance (MR) scan]; neoadjuvant chemotherapy; palliative chemotherapy; best supportive care; other) for each patient. Data were entered into REDCap. The tumors were staged according to the American Joint Committee on Cancer (AJCC) staging with some modifications (the use of Nx and Mx categories was allowed if participants were unable to assess nodal status or presence of metastases based on the available material). To mitigate the impact of missing data, records with either missing values of TNM stage or “other” treatment were manually assessed and replaced with an appropriate value based on the comments provided by the participants. All replacements are documented in the Supplementary Material (Appendix 2-3) and were approved by the respective centers.

Statistical analyses

We tabulated descriptive characteristics for the study population. To assess agreements in TNM-staging, resectability assessment, and treatment allocation for each patient between the participating centers, %-agreements (Stata command kappaetc, using the formula by Kilem Gwet) in addition to Krippendorff’s alpha using nominal data were calculated. Krippendorff’s alpha is a kappa coefficient used to assess inter-rater reliability that can be used with any sample size, number of evaluators, and any kind of data in addition to handling missing data appropriately and accounting for random agreement. Krippendorff’s alpha takes values ranging from 0 (or <0 in extreme cases) to 1 with 0 representing no agreement and 1 representing perfect agreement. The interpretation of
such coefficient varies across research fields and values are arbitrary, but useful benchmarks of agreement are: <0: poor/systematic disagreement, 0-0.2: slight, 0.21-0.40: fair, 0.41-0.60: moderate, 0.61-0.80: substantial, and 0.81-1: near perfect. For the %-agreements and Krippendorff’s alpha, associated 95% confidence intervals (CIs) were also computed with the lower limit truncated at 0. Values with Tx, Nx, or Mx and “unable to decide” for resectability assessment were considered as missing in the main analysis. All statistical analyses were performed using Stata 15 (StataCorp LP, Texas, USA).

**Sensitivity analysis**

First, Krippendorff’s alpha was computed using ordinal instead of nominal data for T-stage, resectability assessment, and treatment allocation. Owing to the use of Krippendorff’s ordinal weights, a standard error and thus a CI could not be calculated. Second, Tx, Nx, or Mx values as well as “unable to decide” for resectability assessment were considered as separate data categories instead of missing values.

**RESULTS**

**Patient characteristics**

Nineteen patients with a median age of 69 years (IQR: 60-79 years) were included. Twelve patients were men (Table 1). Fourteen patients were previous or current smokers and 3 patients had an excessive alcohol consumption.

**TNM-staging**
Graphical presentations of T-, N-, and M-stages are shown in Figures 1A-1C. Detailed information on the TNM-staging for each patient by each participating center is available in the Supplementary material (Appendix 4-6). M-stage had the best agreement, while the lowest agreement was seen for T-stage (Table 2). When using ordinal data instead of nominal, T-stage agreement increased from 46.8% to 91.0%. Krippendorff’s alpha values were generally low, ranging from 0.26 (T-stage) to 0.17 (M-stage). When using ordinal data, Krippendorff’s alpha for T-stage increased to 0.62. When considering evaluations of Tx, Nx, or Mx as separate categories and not as missing values, estimates of Krippendorff’s alpha for N- and M-stage attenuated, while the estimate for T-stage was unchanged (Table 2).

**Resectability assessment**

There were substantial variations in resectability assessments (Figure 2). Seven patients were considered to be resectable by at least one center but unresectable by at least one other center. The %-agreement for resectability assessment was 64.1% (95% CI: 50.0%-78.3%) and Krippendorff’s alpha was 0.46 (95% CI: 0.27-0.65). Using ordinal instead of nominal data did not have an impact on Krippendorff’s alpha (0.51). Considering “unable to decide” as a separate category instead of missing values, a decrease was observed for both the %-agreement and Krippendorff’s alpha (Table 2).

**Treatment allocation**

*Overall*

A fair agreement in treatment allocation between the participating centers was observed (%-agreement: 50.4%; 95% CI: 41.2%-59.6% and Krippendorff’s alpha: 0.31; 95% CI: 0.16-0.45). When using ordinal data instead of nominal data, Krippendorff’s alpha increased to 0.55. The
treatment allocation differed between resection and palliative strategy for six patients (Table 3). Likewise, in four patients the decision differed between potentially curative management after further investigation and palliative treatment (Table 3). In nine patients, there was complete agreement on either a curative (resection or further investigation) or palliative strategy (Table 3).

*According to resectability assessment*

132 evaluations were conducted at the seven MDTs. A substantial consistency between resectability assessment and treatment allocation was observed (Table 3 and Figure 4). Among evaluations registered as resectable, the vast majority were allocated to resection. One evaluation was allocated to best supportive care at one center because of concerns related to performance status. Six out of nine evaluations with metastatic tumors were allocated to further investigation in terms of biopsy, staging laparoscopy, or positron emission tomography scan. The remainder was allocated to a palliative approach. Evaluations with borderline resectable cancer were generally allocated to further investigation, while evaluations with locally advanced tumors were most frequently offered palliative treatment.

**DISCUSSION**

Findings from this study suggest a moderate agreement in resectability assessment among patients with pancreatic cancer between MDTs at seven Northern European pancreatic surgical centers. In less than half of the patients, the participants agreed on a potentially curative or palliative strategy. Contrary to this, patients with identical resectability assessments were generally offered the same treatment approach.
The disagreements on a potentially curative or palliative treatment strategy in nine out of 20 patients is concerning, as selecting on a curative or palliative strategy is the main determinant of the prognosis. As the median survival increases dramatically in pancreatic cancer patients undergoing resection with adjuvant chemotherapy (28 months)\(^\text{19}\) instead of palliative chemotherapy alone (7-11 months depending on treatment)\(^\text{20}\) or best supportive care (2 months)\(^\text{21}\), increasing resection rates is important. However, proper selection of patients with a resectable tumor and, preferably, favorable tumor biology is a difficult yet crucial task as demonstrated in this study. 13 out of 15 borderline resectable tumors were allocated to a potentially curative treatment strategy (resection or further investigation), whereas this applied to only six of 34 patients with locally advanced tumors (further investigation or neoadjuvant chemotherapy). These findings support the importance of an accurate and consistent discrimination between these borderline resectable and locally advanced pancreatic cancer.

Different guidelines to define resectability in pancreatic cancer have been suggested (\textit{e.g.} National Comprehensive Cancer Network\(^\text{6}\), Americas Hepato-Pancreato-Biliary Association\(^\text{7}\), Dutch Pancreatic Cancer Group\(^\text{8}\), and Japan Pancreas Society\(^\text{9}\)). Additionally, there may be differences in local traditions and expertise with some centers potentially being less likely to allocate patients with borderline resectable tumors to surgery. In the present study, it was deliberately not prespecified which criteria to use to classify resectability, as the aim was to describe variations in current clinical practice rather than assessing radiological performance. Accordingly, there was some variation in the resectability criteria used in the different centers, although the majority used guidelines from the National Comprehensive Cancer Network (Appendix 8). This may in part explain the observed disagreements in resectability assessments despite a high agreement in TNM staging, as resectability is largely based on the tumor’s relationship to major blood vessels around the pancreas rather than tumor stage.
The literature on outcome, variations, and consistency of MDT evaluations of pancreatic cancer patients are sparse. Two studies found that around 25% had a change in their treatment allocation following an MDT evaluation.\textsuperscript{10,11} One of the reports did not observe any impact on survival rates. Contrary, a Norwegian study found that patients living in counties with an MDT meeting had an improved survival compared with patients living in counties without an MDT meeting,\textsuperscript{22} suggesting that this was explained by more frequent use of chemotherapy. An actual comparison of tumor staging, resectability assessment, and treatment allocation between different MDT meetings has not been conducted. Given the considerable resources allocated to the MDT conferences, the quality should be evaluated with respect to their efficacy in patient care and treatment decision-making.

Some important strengths of this study should be considered. First, the large number of participating centers from different European countries reinforces the external validity and suggest that the findings are applicable at a wider level. Second, the patients were evaluated at a real-life MDT meeting, which mirrors the routine clinical practice as close as possible. Third, the participating centers were blinded to evaluations performed by any of the other participants. Fourth, the MDT evaluations were examined using both raw \%-agreement and Krippendorff’s alpha calculations. Using both measures is a strength as the \%-agreement overestimates the actual agreement, because it does not account for the agreement that will occur by chance, while Krippendorff’s alpha may underestimate it depending on several conditions such as the amount of missing data, distribution of evaluations, and number of variable categories. Therefore, the differences between \%-agreements and Krippendorff’s alpha for TNM-staging were expected and are probably explained by the number of T-stage categories, the amount of missing values for N- and M-stages, and the evaluations of M-stage skewed towards M0. However, for both N- and M-
stages the CIs for Krippendorff’s alpha were wide and results should thus be interpreted with caution.

Some limitations should also be considered. First, the study was limited to 20 patients. Logistical aspects, with the fact that real-life MDT meetings are extremely busy and resource-demanding, were strongly considered when deciding on the sample size. It was thus a priority to include as many centers as possible at the cost of a smaller study population. Findings from this study therefore require confirmation in larger cohorts. Second, regardless of the efforts to resemble a true clinical setting, participating centers may be more likely to allocate patients for resection than they would have been in a real-life clinical setting, as the choice of resection was without consequences to the particular patient under evaluation. Third, the evaluations were based strictly on the CT scan and clinical information. Some centers may prefer other imaging modalities. Also, some CT scans may have had sub-optimal quality. Fourth, the centers had limited information available on the patient’s clinical performance, which may have affected the treatment allocation. However, this is unlikely to have affected the TNM stage or resectability assessments. In a clinical setting, patient performance obviously plays a major role in the treatment allocation. Furthermore, this may also have underestimated the true incidence of borderline resectable cancers, as this definition also includes biological and conditional factors. Despite the limitations of this study, it provides a novel and important framework and infrastructure for assessing outcomes across different MDTs, which could be applied to other tumors with some modifications.

The findings have important clinical implications, confirming the perception that resectability assessment and treatment allocation in patients with pancreatic cancer differ between centers. Although there were disagreements on a potentially curative or palliative approaches in more than half of the patients, the high agreements in treatment strategy among patients with the same resectability assessments are reassuring. These findings substantiate the need of more uniform
resectability assessments to mitigate disparities in pancreatic cancer management. Another potential solution could be to implement a standardized second-opinion system or European referral network, when a center has doubts on the resectability assessment. Such system has already been implemented in Sweden and could be extended to include other countries as well.

In conclusion, this study reveal a moderate agreement in resectability assessment and only a fair agreement in treatment allocations between participating centers. However, agreements on treatment allocations are high within the same groups of resectability assessment.

**AUTHOR CONTRIBUTIONS**

Study idea and design: JK, FVM

Acquisition, analysis, or interpretation of data: JK, EKA, MaS, SOB, MC, RJdH, MdD, CF, EMH, MBM, MN, JP, JAS, SJW, TW, FVM

Drafting of the paper: JK, FVM

Critical revision of the paper for important intellectual content: JK, EKA, MaS, SOB, MC, RJdH, MdD, CF, EH, MBM, MN, JP, JAS, SJW, TW, FVM

The corresponding author confirms that he had full access to all the data in the study and had final responsibility for the decision to submit for publication. All authors agree to be accountable for all aspects of the work.

**DECLARATION OF INTERESTS**

JK: None.

EKA: None.
MaS: None.
SOB: None.
MC: None.
RJdH: None.
MdD: None.
CF: None.
EMH: None.
MBM: None.
MWN: None.
JP: None.
JAS: None.
SJW: None.
TW: None.
FVM: None.

ACKNOWLEDGEMENTS

This study was initiated at the North Sea HPB Meeting in Tromsø, Norway, June 14-17 2017.

Arthur Revhaug played a large role in initiating this study. Kjetil Søreide has contributed to drafting this paper. C.H.C. Dejong provided valuable comments to the paper. All persons are highly acknowledged for their efforts. We are thankful to all the departments for their time spent collaborating on this study.

The data underlying this manuscript contains individual-level data and cannot be shared without violating Danish law. No preregistration exists for the reported studies reported in this article.
REFERENCES


8. DPCG-definitions of resectability in pancreatic carcinoma.


FIGURE HEADINGS

Figure 1. Distribution of TNM stage assessment in the 19 patients. Each bar represents one patient. The x-axes are sorted with increasing T-, N-, and M-stage towards to right. The Y-axis shows the number of evaluations of each T-, N-, and M-stage for each patient.

Figure 2. Distribution of resectability assessments in the 19 patients. Each bar represents one patient. The x-axis is sorted with the least resectable patients towards the right. The Y-axis shows the number of evaluations of each resectability assessment for each patient.

Figure 3. Distribution of treatment allocations in the 19 patients. Each bar represents one patient. The x-axis is sorted with the patients allocated to palliative strategies towards the right. The Y-axis shows the number of evaluations of each treatment allocation for each patient.

Figure 4. Treatment allocation in relation to resectability assessment in the 19 patients.