

Impact of multidisciplinary team discrepancies on comparative lung cancer outcome analyses and treatment equality

Rasmussen, Torben Riis; Gouliaev, Anja; Jakobsen, Erik; Hjorthaug, Karin; Larsen, Lene Unmack; Meldgaard, Peter; Thygesen, Jesper; Bibi, Rana; Møller, Lars B.; Arshad, Arman; Folkersen, Birgitte; Højsgaard, Anette; Saghir, Zaigham; Larsen, Klaus R.; Ravn, Jesper

Published in:
BMC Cancer

DOI:
10.1186/s12885-024-13188-4

Publication date:
2024

Document version:
Final published version

Document license:
CC BY-NC-ND

Citation for pulished version (APA):

Rasmussen, T. R., Gouliaev, A., Jakobsen, E., Hjorthaug, K., Larsen, L. U., Meldgaard, P., Thygesen, J., Bibi, R., Møller, L. B., Arshad, A., Folkersen, B., Højsgaard, A., Saghir, Z., Larsen, K. R., & Ravn, J. (2024). Impact of multidisciplinary team discrepancies on comparative lung cancer outcome analyses and treatment equality. *BMC Cancer*, 24, Article 1423. <https://doi.org/10.1186/s12885-024-13188-4>

Go to publication entry in University of Southern Denmark's Research Portal

Terms of use

This work is brought to you by the University of Southern Denmark.
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

RESEARCH

Open Access



Impact of multidisciplinary team discrepancies on comparative lung cancer outcome analyses and treatment equality

Torben Riis Rasmussen^{1,14*}, Anja Gouliaev¹, Erik Jakobsen², Karin Hjorthaug³, Lene Unmack Larsen⁴, Peter Meldgaard⁵, Jesper Thygesen⁶, Rana Bibi⁷, Lars B. Møller⁸, Arman Arshad⁹, Birgitte Folkersen¹, Anette Højsgaard¹⁰, Zaigham Saghir¹¹, Klaus R. Larsen¹² and Jesper Ravn¹³

Abstract

Introduction This study aimed to evaluate the consistency of lung cancer case assessments across multidisciplinary team (MDT) sites in Denmark. The goal was to appraise the comparability of outcomes between hospitals in a real-world context.

Methods We prepared sixty comprehensive, fictitious lung cancer case stories, complete with images, and distributed them to the four primary lung cancer MDT conferences in Denmark. These cases were subsequently evaluated as had they been ordinary patients during regular MDT meetings. We compared the conclusions on assigned TNM stage and proposed treatment intent using Kappa statistics.

Results The consensus on assigned stage (Stages IA-B, IIA-B, IIIA-B, IV, and undetermined) corresponded to a Fleiss' Kappa-value of 0.62 (95% CI: 0.52–0.71). The overall assessment of curability, categorized as Curable, Incurable, and Undetermined, corresponded to a Kappa-value of 0.72 (CI: 0.61–0.84). However, for cases unanimously judged by all MDT sites to be Stage III, the concordance on treatment intent was poor, with an agreement coefficient of only 0.32 (95% CI: -0.27–0.97).

Conclusion In detail, the level of agreement on assigned stages was less than desired. In consequence, comparative analyses of treatment results from different hospitals or centres may be prone to bias caused by systematic differences in stage assessment or intent of treatment. The least consensus was observed for cases in Stage III, indicating a need for quality improvement efforts to ensure a higher degree of consistency in MDT decisions.

Additional participants were the many and changing colleagues who attended the MDT conferences where the cases were presented and helped to make the decisions described and analysed in this report.

Rasmussen CSH, Secretary at the Danish Lung Cancer Group, was responsible for collecting and organizing the responses from the participating MDT sites.

*Correspondence:
Torben Riis Rasmussen
torbrasm@rm.dk; charlotte.rasmussen@rsyd.dk

Full list of author information is available at the end of the article



© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

Keywords Multidisciplinary team meeting, MDT, Lung cancer, Real world setting

Introduction

Multidisciplinary team meetings (MDTs) have become the model of care planning for patients with cancer including lung cancer worldwide. These meetings serve as the platform where decisions about diagnosis, stage, and optimal treatment are made, and they ensure most correct conclusion on the stage of the patients and the best decision on treatment [1, 2]. In lung cancer, MDTs improve communication, coordination, decision-making, correct conclusion on stage, and adherence to guidelines [3–5] and has been shown to improve survival [6]. Lung cancer stage is the most significant determinant for treatment options and curability. However, other parameters such as age, comorbidities, and performance status also influence the treatment recommendation. Over the last decades, a rapidly expanding array of treatment options has become available, complicating treatment recommendation at lung cancer MDTs. Although evaluation of the quality and reproducibility of MDTs is imperative, very few studies have assessed peer-reviews of cancer MDTs [7–10]. Agreement on stage and treatment for lung cancer is particularly challenging for non-small cell lung cancer stage III [11]. Moreover, imaging evaluation is not binary and may be influenced by experience and local traditions. Thus, applying guidelines in a real-world setting is not always straightforward and involves some level of subjective evaluations, making real-world comparisons between centres challenging.

In Denmark, a small country with universal tax-funded healthcare for all citizens, 85% of lung cancer patients are discussed on MDTs [12]. The lung cancer MDTs consists of respiratory physicians, oncologists, pathologists, radiologist, thoracic surgeons and specialists in nuclear medicine [13]. The assumption that lung cancer MDTs in Denmark evaluate individual clinical cases similarly according to the applicable national guidelines, has never been formally evaluated. In this study, we assess the consistency in the evaluation of clinical stage and treatment intent for lung cancer cases across the four primary lung cancer MDT meetings in Denmark. This could potentially validate our ongoing comparative analysis of outcomes or guide us towards improved consensus.

Materials and methods

Sixty fictitious lung cancer case stories, each complete with clinical and paraclinical information, were constructed for this study. These cases were modelled after real cases diagnosed at one of the participating hospitals. Some basic patient characteristics, such as gender and previous medical procedures like mastectomy or hip replacement, had to be retained to match their images.

The case stories encompassed fictitious information about comorbidities, general condition, Eastern Cooperative Oncology Group (ECOG) performance status, lung function, smoking habits, diagnostic procedures, and results such as histology, programmed death-ligand 1 (PD-L1) expression, and mutations in the epidermal growth factor receptor (EGFR) or anaplastic lymphoma kinase (ALK) gene rearrangements. The actual patients' computed tomography (CT) and positron emissions tomography (PET)/CT scans were anonymized, stored on external hard drives, and distributed to the participating hospitals where the imaging was loaded into the hospitals' PACS system. Each case was assigned a randomized number for each MDT meeting to prevent discussion of cases across MDT sites. The participating MDT meetings included specialists within pulmonology, radiology, nuclear medicine, oncology, thoracic surgery and pathology. Each MDT site evaluated the cases as they would any ordinary clinical cases during their regular MDT meetings and reported their conclusions regarding TNM stage according to the IASLC 8th edition for lung cancer, the suggested treatment, and whether they considered the proposed treatment to be with curative intent to a database in the Danish Lung Cancer Register (DLCR) at Odense University Hospital. For cases, where the MDT found that the results of the diagnostic work-up presented for each case were insufficient to assign a final stage they could provide comments on why they could not reach a final decision about stage or treatment.

Selection of cases

The case stories and corresponding images were selected to cover the full spectrum of cases typically seen at lung cancer MDT meetings, in terms of clinical stage, histology, PD-L1 expression, EGFR mutations, and expected treatment choice, using the national results from the DLCR as a reference. The stage distribution was enriched with cases in clinical stage III, anticipating that this would be the stage with the most significant discussion of curative treatment options.

Validation of constructed cases

All the constructed cases with corresponding images were evaluated by a reference group of experienced clinicians in lung cancer diagnosis and treatment from radiology, nuclear medicine, pulmonology, thoracic surgery, and oncology before the final selection, to validate that the individual cases were realistic and neither overly simplistic nor too ambiguous. Table 1 lists the characteristics of the cases and Fig. 1 presents an example of a case.

Statistical evaluation

The conclusions from the four participating MDT sites on each of the sixty cases were primarily compared using Kappa statistics, introduced by Cohen in 1960 [14]. Kappa statistics is a statistical measure of inter-rater reliability for categorical outcomes, correcting for agreement by chance, which simple percent agreement does not. It is suggested that Kappa results be interpreted as follows: values ≤ 0 indicate no agreement, 0.01–0.20 none to slight, 0.21–0.40 fair, 0.41–0.60 moderate, 0.61–0.80 substantial, and 0.81–1.00 almost perfect agreement [14, 15]. For the current study, Scott/Fleiss' Kappa was used as it provided the highest flexibility with respect to the number of raters and categories in the statistical software package used for the analyses [15]. The results from the study were analysed according to the intention-to-treat principle, thus including in the analyses results from cases where one or several MDT sites could not reach a conclusion on stage or treatment.

A statistical power calculation prior to the study was based on an estimated Kappa value assumed to be 0.9 under the null hypothesis of a high degree of consensus between the four MDT meetings. Thus, with a significance level of 5% and a test power of 80%, it was calculated that at least 52 cases were needed to measure/test whether the agreement between the four multi-disciplinary teams regarding the decision on treatment with curative intent versus palliative treatment was at least equivalent to a Kappa value of 0.8. To compensate for the risk of missed data during the study, a population of sixty cases was chosen.

The primary evaluation of the results in terms of agreement between MDT meetings is based on the Kappa value for the assessment of curative treatment options. Additionally, the agreement on assessments for T-, N-, and M-categories and for the resulting overall stage was evaluated. The results also identify the specific cases where the MDT meetings show disagreement to a greater or lesser extent. Supplementary Chi2 statistics on cross-tabulations were calculated to describe this. Comparison between MDT sites in stage assessment was done with the Kruskal-Wallis test. The statistical analyses were performed with STATA, ver. 17 (StataCorp, College Station, Texas 77845 USA).

Results

The study was conducted over an 18-month period from 2021 to 2022. We received responses for all sixty cases from each of the four MDT sites. In the ensuing overview of the results and subsequent statistical analyses, the four participating MDT sites are anonymously identified by numbers.

Table 2 presents the degree of concordance for the assessment of individual T-, N-, and M-categories, as

well as the overall agreement for each category. The concordance for the T-category was significantly below the desired threshold of 0.80, while the confidence intervals for the N- and M-categories encompassed 0.80. Within each category, there seems to be a higher level of agreement at the extremes, most notably for the N- and M-categories.

Stage assessment

The agreement between the MDTs regarding the stage assessment of each case, with eight stage steps (Stages IA-B, IIA-B, IIIA-C, and IV) and an additional ninth option for undetermined stage, resulted in a combined Kappa-value of 0.62 (95% CI: 0.52–0.71). This value is significantly below the desired level of agreement. The results for each stage step are presented in Tables 3 and 4.

When the stages were grouped into three clinically relevant categories - Localised (Stage IA-IIB), Locally Advanced (Stage IIIA-IIIC), and Disseminated Disease (Stage IV), along with a group for undetermined stage - the combined Kappa-value significantly increased to 0.82 (95% CI: 0.72–0.92), as also shown in Tables 3 and 4 visually represents the variation in stage assessments across different MDT sites within stages IA to IIB. This variation is largely eliminated when Stages IA to IIB are consolidated into a single stage group, Stages I+II.

Assessment on curability

Table 5 illustrates the agreement among MDT sites on the assessment of curability, independent of stage, with the options being Curable, Incurable, and Undetermined. The agreement between assessments corresponds to a Kappa-value of 0.72 (95% CI: 0.61–0.84).

Stage assessment versus curability assessment

The determination of curability is intrinsically linked to the stage assessment, although minor differences in stage assessment, such as between Stage I and II, do not impact the potential for curative treatment, even though they may influence the prognosis.

Table 6 presents the concordance among MDT sites on the potential for curative treatment, contingent on whether there is full agreement on the stage group among all MDT sites. When all MDT sites concurred on the stage group, there was a high level of agreement on the potential for a treatment with curative intent, corresponding to a Kappa-value of 0.84 (95% CI: 0.73–0.95). Conversely, if the MDTs did not unanimously agree on the stage group, the agreement was very low, with a Kappa-value of 0.28 (95% CI: 0.06–0.49).

Table 1 Characteristics of cases. ECOG= Eastern Cooperative Oncology Group. PD-L1 = programmed cell death Ligand 1. EGFR=epidermal growth factor receptor. ALK=anaplastic lymphoma kinase. NOS= not otherwise specified

Characteristics of cases	
Number of cases	60
Males/Females	32/28
Mean Age (Max-Min)	70 y (50–84)
Tobacco smoking	
Never smoker	0
Former smoker	36
Current smoker	22
Unknown	2
ECOG Performance status	
0	20
1	28
2	9
3	3
T-categories	
T1	7
T2	18
T3	6
T4	26
Tx	3
N-categories	
N0	23
N1	2
N2	13
N3	19
Nx	3
M-categories	
M0	32
M1a	4
M1b	3
M1c	17
Mx	4
Stage	
Stage I	17
Stage II	2
Stage III	13
Stage IV	26
Stage undetermined	2
Histology	
Adenocarcinoma	41
Non-small cell carcinoma, NOS	1
Squamous cell carcinoma	11
Small-cell carcinoma	6
Small-cell carcinoma + Adenocarcinoma	1
PD-L1 expression	
< 1%	20
1 – 25%	8
1 – 50%	7
>=50%	19
Not measured	6
EGFR mutations	
exon 19 deletion	4
exon 20 insertion	1

Table 1 (continued)

Characteristics of cases	
p.Leu858Arg in exon 21	1
ALK translocation	
Positive	1

T-, N- & M-categories and Stages according to the IASLC 8th edition for lung cancer is as assessed at one of the four MDT sites

Histology, PD-L1 expression, EGFR mutations, and ALK translocation were given as part of the case information

Table 7 presents a cross-tabulation comparing the consensus among all MDT sites on the main stage groups I-IV versus the agreement on the potential for treatment with curative intent. For only forty-two out of the sixty cases (70%), did all MDT sites reach a consensus on both the stage group and the potential for curative treatment.

Limiting cases to those classified as localised stage (Stages IA-IIIB) by the MDT sites, the agreement on curability was very high. For these seventeen cases, the agreement percentage was 94% (95% CI: 86 – 100%). For the twenty-five cases where all MDT sites agreed on Stage IV, the agreement on non-curability was 98% (95% CI: 94 – 100%). The least concordance on the potential for curative treatment, even when all MDT sites agreed on the stage, was observed for Stage III where the agreement percentage was 67% (95% CI: 37 – 96%), corresponding to a Kappa value of 0.32 (95% CI: -0.27–0.91). Stages IIIA and IIIB were almost evenly split between curable and incurable, reflecting the low level of agreement.

The 16 cases, where the four MDT disagreed on treatment intent, are shown in Table 8, including details on the evaluation of each of these cases. In 11 out of the 16 cases, the disagreement on stage revolved around stage III. In two cases, the MDTs agreed on the stage, however, disagreed on treatment intent (case 3 and 12). There was no systematic overall difference between the MDT sites in their assessment of the stages of the sixteen cases (Kruskal-Wallis, $P > 0.5$).

Discussion

This is the first study to compare decisions on TNM stage and treatment recommendations for cases of lung cancer between MDT centres in Denmark. Additionally, it is the first study to include the whole spectrum of lung cancer stages and the largest in number of cases world-wide. Overall, the study reveals a less-than-optimal agreement among the four primary MDT centres in Denmark when assessing sixty cases spanning the full spectrum of clinical stages in terms of assessment of stage and potential for curative treatment. However, when stage assessments were grouped into the three clinically most relevant groups: Localised stages (Stages I-II), Locally advanced stages (Stages IIIA-IIIC), and Disseminated stage (Stage IV), the agreement improved significantly.

For cases, where MDTs concurred on either localised stage or disseminated disease, there was a high level of agreement on whether the treatment was with curative intent or not – even when including the third possibility of being unable to determine whether curative treatment would be possible.

The study underscores that while it may be straightforward to follow the guidelines and agree on stage and treatment recommendations if a case is described on paper where, for instance, it is stated that there are no signs of metastases or that transbronchial needle aspiration from a certain lymph node station contained malignant cells, it becomes more challenging when an evaluation of images is included. Interpretation of images involves a subjective assessment where for instance a focus with moderately increased fluorodeoxyglucose (FDG) uptake could be judged to represent either inflammation or metastasis depending on the experience or tendency of the clinician interpreting the scan. For the forty-two cases for which all MDTs agreed fully on either localised disease or disseminated disease, there was high concordance between the MDT sites on the possibility or not for treatment with curative intent. This highlights the significance of both the stage and consensus on assigned stage, suggesting that treatment is dictated by guidelines once a decision on the stage is established. However, in contrast, for the eighteen cases without complete agreement on stage groups, the concordance among the four MDT sites regarding curative intent or palliative treatment was notably low.

To date, there have been very few publications on peer-reviews for assessment of differences between MDT sites. Previous papers have reported less than perfect agreement between individual MDT centres. A recent study on retroperitoneal sarcoma with twenty-one cases assessed by twelve MDTs in Great Britain, revealed that agreement was merely slight to fair [7]. An inter-MDT assessment of twenty patients with oesophageal cancer in Denmark revealed, that the disagreement would have impacted treatment for twelve out of twenty patients [8]. A study performed across seven northern European MDT centres, found that seven out of nineteen patients with non-metastatic pancreatic cancers, were considered resectable by one MDT but unresectable by another [10]. Just one previous study has evaluated agreement of MDTs within lung cancer. In a Dutch study, ten patients

Example of a case

A 71-year-old man with chronic back pain after surgery 25 years ago and diabetes type 2 treated with metformin.

Initial investigation started 3 years ago when he presented with fatigue and anemia. He had a CT scan, gastroscopy, and colonoscopy performed, all with normal findings except for the discovery of a small, elongated consolidation in the left upper lobe, which has since been followed with control CT scans.

On the latest CT scan, there was progression of the consolidation which led to the current diagnostic work-up.

He is a current smoker with fifteen cigarettes daily and forty packyears.

FEV1 = 2.3 L (81%) ; DLco of 71%.

Regional lung perfusion scan: Left lung 46%.

ECOG PS 0-1

PET with diagnostic CT:

1. A consolidation of 29mm in the apex of the left upper lobe with moderately increased FDG uptake, suspicious for malignancy.
2. No signs on CT or PET of lymph node involvement or distant metastases.
3. The right flexur of the colon has a wall thickening with increased FDG uptake, colonoscopy is suggested.
4. A tumor in the urinary bladder at the right ostium which cannot be assessed further by PET, suggests a cystoscopy.

EBUS: Endobronchial, there is bilaterally as far as can be seen macroscopically normal conditions. EBUS-TBNA with confirmed lymph node from small lymph nodes in lymph node station 4R and 7. No lymph node is found in 4L. No detected malignancy in TBNA.

CT-guided needle biopsy from the left upper lobe: Lung adenocarcinoma. PDL1 < 1%, EGFR mutation not detected. ALK or ROS1 translocation not detected.

Colonoscopy: Nothing abnormal found.

Cystoscopy with detection of a tumor of 3 cm, which must be further examined.

Sample images [At the MDT all images were available]

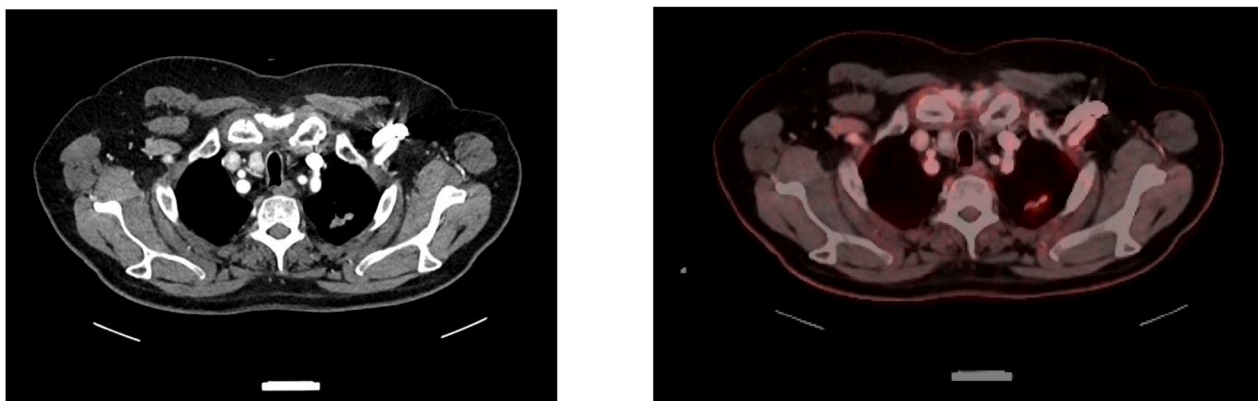


Fig. 1 Example of a case presentation

with stage IIIA non-small cell lung cancer were discussed on eleven MDTs, and agreement was found to be merely moderate [9]. The results of the present study are in line with or higher than the above-mentioned studies. The results regarding concordance in assessments of

T-category are similar to the results in the Dutch study, while the assessments of the N-category had a higher concordance than was found in the Dutch study. However, it is plausible that the concurrence on the N-category in our study could have mirrored that of the Dutch

Table 2 Assessments of T-, N-, and M-categories

T-categories	Kappa (95% CI)
T0	0.49
T1	0.57
T2	0.44
T3	0.40
T4	0.71
Tx	0.02
Kappa combined for T-category	0.54 (0.44–0.65)
N-categories	Kappa (95% CI)
N0	0.91
N1	0.60
N2	0.69
N3	0.84
Nx	-0.01
Kappa combined for N-category	0.79 (0.69–0.88)
M-categories	Kappa (95% CI)
M0	0.81
M1a	0.60
M1b	0.59
M1c	0.90
Mx	0.04
Kappa combined for M-category	0.75 (0.63–0.87)

Table 3 Assessments of TNM single stages and stage groups

TNM single stages	Kappa (95% CI)
IA	0.61
IB	0.54
IIA*	–*
IIB	0.38
IIIA	0.43
IIIB	0.60
IIIC	0.25
IV	0.88
x	0.05
Kappa combined	0.62 (0.52–0.71)
TNM stage groups	Kappa (95% CI)
Stage I + II	0.92
Stage III	0.75
Stage IV	0.88
Stage x	0.05
Kappa combined	0.82 (0.72–0.92)

*) too few cases assessed to this stage

Table 4 Distribution of cases within stages IA-IIIB

Number cStage	MDT Site				Total
	1	2	3	4	
IA	12	9	5	10	36
IB	4	5	12	5	26
IIA	-	-	-	1	1
IIB	4	3	2	4	13
Total	20	17	19	20	76

Table 5 Assessment of curability

Treatment intent	Kappa (95% CI)
Curable	0.74
Incurable	0.79
Undetermined	0.14
Kappa combined	0.72 (0.61–0.84)

Table 6 Agreement on curability dependent on whether all MDT sites agreed on the stage group

Full agreement on stage group:	
Curative intent	Kappa (95% CI)
No	0.86
Yes	0.86
?	-0.01
Kappa combined	0.84 (0.73–0.95)
Without full agreement on stage group:	
Curative intent	Kappa (95% CI)
No	0.45
Yes	0.19
?	0.10
Kappa combined	0.28 (0.06–0.49)

Table 7 The relationship between agreement on Stage Group versus full agreement by all MDT sites on curability ($P < 0.001$)

All MDTs agree on Stage Group	All MDTs agree on curability		
	No	Yes	Total
No	9	2	11
Yes	7	42	49
Total	16	44	60

study more closely, had we confined our analysis for the N-category to Stage III. The concordance for the M-category was lower than in the Dutch study, probably because the Dutch study only included cases in pathological stage IIIA and thus by definition should be without metastases while the current study included a broad spectrum of stages, both with and without metastases.

The current results emphasize that the clinically most challenging stages are the locally advanced stages, stages IIIA-IIIC, as previously reported [11]. First of all, the number of cases assessed to be in Stage III varied between MDT sites from 9 to 13. Secondly it was within this stage group the least concordance between MDT sites in decision on stage and the possibility for curative treatment was found. Even for cases unanimously assessed to be Stage III, the concordance on the possibility of treatment with curative intent was low. The Dutch study also found a wide variation in treatment recommendations but did not pose the question of whether the treatment was with curative intent or not.

However, for cases with differences between MDTs on stage or intent of proposed treatment, MDTs had added comments explained how additional investigation should decide whether a treatment with curative intent could

Table 8 Detailed information about the evaluation of each of the sixteen cases without agreement on stage and/or intent of treatment between the four MDT sites

Case	Stage	Curative/Palliative								Treatment			
		MDT1	MDT2	MDT3	MDT4	MDT1	MDT2	MDT3	MDT4	MDT1	MDT2	MDT3	MDT4
1	IIIB	IIIA	IIIA	IIIB	IIIB	Cur	Cur	Cur	Cur	CRT+adj im	Pal im	Surg ⁷ or CRT+adj im	CRT+adj im
2	? ³	IIIA	? ⁴	IIB	IIB	? ³	Cur	? ⁴	Cur	Surg or Pal C+im	Surg	Surg or Pal C+im	Surg (+adj C)
3	IIIB	IIIB	IIIB	IIIB	IIIB	Pal	Cur ⁵	Pal	Cur	Pal C+im	CRT ⁵ or Pal	Pal C+im	Pal C+im
4	IV	IV	IV	IV	IV	Pal	? [*]	Pal	Pal	Pal C+im	? [*]	Pal C+im	Pal C+im
5	IA	IV	IV	IV	IV	Cur ⁶	Pal	Pal	Pal	Surg ⁶	Pal C	Pal C	Pal C
6	IB	IB	IB	IA	IA	Cur	Cur	Cur	? ⁷	Surg	Surg	Surg	Surg ⁷
7	IIIA	IV	IIIA	IIIA	IIIA	Cur	Pal	Cur	Cur	CRT+adj im	Pal C+im	CRT+adj im	CRT+adj im
8	IIIA	? ⁸	IIIA	IIIA	IIIA	Pal	? ⁸	Pal	Pal	Pal C	? ⁸	Pal C	Pal C
9	? ⁹	IIIA	IIIA	IIIA	? ¹⁰	? ⁹	Cur	Cur	? ¹⁰	Surg+adj C or Pal im ⁹	Surg	Surg	Surg+adj C or Pal im ¹⁰
10	IV	IIIC	IV	IIIC	IIIC	Pal	Pal	Pal	Cur	Pal im	Surg	Pal im	CRT+adj im
11	IA	IA	IA	IIB	IIB	Cur	Cur	Cur	Pal ¹¹	Surg	Surg	Surg	Pal TKI
12	IIIB	IIIB	IIIB	IIIB	IIIB	Cur	Pal	Pal	Cur	CRT+adj im	Pal im	Pal im	CRT+adj im
13	IV	IIIC	IIIC	? ¹²	IIIC	Pal	Pal	Pal	Cur ¹²	Pal C	Pal C	Pal C	CRT
14	IIIB	IIIB	IIIA	IIIA	IIIA	Pal	Cur	Pal	Pal	Pal C	CRT	Pal C	Pal C
15	IV	IV	IIIA ¹³	IIIA	IIIA	Pal	Pal	Cur ¹³	Pal	Pal im	Pal im	Pal C	CRT or Pal im
16	IIIB	? ¹⁴	IIIB	IIIB	IIIB	Cur	Cur	Cur	Cur	Surg	Surg ¹⁴	Surg	Surg

CRT: Chemo-Radiation-Therapy. Surg: Surgery. Pal: Palliative. adj: adjuvant. im: Immunotherapy. C: Chemotherapy. TKI: Tyrosine Kinase Inhibitor

1) PET-positive lymph node (LN) in 1R makes radiation impossible

2) IF TBNA from LN 2R is benign then surgery (single station N2)

3) Examination of small pleural effusion and PET-pos. pericardial LN. If both benign then stage IIIA and surgery. Otherwise, stage IIIB or IV

4) Examination of small pleural effusion and PET-pos. LN. If both benign then stage IIB and surgery. Otherwise, stage IIIB or IV

5) If ECOG Performance status sufficiently good. Otherwise, palliative

6) If successfully treated for cardia-cancer then the patient may be offered surgery for bilateral lung cancer

7) 10 mm PET-pos. mass in breast and inguinal PET-pos. LN must be further investigated before decision possible

8) Further investigations are needed before decision about treatment and prognosis is possible

9) If pleuracensis and biopsy of small nodule benign then Surgery+adj C. Otherwise palliative immunotherapy

10) If biopsy of small nodule benign then surgery+adj C. Otherwise palliative immunotherapy

11) Considered to have multifocal lung cancer

12) If no metastases in liver

13) If metastases are excluded then possibly curative

14) Additional consolidations must be verified as infectious before offered surgery

*) Several lung metastases but no obvious primary tumour and histology non-specific, so in doubt if lung cancer

be offered. However, the inability to decide on stage or treatment options because further information is deemed essential before a decision can be reached is part of the clinical reality of MDT meetings and was therefore included in the analysis as a valid response. From the MDTs comments, it is clear that the main reason for disagreement was due to different interpretations of finding on the images. Our findings are in line with the previous Dutch study by Hoeijmakers et al. [9] on MDT consensus of Stage III cases. They similarly found that a wide range of additional diagnostic procedures was also proposed for the patients to be fully diagnosed. However, agreement or disagreement with respect to pathology was not assessed in the current study. Inclusion might have diminished the degree of concordance between MDT sites further.

In the present study, proposed treatments were categorized into just two categories, curative intent and palliative, as this is the most important distinction for the patients. However, this is obviously a simplification of the real-world situation. A treatment may set out to aim for cure but is later found to fail, either because the disease was more advanced than originally thought, or the tumour turns out to be more resistant to the treatment, or the patient experiences intolerable side effects which leads to termination of the initial treatment. In the example of the small pleural effusion, it may turn out that it did in fact represent pleural metastases and should not have been ignored. On the other hand, if the hope for cure is abandoned from the start and the patient is assigned to palliative treatment, they may have missed the opportunity for curative treatment. In real-life situations, it is often difficult, if not impossible, to say what is the right decision when treatment is initiated. In addition, if similar patients (or cases in this study) tend to be assigned to palliative treatment at one MDT sites while he/she would have been offered the possibility of cure at another MDT it may lead to a biased outcome when comparing the treatment results from the two different hospitals.

The clinicians participating in the current study all had several years of experience in lung cancer diagnostics and treatment. But this also means that they may have accumulated dissimilar experience of bad or good from past cases which may have influenced their evaluation of cases with similarities with past patients.

The less than perfect concordance between MDT sites emphasizes that comparative analyses of results from different hospitals or centres may be prone to bias caused by differences in stage assessment. If the discrepancies observed between different MDT sites or hospitals were merely due to random variations, akin to the random inaccuracies in a laboratory measurement, then we could rely on the average values, provided we have a sufficiently large patient cohort. But besides random variation, it is probable that certain MDT meetings have a propensity

to interpret findings differently. This was evident in our current investigation regarding the decision of stage for cases with localized disease. It appears highly likely that one of the MDT sites categorizes cases differently compared to the other sites. Comparative analyses of treatment results from different hospitals will normally be adjusted for sex, age, histology, and stage while there is an inherent assumption that the MDT meetings at the different hospitals will assign stages to the patients in the same way. But as is shown here this is not necessarily so. Thus, in comparative analyses of treatment results it would be wise to test for differences in MDT evaluations.

One feasible way to address and reduce differences between MDT meetings could be to create a nationally, or even internationally, accessible MDT learning portal with a set of fictional cases that can be evaluated and discussed in a collaborative effort to reach consensus, the consensus being the nearest to a reference. Another model could involve randomly selecting a sample of patients throughout the year, exchanging their medical record and diagnostic results with another MDT site and discuss and resolve any differences in a joint MDT evaluation. If some time has passed since start of treatment, it may also be possible to judge if the initial assessment of the patient and the potential for cure was correct.

Conclusion

In summary, we found that the level of consensus on cases when evaluated at a single stage level was barely within the range strived for. The measures of agreement improved significantly to a level above the lower limit of what is considered desirable when stages were grouped into the three clinically most pertinent groups: Localised stages (Stages I-II), Locally advanced stages (Stages IIIA-IIIIC), and Disseminated stage (Stage IV). If cases were unanimously assessed by all MDT sites to be in either a localised stage (stages I-II) or in stage IV, the consensus on whether treatment with curative intent could be offered was very high. However, complete agreement on stage group was only achieved for forty-two of the sixty cases. For the remaining eighteen cases, the concordance between MDT sites was notably low. Cases judged to be in Stage III particularly fell into this category. Overall, there seems to be potential for improvement to ensure that patients will receive more uniform evaluations of stage and treatment recommendations regardless of where they are diagnosed and evaluated at MDT meetings. Additionally, enhancing these processes could reduce potential bias when comparing outcomes between different hospitals by addressing possible systematic differences in stage assessment and corresponding variations in treatment offered.

However, it is important to note that several other aspects, such as disagreement regarding the histological

or molecular evaluation of biopsies, were not included in the study. Additionally, it was not possible to perform the extra diagnostic procedures that some of the MDT meeting had desired, which might have resolved differences in opinion regarding stage or treatment options.

Ethical and legal considerations

The cases evaluated were not actual clinical cases, but fictional cases constructed to align with the selected images. The essential information necessary to match the images used was preserved, but otherwise, the cases were fictitious to prevent identification of actual patients. The images were fully and irrevocably stripped of any personal identification and could not be linked to a specific individual. The use of fully and irrevocably anonymized data for research is permissible under Danish and European General Data Protection Regulation (GDPR) rules.

Acknowledgements

Not applicable.

Author contributions

Rasmussen TR: Conceptualization, Funding acquisition, Methodology, Project administration, Analysis, Writing - original draft. Gouliava A: Writing - review & editing. Jakobsen E: Validation of cases used, Data collection, Writing - review & editing. Hjorthaug K: Validation of cases used, Data collection, Writing - review & editing. Larsen LU: Validation of cases used, Data collection, Writing - review & editing. Meldgaard P: Validation of cases used, Data collection, Writing - review & editing. Thygesen J: Preparation of cases imaging studies, Writing - review & editing. Bibi R: Data collection, Writing - review & editing. Møller LB: Data collection, Writing - review & editing. Arshad A: Data collection, Writing - review & editing. Folkersen B: Data collection, Writing - review & editing. Højgaard A: Data collection, Writing - review & editing. Saghir Z: Data collection, Writing - review & editing. Larsen KR: Data collection, Writing - review & editing. Ravn J: Data collection, Writing - review & editing.

Funding

AstraZeneca Nordic provided financial support for the project. Aside from the statistical assistance for calculating the test power for appropriate sizing of the study, they had no influence on the project.

Data availability

The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Declarations

Ethics approval and consent to participate

Not applicable. The study is based on fictitious case stories and associated imaging used in the study are fully and irrevocably anonymized. The Danish National Medical Research Ethics Committee (<https://researchethics.dk/>) waived the need for ethical approval and referred to the local Institutional Review Board at Aarhus University Hospital (IRB-AUH) for permission to access selected imaging for full and irrevocable anonymization and subsequent use in the study, for which permission was granted. The need for consent to participate was waived by IRB-AUH as it was deemed unnecessary for the intended use according to Danish national regulations (Danish Health Act, § 42 d (<https://www.retsinformation.dk/eli/ta/2018/1286#P42d>)). The use of fictitious case stories and fully and irrevocably anonymized data for research are permissible under Danish and European General Data Protection Regulation (the GDPR rules – REGULATION (EU) 2016/679 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 27 April 2016, article 26 (<https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32016R0679>)).

Consent for publication

Not applicable. The study is based on fictitious case stories and fully and irrevocably anonymized images, the use of which is permissible under Danish and European GDPR rules as stated and explained above for 'Ethics approval and consent to participate'.

Competing interests

The authors declare no competing interests.

Authors information

Each author is a senior consultant with extensive experience in lung cancer diagnostics at university hospitals or is conducting research into the function of MDT in lung cancer.

Author details

¹Department of Respiratory Diseases and Allergy, Aarhus University Hospital (AUH), Aarhus N DK-8200, Denmark

²Department of Thoracic Surgery, Odense University Hospital (OUH), Odense DK-5000, Denmark

³Department of Nuclear Medicine, AUH, Aarhus N DK-8200, Denmark

⁴Department of X-ray & Scanning, AUH, Aarhus N DK-8200, Denmark

⁵Cancer Department, AUH, Aarhus N DK-8200, Denmark

⁶Department of Clinical Engineering, AUH, Aarhus N DK-8200, Denmark

⁷Department of Pulmonary Diseases, Aalborg University Hospital, Aalborg DK-9000, Denmark

⁸Department of Thoracic Surgery, Aalborg University Hospital, Aalborg DK-9000, Denmark

⁹Department of Pulmonary Diseases, OUH, Odense DK-5000, Denmark

¹⁰Department of Thoracic Surgery, AUH, Aarhus N DK-8200, Denmark

¹¹Pulmonary Diseases Section, Medical Department, Gentofte Hospital, Hellerup DK-2900, Denmark

¹²Department of Pulmonary Infections & Medicine, Bispebjerg Hospital, Copenhagen DK-2400, Denmark

¹³Department of Thoracic Surgery, Rigshospitalet, Copenhagen DK-2100, Denmark

¹⁴The Danish Lung Cancer Group, Odense University Hospital (OUH), Odense DK-5000, Denmark

Received: 23 April 2024 / Accepted: 11 November 2024

Published online: 18 November 2024

References

- Scott Bir B. Multidisciplinary Team Approach in Cancer Care: A Review of the Latest Advancements Featured at ESMO 2021. 2022 Mar 21 [cited 2024 Sep 21]; <https://www.emjreviews.com/oncology/symposium/multidisciplinary-team-approach-in-cancer-care-a-review-of-the-latest-advancements-feature-d-at-esmo-2021-s130621/>
- Walraven JEW, Desar IME, van der Hoeven JJM, Aben KKH, Hillegersberg van R, Rasch CRN, et al. Analysis of 105,000 patients with cancer: have they been discussed in oncologic multidisciplinary team meetings? A nationwide population-based study in the Netherlands. *Eur J Cancer Oxf Engl* 1990. 2019;121:85–93.
- Forrest LM, McMillan DC, McArdle CS, Dunlop DJ. An evaluation of the impact of a multidisciplinary team, in a single centre, on treatment and survival in patients with inoperable non-small-cell lung cancer. *Br J Cancer*. 2005;93(9):977–8.
- Heinke MY, Vinod SK. A review on the impact of lung cancer multidisciplinary care on patient outcomes. *Transl Lung Cancer Res*. 2020;9(4):1639–53.
- Schmidt HM, Roberts JM, Bodnar AM, Kunz S, Kirtland SH, Koehler RP, et al. Thoracic multidisciplinary tumor board routinely impacts therapeutic plans in patients with lung and esophageal cancer: a prospective cohort study. *Ann Thorac Surg*. 2015;99(5):1719–24.
- Stone E, Rankin N, Kerr S, Fong K, Currow DC, Phillips J, et al. Does presentation at multidisciplinary team meetings improve lung cancer survival? Findings from a consecutive cohort study. *Lung Cancer Amst Neth*. 2018;124:199–204.
- Tirotta F, Hodson J, Alcorn D, Al-Mukhtar A, Ayre G, Barlow A, et al. Assessment of inter-centre agreement across multidisciplinary team meetings for patients with retroperitoneal sarcoma. *Br J Surg*. 2023;110(9):1189–96.

8. Achiam MP, Nordmark M, Ladekarl M, Olsen A, Loft A, Garbyal RS, et al. Clinically decisive (dis)agreement in multidisciplinary team assessment of esophageal squamous cell carcinoma; a prospective, national, multicenter study. *Acta Oncol Stockh Swed*. 2021;60(9):1091–9.
9. Hoeijmakers F, Heineman DJ, Daniels JM, Beck N, Tollenaar RAEM, Wouters MWJM, et al. Variation between multidisciplinary tumor boards in clinical staging and treatment recommendations for patients with locally Advanced Non-small Cell Lung Cancer. *Chest*. 2020;158(6):2675–87.
10. Kirkegård J, Aahlin EK, Al-Saiddi M, Bratlie SO, Coolsen M, de Haas RJ, et al. Multicentre study of multidisciplinary team assessment of pancreatic cancer resectability and treatment allocation. *Br J Surg*. 2019;106(6):756–64.
11. Tanner NT, Gomez M, Rainwater C, Nietert PJ, Simon GR, Green MR, et al. Physician preferences for management of patients with stage IIIA NSCLC: impact of bulk of nodal disease on therapy selection. *J Thorac Oncol off Publ Int Assoc Study Lung Cancer*. 2012;7(2):365–9.
12. DLCR. The Danish Lung Cancer registry. [Internet]. [cited 2024 Apr 23]. <https://www.lungecancer.dk/wp-content/uploads/2023/06/%C3%85rsrapport-2022-DLCR-offentlig.pdf>
13. Gouliaev A, Berg J, Bibi R, Arshad A, Leira HO, Neumann K, et al. Multi-disciplinary team meetings for lung cancer in Norway and Denmark: results from national surveys and observations with MDT-MODE. *Acta Oncol Stockh Swed*. 2024;63:678–84.
14. Cohen J. A coefficient of agreement for nominal scales. *Educ Psychol Meas* 1960(20):37–46.
15. Zapf A, Castell S, Morawietz L, Karch A. Measuring inter-rater reliability for nominal data - which coefficients and confidence intervals are appropriate? *BMC Med Res Methodol*. 2016;16:93.

Publisher's note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.