

Contents lists available at [ScienceDirect](https://www.sciencedirect.com)

## European Journal of Cancer

journal homepage: [www.ejccancer.com](http://www.ejccancer.com)

Original research



## TTF-1 status in early-stage lung adenocarcinoma is an independent predictor of relapse and survival superior to tumor grading

Simon Schallenberg<sup>a,\*</sup>,<sup>1,2</sup> Gabriel Dernbach<sup>a,b,c,\*\*</sup>,<sup>1,3</sup> Mihnea P. Dragomir<sup>a,d,e</sup>, Georg Schlachtenberger<sup>f</sup>, Kyrill Boschung<sup>g</sup>, Corinna Friedrich<sup>a,h,i</sup>, Kai Standvoss<sup>b</sup>, Lukas Ruff<sup>b</sup>, Philipp Anders<sup>j</sup>, Christian Grohé<sup>k</sup>, Winfried Randerath<sup>g</sup>, Sabine Merkelbach-Bruse<sup>l</sup>, Alexander Quaas<sup>l</sup>, Matthias Heldwein<sup>f</sup>, Ulrich Keilholz<sup>m</sup>, Jens Khosro Hekmat<sup>f</sup>, Carsten Rückert<sup>n</sup>, Reinhard Büttner<sup>l</sup>, David Horst<sup>a</sup>, Frederick Klauschen<sup>a,c,d,o,p,1</sup>, Nikolaj Frost<sup>q,1</sup>

<sup>a</sup> Institute of Pathology, Charité – Universitätsmedizin Berlin, corporate member of Freie Universität Berlin, Humboldt-Universität zu Berlin and Berlin Institute of Health, Germany

<sup>b</sup> Aignostics GmbH, 10555 Berlin, Germany

<sup>c</sup> BIFOLD - Berlin Institute for the Foundations of Learning and Data, Berlin, Germany

<sup>d</sup> German Cancer Consortium (DKTK), Partner Site Berlin, and German Cancer Research Center (DKFZ), Heidelberg, Germany

<sup>e</sup> Berlin Institute of Health (BIH), Berlin, Germany

<sup>f</sup> Department of Cardiothoracic Surgery, University Hospital Cologne, Germany

<sup>g</sup> Bethanien Hospital, Clinic of Pneumology and Allergology, Center for Sleep Medicine and Respiratory Care, Institute of Pneumology at the University of Cologne, Solingen, Germany

<sup>h</sup> Freie Universität Berlin, Humboldt-Universität zu Berlin, and Berlin Institute of Health, Berlin, Germany

<sup>i</sup> Max Delbrück Center for Molecular Medicine in the Helmholtz Association (MDC), Proteomics Platform, Berlin, Germany

<sup>j</sup> Faculty of Medicine, Semmelweis University, 1085 Budapest, Hungary

<sup>k</sup> Klinik für Pneumologie, Evangelische Lungenklinik Berlin Buch, Berlin, Germany

<sup>l</sup> Institute of Pathology, University Hospital Cologne, Germany

<sup>m</sup> Charite Comprehensive Cancer Center, Charité – Universitätsmedizin Berlin, corporate member of Freie Universität Berlin, Humboldt-Universität zu Berlin and Berlin Institute of Health, Germany

<sup>n</sup> Department of General, Visceral, Vascular and Thoracic Surgery, Charité-Universitätsmedizin Berlin, Germany

<sup>o</sup> Institute of Pathology, Ludwig-Maximilians-University Munich, Thalkirchner Str. 36, 80337 München, Germany

<sup>p</sup> German Cancer Consortium (DKTK), German Cancer Research Center (DKFZ), Munich Partner Site, Heidelberg, Germany

<sup>q</sup> Department of Infectious Diseases and Respiratory Medicine, Charité – Universitätsmedizin Berlin, corporate member of Freie Universität Berlin, Humboldt-Universität zu Berlin and Berlin Institute of Health, Germany

## ARTICLE INFO

## Keywords:

TTF-1

LUAD

Tumor grade

Prognosis

Prognostic factor

Bayesian network

Causal inference

## ABSTRACT

**Objectives:** Thyroid transcription factor 1 (TTF-1) is a well-established independent prognostic factor in lung adenocarcinoma (LUAD), irrespective of stage. This study aims to determine if TTF-1's prognostic impact is solely based on histomorphological differentiation (tumor grading) or if it independently relates to a biologically more aggressive phenotype. We analyzed a large bi-centric LUAD cohort to accurately assess TTF-1's prognostic value in relation to tumor grade.

**Patients and methods:** We studied 447 patients with resected LUAD from major German lung cancer centers (Berlin and Cologne), correlating TTF-1 status and grading with clinical, pathologic, and molecular data, alongside patient outcomes. TTF-1's impact was evaluated through univariate and multivariate Cox regression.

\* Corresponding author.

\*\* Corresponding author at: Institute of Pathology, Charité – Universitätsmedizin Berlin, corporate member of Freie Universität Berlin, Humboldt-Universität zu Berlin and Berlin Institute of Health, Germany

E-mail addresses: [simon.schallenberg@charite.de](mailto:simon.schallenberg@charite.de) (S. Schallenberg), [gabriel.dernbach@charite.de](mailto:gabriel.dernbach@charite.de) (G. Dernbach).

<sup>1</sup> Indicates equal contribution of respective authors.

<sup>2</sup> 36900

<sup>3</sup> 36900

<https://doi.org/10.1016/j.ejca.2023.113474>

Received 30 October 2023; Received in revised form 3 December 2023; Accepted 4 December 2023

Available online 10 December 2023

0959-8049/© 2023 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

Causal graph analysis was used to identify and account for potential confounders, improving the statistical estimation of TTF-1's predictive power for clinical outcomes.

**Results:** Univariate analysis revealed TTF-1 positivity associated with significantly longer disease-free survival (DFS) (median log HR  $-0.83$ ;  $p = 0.018$ ). Higher tumor grade showed a non-significant association with shorter DFS (median log HR  $0.30$ ;  $p = 0.62$  for G1 to G2 and  $0.68$ ;  $p = 0.34$  for G2 to G3). In multivariate analysis, TTF-1 positivity resulted in a significantly longer DFS (median log HR  $-0.65$ ;  $p = 0.05$ ) independent of all other parameters, including grading. Adjusting for potential confounders as indicated by the causal graph confirmed the superiority of TTF-1 over tumor grading in prognostics power.

**Conclusions:** TTF-1 status predicts relapse and survival in LUAD independently of tumor grading. The prognostic power of tumor grading is limited to TTF-1-positive patients, and the effect size of TTF-1 surpasses that of tumor grading. We recommend including TTF1 status as a prognostic factor in the diagnostic guidelines of LUAD.

## 1. Introduction

After a steady increase in cases over the last decades, lung adenocarcinoma (LUAD) now accounts for more than 50% of all lung cancers [1,2]. Disease-related prognostic factors include performance status, sex, tumor stage, targetable molecular alterations, and the tumor grade, classified according to the predominant tumor growth pattern [3–7].

The thyroid transcription factor 1 (TTF-1), also known as NK2 homeobox 1 (NKX2-1), is a 38 kDa nuclear protein encoded by the *NKX2-1* gene [8], and physiologically expressed by surfactant-producing type 2 pneumocytes and Club cells, formerly known as Clara cells, in the lung. TTF-1 is expressed in approximately 70–80% of LUAD and is routinely used to distinguish LUAD from metastases from extrapulmonary adenocarcinomas [9,10]. Having excluded metastases from extrapulmonary adenocarcinomas, the 20–25% of TTF-1 negative carcinomas are nevertheless classified as LUAD. A negative TTF-1 status is associated with an unfavorable cancer phenotype, including a lack of actionable genomic alterations, a reduced performance status, and a higher tumor burden [11–13]. Thus, TTF-1 expression can be a prognostic factor in early-stage lung cancer as well as metastatic disease [11,14–17]. However, it remains to be clarified whether a negative TTF-1 status represents a unique LUAD subtype characterized by higher biological aggressiveness or whether it can be predominantly explained by poor histomorphological differentiation (higher tumor grade).

To address this gap and improve prognosis evaluation, we conducted a causal effect estimation in a large bi-centric cohort of 447 patients with resected LUAD to evaluate the effect sizes and mutual moderation concerning TTF-1-status, grading, and several tumors- and patient-specific parameters.

## 2. Patients and methods

### 2.1. Patient cohort

447 treatment-naïve patients from two German large-volume lung cancer centers (Department of General, Visceral, Vascular and Thoracic Surgery, Charité-Universitätsmedizin Berlin, Germany; Department of Cardiothoracic Surgery, University Hospital Cologne, Germany) having received surgery for stage I-III LUAD between 2009 and 2019 were included into this retrospective study. The reviewed diagnosis was based on the current WHO classification criteria for lung cancer, and grading was determined following recommendations from IASLC/ATS/ERS [18]. Only patients with complete resection (R0) were included, whereas distinct LUAD variants such as mucinous, enteric, colloid, or fetal cancers were excluded, as suggested by IASLC/ATS/ERS [18,19]. Baseline patient demographic data and cancer-related characteristics such as pTNM classification, tumor stage, genomic alterations, type of surgery, and adjuvant therapy were collected. The study was performed according to the ethical principles for medical research of the Declaration of Helsinki and was approved by the Ethics Committee of the Charité University Medical Department in Berlin (EA4/243/21).

### 2.1.1. Histologic evaluation

Hematoxylin and eosin, periodic acid-Schiff, and Elastica-van-Gieson stained slides, as well as immunohistochemical staining of formalin-fixed and paraffin-embedded LUAD specimens, were reviewed by an experienced pathologist (S.S.) for diagnosis, grading, pTNM classification, angioinvasion, lymphatic invasion, and tumor stage according to the 8th edition of the TNM classification (AJCC). Uncertainties were discussed with other pathologists (A.Q., M.P.D., F.K.). TTF-1 immunohistochemistry was performed for primary diagnosis using the same antibody at both institutions (8G7G3/1, Dako, 1:100). Specimens containing artifacts and heavily bleached were newly cut, stained, and reassessed. Positive TTF-1 staining was defined as any positive nuclear staining in the tumor cells (see [Supplementary Figure 1](#) for exemplary staining and an overview of the respective tumor grades).

### 2.1.2. Statistics

Demographics and disease data were described and compared using the Chi2-test. Categorical outcomes for differences in the TTF-1 positive and negative subgroups were assessed in univariate and multivariate Cox regression. Survival outcome was evaluated with the Kaplan-Meier estimator, and follow-up time was calculated with the reverse Kaplan-Meier method. Disease-free survival (DFS) and overall survival (OS) were defined as the time in months from surgery to histologically confirmed or radiologically detected relapse or death from any cause, whichever occurred first. Confidence intervals of the Cox regression models were calculated using bootstrap resampling with  $n = 1000$  iterations, including a missing value imputation from the respective variable's marginal distribution within each bootstrap sampling iteration. P-values were derived from the bootstrap mean and standard deviation of the regression coefficients, which we take as an estimate of the standard error. All analyses were performed with Python using pandas, scipy, seaborn, and lifelines packages. The code is available on 'github.gabrielernbach...'. If not otherwise specified, values are given as median with a 95% confidence interval (95% CI).

### 2.1.3. Causal effect estimation

In randomized controlled trials, statistical confounding in the estimation of treatment effects is excluded by the study design. In contrast, the present study is observational, and possible confounding variables must be considered when estimating statistical effects. To identify these confounding variables, the interactions of known relevant variables were evaluated using a causal Bayesian network to derive the necessary adjustment sets for a reliable estimation of the total causal effects [20]. The respective adjustment sets for a given effect of interest were used as additional predictors in multivariate Cox regression.

## 3. Results

### 3.1. Association of TTF-1-status and grading with clinicopathologic parameters and DFS

Classification of the 447 patients resulted in 397 TTF-1-positive (88.8%) and 50 TTF-1-negative (11.2%) LUAD samples. The

distribution of patient demographic and clinicopathologic characteristics by TTF-1 status is provided in Table 1a. The mutation status of patients according to TTF-1 status is shown in Table 1b.

First, we analyzed the association between TTF-1 status and grading. TTF-1 positive LUADs mainly showed a low- or intermediate tumor grade (70.3% vs. 50%), whereas TTF-1 negative LUADs were more frequently high-grade tumors (50% vs. 29.7%) (Fig. 1A). Second, to examine the association of grading and TTF-1 status in their association with DFS, we compared Kaplan-Meier curves of the different TTF-1 status within grading subgroups (Fig. 1B-D) and vice versa (Fig. 1E-F). In a direct comparison of TTF-1 and grading, a positive TTF-1 status was associated with longer DFS independent of grading (Fig. 1B-D).

Conversely, a strong association of DFS with grading was observed only in TTF-1-positive patients (Fig. 1E-F).

### 3.1.1. Effect of TTF-1, grading, and clinicopathologic parameters on DFS

The univariate analyses revealed a significantly longer DFS (median log HR  $-0.83$  [ $-1.43$ ;  $-0.20$ ];  $p = 0.018$ ) for TTF-1 positive patients (Fig. 2A). Grading did not test significantly for a shorter DFS (median log HR  $0.30$  [ $-0.58$ ;  $1.60$ ];  $p = 0.62$  for G1 to G2 /  $0.68$  [ $-0.24$ ;  $1.89$ ];  $p = 0.34$  for G2 to G3). Among the other clinicopathologic parameters, only the contrast of tumor stage 2 versus tumor stage 3 revealed a significantly shorter DFS (log HR  $1.30$  [ $0.82$ ;  $1.84$ ];  $p < 0.0001$ ). In multivariate analysis, TTF-1 positivity was associated with a

**Table 1a**

Distribution of patients' demographic and clinicopathological characteristics according to the TTF-1 status and tumor grade.

Characteristic	TTF-1 positive ADCs								TTF-1 negative ADCs						P-value
	All patients		G1		G2		G3		G1		G2		G3		
	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%	
<b>Patients</b>	447	100	30	7	249	56	118	26	3	1	22	5	25	5	
<b>Location</b>															
Berlin	318	71	30	7	179	40	77	17	3	1	12	3	17	4	
Cologne	129	29	0	0	70	16	41	9	0	0	10	2	8	2	
<b>Age at diagnosis (median)</b>	65.66		70.35		65.93		64.08		70.87		64.63		66.18		
<b>Sex</b>															
Male	259	58	14	3	145	32	71	16	1	0	10	2	18	4	
Female	188	42	16	4	104	23	47	11	2	0	12	3	7	2	
<b>ECOG</b>															
0	258	58	8	2	152	40	67	18	1	0	13	3	17	4	
1	108	24	14	4	56	15	28	7	2	1	4	1	4	1	
2	13	3	0	0	6	2	4	1	0	0	2	1	1	0	
3	1	0	0	0	0	0	0	0	0	0	0	0	0	0	
NA	67	15													
<b>Smoking status</b>															
Smoker	142	32	7	2	76	24	43	14	0	0	6	2	10	3	
Non-smoker	139	31	11	4	75	24	36	12	1	0	8	3	8	3	
Never-smoker	32	7	3	1	25	8	2	1	1	0	1	0	0	0	
NA	134	30													
<b>Pack-years</b>															
≤ 30	86	19	11	5	49	21	16	7	1	0	4	2	5	2	
> 30	146	33	4	2	78	34	47	20	0	0	7	3	10	4	
NA	215	48													
<b>UICC stage</b>															
I	243	54	23	5	146	33	50	11	1	0	13	3	10	2	
II	101	23	4	1	56	13	30	7	2	0	5	1	4	1	
III	103	23	3	1	47	11	38	9	0	0	4	1	11	2	
<b>Tumor diameter (median in mm)</b>	25		20.5		22		30		13		25		35		
<b>Tumor stage</b>															
T1	204	46	18	4	122	27	46	10	1	0	8	2	9	2	
T2	155	35	9	2	86	19	43	10	0	0	10	2	7	2	
T3	59	13	3	1	30	7	15	3	2	0	4	1	5	1	
T4	29	6	0	0	11	2	14	3	0	0	0	0	4	1	
<b>Nodal status</b>															
N0	303	68	25	6	177	41	70	16	2	0	16	4	13	3	
N1	56	12	3	1	30	7	19	4	0	0	2	0	5	1	
N2	71	16	2	0	35	8	24	6	0	0	4	1	6	1	
N3	3	1	0	0	1	0	2	0	0	0	0	0	0	0	
NA	14	3													
<b>Lymphangioinvasion</b>															
L0	357	80	27	6	208	47	85	19	3	1	16	4	18	4	
L1	90	20	3	1	41	9	33	7	0	0	6	1	7	2	
<b>Hemangioinvasion</b>															
V0	411	92	29	6	232	52	106	24	3	1	20	4	21	5	
V1	36	8	1	0	17	4	12	3	0	0	2	0	4	1	
<b>Adjuvant therapy</b>															
Not done	301	67	27	3	176	40	69	16	2	0	14	3	13	3	
Done	144	32	3	2	72	16	49	11	1	0	8	2	11	2	
NA	2	1													
<b>Relapse</b>															
No	314	70	25	6	183	41	80	18	1	0	12	3	13	3	
Yes	133	30	5	1	66	15	38	9	2	0	10	2	12	3	

Abbreviations: TTF-1, Thyroid transcription factor 1; ADCs, adenocarcinomas; G1, grade 1; G2, grade 2; G3, grade 3; ECOG, Eastern Cooperative Oncology Group Performance Status Scale; NA, not available; UICC, International Union Against Cancer. P-value < 0.05 was considered significant and was flagged with one star (\*). P-value < 0.01 was flagged with two stars (\*\*). #Percentage of tumor grades refers to each individual characteristic (the total of 100% is per row).

**Table 1b**  
Distribution of patient mutation status according to the TTF-1 status and tumor grade.

Characteristic	All patients		TTF-1 positive ADCs						TTF-1 negative ADCs						P
			Tumor grade <sup>#</sup>												
			G1		G2		G3		G1		G2		G3		
No.	%	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%		
<b>Patients</b>	447	100	30	7	249	56	118	26	3	1	22	5	25	5	
<b>EGFR</b>															
WT	171	38	3	2	97	49	43	22	1	1	14	7	13	7	0.449
Mutated	35	8	2	1	14	7	5	3	0	0	1	1	3	2	
Not done	241	54													
<b>ALK</b>															
Negative	236	53	14	6	135	56	63	26	1	0	11	5	12	5	0.707
Positive	4	1	0	0	4	2	0	0	0	0	0	0	0	0	
Not done	207	46													
<b>KRAS</b>															
WT	65	15	1	1	37	26	17	12	1	1	5	3	4	3	0.869
Mutated	79	18	1	1	41	28	23	16	0	0	8	6	6	4	
Not done	303	67													
<b>BRAF</b>															
WT	9	2	2	1	70	50	38	27	1	1	10	7	11	8	0.912
Mutated	132	30	0	0	6	4	2	1	0	0	1	1	0	0	
Not done	306	68													
<b>PIK3CA</b>															
WT	110	25	2	2	58	50	31	27	1	1	11	9	7	6	0.704
Mutated	6	1	0	0	2	2	3	3	0	0	0	0	1	1	
Not done	331	74													
<b>TP53</b>															
WT	63	14	1	1	41	34	11	9	1	1	5	4	4	3	0.168
Mutated	58	13	1	1	25	21	20	17	0	0	6	5	6	5	
Not done	326	73													

Abbreviations: TTF-1, Thyroid transcription factor 1; ADCs, adenocarcinomas; G1, grade 1; G2, grade 2; G3, grade 3; *EGFR*, Epidermal Growth Factor Receptor; *ALK*, Anaplastic Lymphoma Kinase; *KRAS*, Kirsten RAT Sarcoma virus; *BRAF*, proto-oncogene B-Raf; *PIK3CA*, Phosphatidylinositol-4,5-bisphosphate 3-kinase, catalytic subunit alpha; *TP53*, Tumor Suppressor P53. <sup>#</sup>Percentage of tumor grades refers to each individual characteristic (the total of 100% is per row).

significantly longer DFS (log HR  $-0.65$  [ $-1.13$ ;  $-0.09$ ];  $p = 0.05$ ), whereas grading showed again no significant association with shorter DFS again (log HR  $0.12$  [ $-0.63$ ;  $1.14$ ];  $p = 0.73/0.22$  [ $-0.60$ ;  $1.32$ ];  $p = 0.66$ ). In contrast, tumor stage 2 to tumor stage 3 indicated a significant association with a shorter DFS (log HR  $1.14$  [ $0.66$ ;  $1.62$ ];  $p = <0.0001$ ) (Fig. 2B).

### 3.1.2. Comparison of the respective total causal effect of TTF-1 and grading

For a causally informed analysis of the clinicopathological variable interactions, we performed a Bayesian network analysis (Fig. 3A). We listed all known related clinical variables and arranged them by their expert-believed interactions in a graph. Variables (graph nodes) were connected (graph edge) if they were known to affect each other in a causal fashion, additionally their direction of interaction was depicted. The graph can then be evaluated to identify variable pathways that inform how an intervention on one of the variables affects other downstream variables. Knowledge about the variable pathways helps us select the exact set of variables necessary to estimate the unique contribution of an exposure variable on an outcome variable. Using causally informed Bayesian networks to determine adjustment sets provides a structured and principled method for covariate selection, mitigating risks such as collider bias that can arise from naive confounding control in conventional multivariate analyses. Using the network, we identified sex as a necessary adjustment set for estimating the total causal effect of TTF-1 positivity on DFS. The corresponding Cox regression showed a significant hazard ratio (HR  $-0.86$  [ $-1.25$ ;  $-0.41$ ]). A similar network analysis for estimating the total effect of grading on DFS indicated the necessary adjustment set of neoadjuvant therapy and TTF-1. The corresponding Cox regression demonstrated no significant effect on DFS (median log HR  $0.31$  [ $-0.32$ ;  $1.30$ ]/  $0.61$  [ $-0.07$ ;  $1.65$ ]) (Fig. 3B). Finally, for the fitted model, we show what the survival curves look like when we vary a single covariate and keep everything else the same. This is useful for understanding the effect of a covariate in a particular model. In Fig. 3C we simulate such a

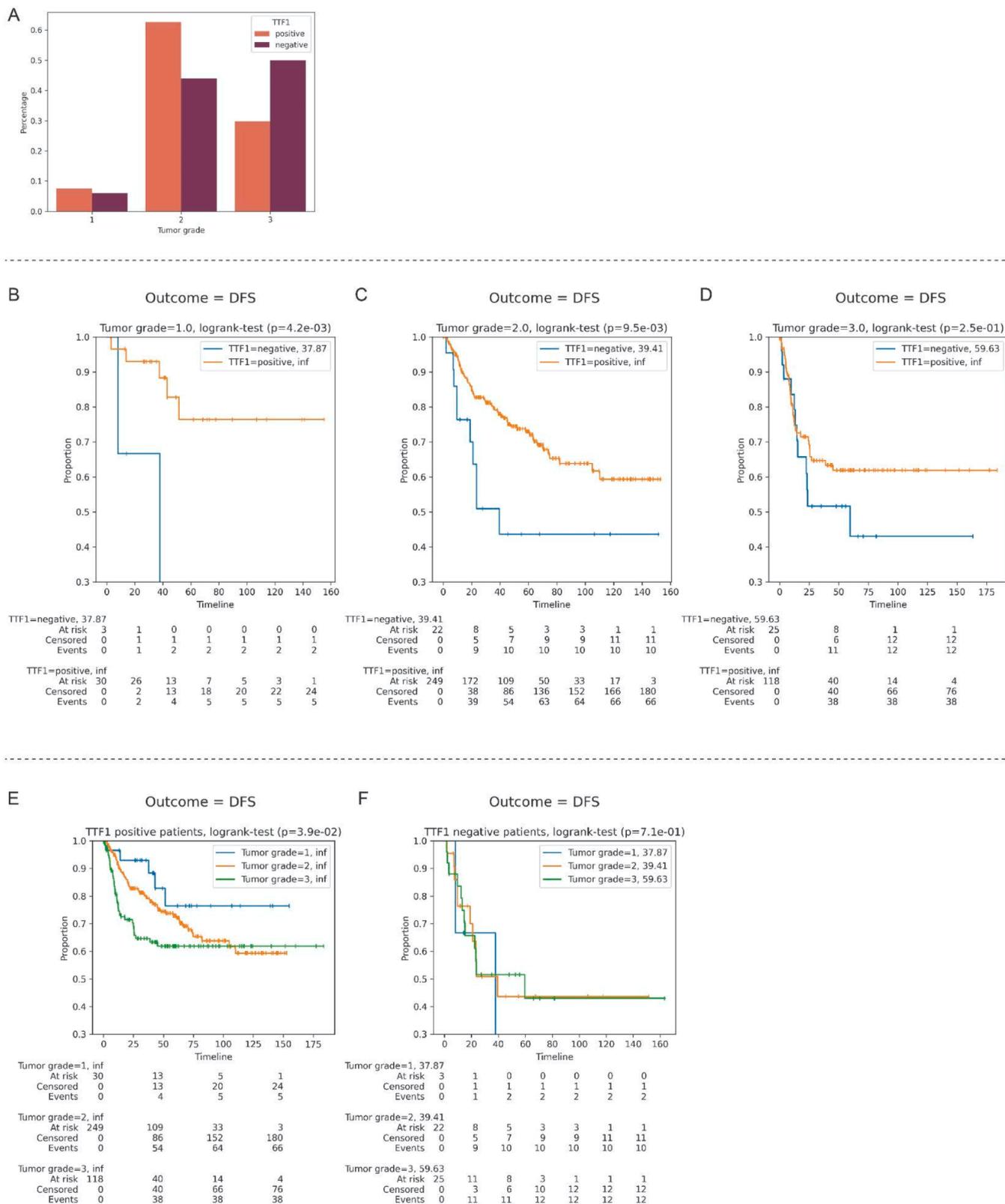
comparison by varying the covariate TTF-1 (Fig. 3C) to the effect of varying the covariate grading (Fig. 3D). The impact of varying TTF-1 is superior in absolute effect to the impact of varying grading, indicating that TTF-1 is a more relevant prognostic parameter than tumor grading.

## 4. Discussion

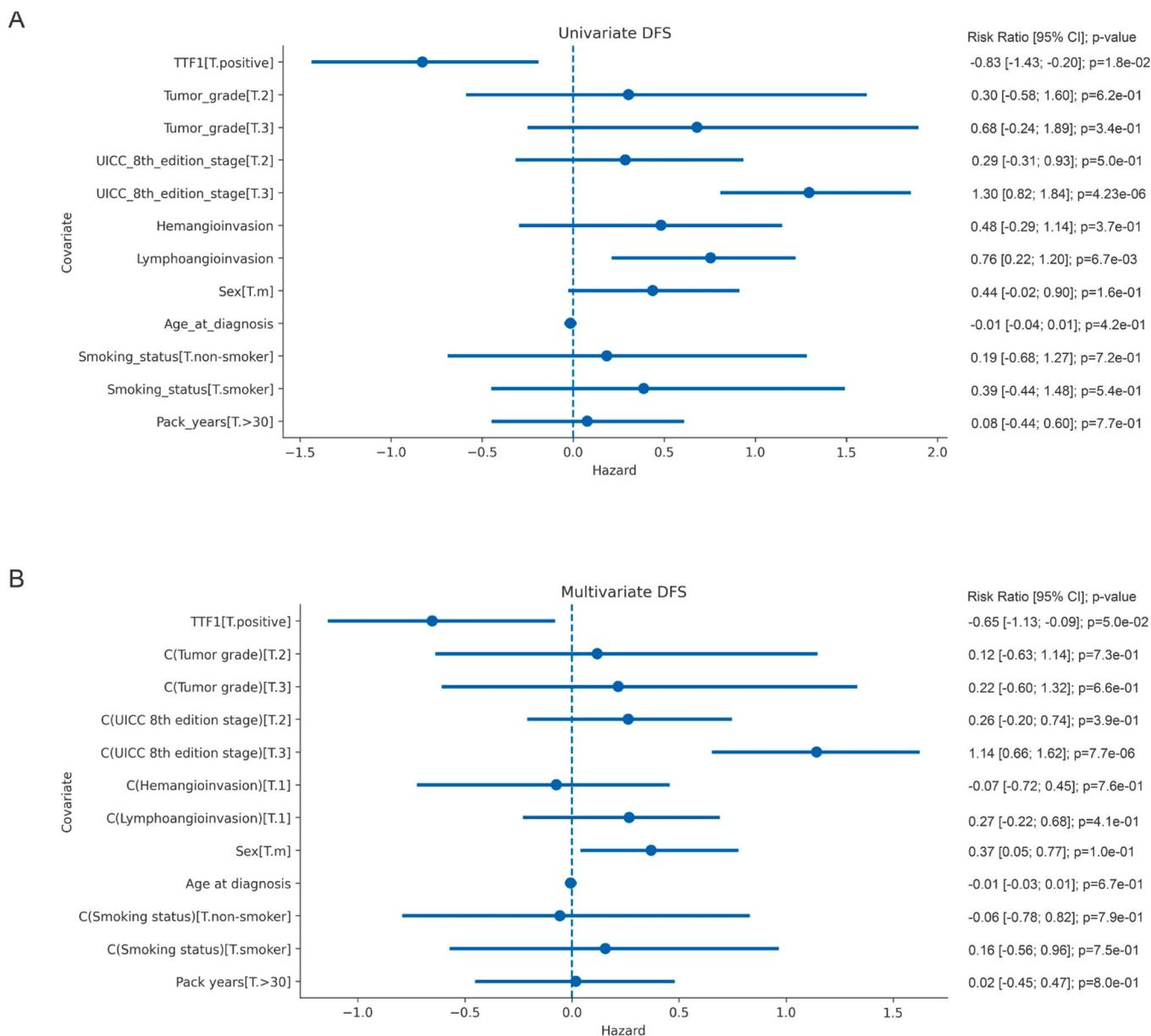
In this retrospective cohort analysis of 447 patients with fully resected, early-stage LUAD (stage I-III), we investigated the prognostic value of TTF-1 expression and found it to be independent of grading. Moreover, our study determined the prognostic contributions of TTF-1 and grading and could show that TTF-1 has a larger total effect on patient outcome than grade.

In LUAD, the predominant histological growth pattern has a clear prognostic impact on OS and DFS, independent of age, sex, tumor stage, and treatment [4,21–23], which has led to the development of a three-tiered histological grading system for LUAD [18]. The latter aligns with our results that tumor grade predicted the DFS in the whole cohort (Fig. S2Y). However, our analyses on a large bi-centric cohort have shown that grading only has predictive value for TTF-1 positive patients, whereas no effect was observed in TTF-1 negative samples. Moreover, the validity of the tumor grading is limited after neoadjuvant treatment due to potential changes in the growth pattern that could lead to misclassification.

TTF-1 status has been reported to correlate with the Karnofsky Performance Status, sex, smoking status, as well as distinct molecular alterations like sensitizing EGFR mutations in previous studies [11,12]. In contrast, our cohort did not show any statistically significant association between TTF-1 and these parameters (Fig. S2V). Furthermore, several studies described TTF-1 as a prognostic indicator for patients' survival in early and late-stage LUAD [11,14,15,17,24]. This is consistent with our results, where a positive TTF-1 status is correlated with a significantly longer DFS in univariate analysis. In addition to previous studies, we identified TTF-1 as a prognostic indicator of patient outcome



**Fig. 1.** Association of TTF-1-status and tumor grade with DFS: (A) The histogram shows the share of TTF-1 status per tumor grading. (B-D) DFS Kaplan-Meier curves showing a comparison of TTF-1 status for the different tumor grades and (E, F) vice versa. The numbers in the upper right corner indicate the median survival time in months if an event was observed in more than 50% of cases (otherwise = inf). SubFig. F shows that further differentiation of patients by tumor grade has no effect once it is known that a patient is TTF-1 negative. The reverse is not true (subfigure B-D), where TTF-1 status remains informative even if we initially stratify by tumor grade.



**Fig. 2.** Effect of TTF-1, tumor grade, and clinicopathologic parameters on DFS by univariate and multivariate analyses: (A) Univariate analysis and (B) multivariate analysis of TTF-1 status (switch from negative to positive), tumor grade (switch from grade 1 to grade 2 and grade 2 to grade 3 respectively), sex (switch from female to male), UICC stage (switch from UICC stage 1 to UICC stage 2 and UICC stage 2 to UICC stage 3 respectively), smoking status, pack-years, lymphangioinvasion (switch from L0 to L1), hemangioinvasion (switch from V0 to V1) and age at diagnosis regarding DFS. Hazard ratios are log scaled for better readability.

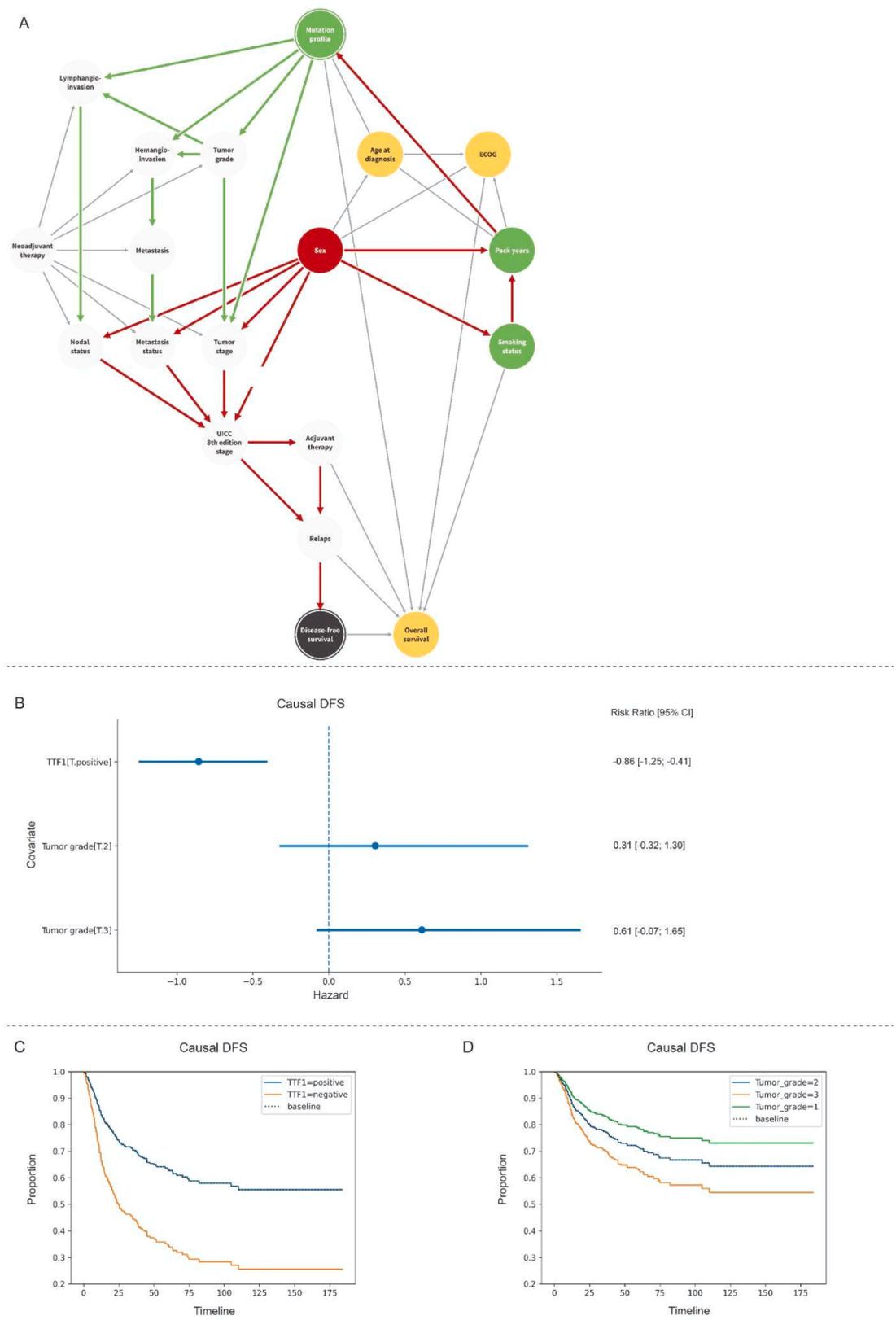
independent of grading in multivariate analysis.

Finally, to measure the effect of TTF-1 in the context of known prognostic parameters, particularly tumor grading, we performed a causally informed effect estimation, for the first time, showing that the impact of varying TTF-1 is superior to that of varying grading. Performing the Bayesian network analysis critically controls for possible confounding factors, which otherwise often remains under-discussed.

It may help to compare the network analysis approach to the popular practice of multivariate analysis in which an analyst would first score each variable individually for its statistical link with the outcome and then use the variables that met a significance threshold for the multivariate model. There are two problems with this approach that Bayesian network analysis helps us address. First, the prefiltering step could mistakenly exclude some variables that could be relevant in combination and would not be included in the multivariate model. Second, the multivariate model may include numerous symptomatic variables that

together overrule the effect of the actual effector variable that caused them. In theory, it is possible to hide the association of a causal factor with the outcome by providing countless related symptomatic variables of that causal factor. For example, if the correlation of an aggressive molecular NSCLC genotype (causal factor), as well as clinical tumor symptoms (related symptomatic variables of an aggressive molecular NSCLC genotype) with the survival time of the patients (outcome), were examined simultaneously, the clinical tumor symptoms would obscure or reduce the true effect size of the aggressive molecular NSCLC genotype in a multivariate analysis.

With Bayesian network analysis, one can establish the cause-and-effect relationships among variables a priori and employ the resulting network to choose the necessary set of variables in the regression. If implemented accurately, this method prevents obscure and misleading regression outcomes but introduces dependence on the analyst's judgment of the causal connections between the variables. The predictive



**Fig. 3.** Evaluation of the true effect size of TTF-1 on DFS and comparison with grading by causal effect estimation: (A). Bayesian network modeling the conditional dependency among variables in a directed acyclic graph. The mutational profile, including TTF-1 status, is defined as exposure (input) and DFS as outcome, respectively. Green circles represent ancestors of exposure, light gray circles ancestors of outcome, red circles ancestors of exposure and outcome, yellow circles represent other variables, green arrows causal pathways, and red arrows biasing pathways respectively. (B) Forrest plot showing the computed effects of TTF-1 status and tumor grade on DFS. Computed survival curves showing the effects of (C) TTF-1 and (D) tumor grading.

power of a Bayesian network highly depends on the accuracy of its design. Therefore, in the appendix, we have included alternative instances of the Bayesian network that we consider having an equal likelihood of being true. We find that discrepancies between these networks have no impact on variable selection in our analysis and presented results, so the equally likely networks show the same significant effects. We believe that this is an important step for standard medical analysis efforts as it facilitates constructive discussions around variable selection. While dependent on prior knowledge, the proposed Bayesian network allows others to adjust and refine estimations by incorporating their domain knowledge and experimental findings. Ultimately, a Bayesian network analysis is a modeling tool that allows for incremental improvement.

## 5. Limitations

This retrospective study has some limitations. First, the relatively small number of TTF-1 negative cases might have mimicked the effects of grading in this population. Second, the setup of the Bayesian network followed the investigators' discretion, thus being prone to selection bias and resulting in potentially missing confounders.

## 6. Conclusion

This study underscores the independent prognostic value of TTF-1 expression for tumor relapse and survival in LUAD regardless of tumor grade. Moreover, our analysis indicates that the prognostic power of tumor grading is limited to TTF-1-positive patients. Furthermore, our findings reveal that the effect size of TTF-1 surpasses that of tumor grading. Thus, TTF-1 may serve as a rapidly evaluable, practical, cost-effective discriminator for patient prognosis in LUAD, which, in contrast to the grading, is also usable after neoadjuvant therapy. To translate the findings into the pathology report, we recommend distinguishing between TTF-1-positive and TTF-1-negative LUADs in tumor coding. Tumor grading should only be applied to TTF-1-positive LUADs (TTF-1 + / G1–3). TTF-1-negative LUADs should either not be graded or always be classified as high grade (TTF-1- / G3).

## Grant numbers and/or funding information

The work of Simon Schallenberg was supported by the Federal Ministry of Education and Research (BMBF) as part of the National Research Initiatives for Mass Spectrometry in Systems Medicine, under grant agreement number 16LW0239K (Charité – University Medicine Berlin) of funding phase B, i.e. MSTAR-2.

## Disclaimer

The project the report is based on is supported by the Federal Ministry of Education and Research (BMBF) under grant agreement number 16LW0239K (Charité – University Medicine Berlin) of funding phase B, i.e., MSTAR-2. The author (Simon Schallenberg) is responsible for the content of the publication.

## CRediT authorship contribution statement

**Nikolaj Frost, Frederick Klauschen, Simon Schallenberg, Gabriel Dernbach:** Conceptualization. **Simon Schallenberg, Gabriel Dernbach, Georg Schlachtenberger, Philipp Anders, Kyrill Boschung, Sabine Merkelbach-Bruse, Ulrich Keilholz:** Data curation. **Simon Schallenberg, Gabriel Dernbach:** Formal analysis. **Simon Schallenberg:** Funding acquisition. **Nikolaj Frost, Frederick Klauschen, Simon Schallenberg, Gabriel Dernbach:** Investigation. **Simon Schallenberg, Gabriel Dernbach:** Methodology. **David Horst, Reinhard Büttner, Frederick Klauschen:** Resources. **Nikolaj Frost, Frederick Klauschen:** Supervision. **Simon Schallenberg, Gabriel**

**Dernbach:** Visualization. **Simon Schallenberg, Gabriel Dernbach, Nikolaj Frost, Frederick Klauschen, Mihnea P. Dragomir:** Roles/ Writing – original draft. **Nikolaj Frost, Frederick Klauschen, Simon Schallenberg, Gabriel Dernbach, Mihnea P. Dragomir, Lukas Ruff, Corinna Friedrich, Jens Khosro Hekmat, Ulrich Keilholz, Georg Schlachtenberger, Matthias Heldwein, Reinhard Büttner, Sabine Merkelbach-Bruse, Alexander Quaas, Kyrill Boschung, Christian Grohé, Carsten Rückert, Winfried Randerath, David Horst, Philipp Anders, Kai Standvoss:** Writing – review & editing.

## Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at [doi:10.1016/j.ejca.2023.113474](https://doi.org/10.1016/j.ejca.2023.113474).

## References

- [1] Meza R, et al. Lung cancer incidence trends by gender, race and histology in the United States, 1973-2010. *PLoS One* 2015;10(3):e0121323.
- [2] Lortet-Tieulent J, et al. International trends in lung cancer incidence by histological subtype: adenocarcinoma stabilizing in men but still increasing in women. *Lung Cancer* 2014;84(1):13–22.
- [3] Goldstraw P, et al. The IASLC Lung Cancer Staging Project: Proposals for Revision of the TNM Stage Groupings in the Forthcoming (Eighth) Edition of the TNM Classification for Lung Cancer. *J Thorac Oncol* 2016;11(1):39–51.
- [4] Russell PA, et al. Does lung adenocarcinoma subtype predict patient survival?: A clinicopathologic study based on the new International Association for the Study of Lung Cancer/American Thoracic Society/European Respiratory Society international multidisciplinary lung adenocarcinoma classification. *J Thorac Oncol* 2011;6(9):1496–504.
- [5] Tsao MS, et al. Subtype classification of lung adenocarcinoma predicts benefit from adjuvant chemotherapy in patients undergoing complete resection. *J Clin Oncol* 2015;33(30):3439–46.
- [6] Hoang T, et al. Clinical model to predict survival in chemo-naïve patients with advanced non-small-cell lung cancer treated with third-generation chemotherapy regimens based on eastern cooperative oncology group data. *J Clin Oncol* 2005;23(1):175–83.
- [7] Wheatley-Price P, et al. The influence of sex and histology on outcomes in non-small-cell lung cancer: a pooled analysis of five randomized trials. *Ann Oncol* 2010; 21(10):2023–8.
- [8] Kimura S, et al. The T/eBP null mouse: thyroid-specific enhancer-binding protein is essential for the organogenesis of the thyroid, lung, ventral forebrain, and pituitary. *Genes Dev* 1996;10(1):60–9.
- [9] Lau SK, et al. Thyroid transcription factor-1: a review. *Appl Immunohistochem Mol Morphol* 2002;10(2):97–102.
- [10] Stenhouse G, et al. Thyroid transcription factor 1 in pulmonary adenocarcinoma. *J Clin Pathol* 2004;57(4):383–7.
- [11] Schilsky JB, et al. Prognostic impact of TTF-1 expression in patients with stage IV lung adenocarcinomas. *Lung Cancer* 2017;108:205–11.
- [12] Shanzhi W, et al. The relationship between TTF-1 expression and EGFR mutations in lung adenocarcinomas. *PLoS One* 2014;9(4):e95479.
- [13] Frost N, et al. Pemetrexed-Based Chemotherapy Is Inferior to Pemetrexed-Free Regimens in Thyroid Transcription Factor 1 (TTF-1)-Negative, EGFR/ALK-Negative Lung Adenocarcinoma: A Propensity Score Matched Pairs Analysis. *Clin Lung Cancer* 2020;21(6):e607–21.
- [14] Anagnostou VK, et al. Thyroid transcription factor 1 is an independent prognostic factor for patients with stage I lung adenocarcinoma. *J Clin Oncol* 2009;27(2): 271–8.
- [15] Park JY, et al. Thyroid transcription factor-1 as a prognostic indicator for stage IV lung adenocarcinoma with and without EGFR-sensitizing mutations. *BMC Cancer* 2019;19(1):574.
- [16] Kim JH, et al. Prognostic Impact of TTF-1 Expression in Non-Squamous Non-Small-Cell Lung Cancer: A Meta-Analysis. *J Cancer* 2018;9(22):4279–86.
- [17] Qian HH, et al. Prognostic value of TTF-1 expression in patients with non-small cell lung cancer: A meta-analysis. *Clin Chim Acta* 2015;451(Pt B):208–14.
- [18] Moreira AL, et al. A grading system for invasive pulmonary adenocarcinoma: a proposal from the international association for the study of lung cancer pathology committee. *J Thorac Oncol* 2020;15(10):1599–610.
- [19] Travis WD, et al. International association for the study of lung cancer/american thoracic society/european respiratory society international multidisciplinary classification of lung adenocarcinoma. *J Thorac Oncol* 2011;6(2):244–85.
- [20] Textor J, et al. Robust causal inference using directed acyclic graphs: the R package "dagitty". *Int J Epidemiol* 2016;45(6):1887–94.

- [21] Warth A, et al. The novel histologic International Association for the Study of Lung Cancer/American Thoracic Society/European Respiratory Society classification system of lung adenocarcinoma is a stage-independent predictor of survival. *J Clin Oncol* 2012;30(13):1438–46.
- [22] Sica G, et al. A grading system of lung adenocarcinomas based on histologic pattern is predictive of disease recurrence in stage I tumors. *Am J Surg Pathol* 2010;34(8): 1155–62.
- [23] Yoshizawa A, et al. Impact of proposed IASLC/ATS/ERS classification of lung adenocarcinoma: prognostic subgroups and implications for further revision of staging based on analysis of 514 stage I cases. *Mod Pathol* 2011;24(5):653–64.
- [24] Saad RS, et al. Prognostic significance of thyroid transcription factor-1 expression in both early-stage conventional adenocarcinoma and bronchioloalveolar carcinoma of the lung. *Hum Pathol* 2004;35(1):3–7.