

Keywords: lung cancer; proliferation; Ki-67; prognosis

# Tumour cell proliferation (Ki-67) in non-small cell lung cancer: a critical reappraisal of its prognostic role

A Warth<sup>\*1</sup>, J Cortis<sup>1</sup>, A Soltermann<sup>2</sup>, M Meister<sup>3,8</sup>, J Budczies<sup>4</sup>, A Stenzinger<sup>1</sup>, B Goeppert<sup>1</sup>, M Thomas<sup>5,8</sup>, F J F Herth<sup>6,8</sup>, P Schirmacher<sup>1</sup>, P A Schnabel<sup>1,8</sup>, H Hoffmann<sup>7</sup>, H Dienemann<sup>7,8</sup>, T Muley<sup>3,8,9</sup> and W Weichert<sup>\*1,9</sup>

<sup>1</sup>Institute of Pathology, University Hospital Heidelberg, Heidelberg, Germany; <sup>2</sup>Institute of Pathology, University Hospital Zürich, Zürich, Switzerland; <sup>3</sup>Translational Research Unit, Thoraxklinik at University Hospital Heidelberg, Heidelberg, Germany; <sup>4</sup>Institute of Pathology, Charité Berlin, Berlin, Germany; <sup>5</sup>Department of Thoracic Oncology, Thoraxklinik at University Hospital Heidelberg, Heidelberg, Germany; <sup>6</sup>Department of Pneumology and Critical Care Medicine, Thoraxklinik at University Hospital Heidelberg, Heidelberg, Germany and <sup>7</sup>Department of Thoracic Surgery, Thoraxklinik at University Hospital Heidelberg, Heidelberg, Germany

**Background:** Uncontrolled proliferation is a hallmark of malignant tumour growth. Its prognostic role in non-small cell lung cancer (NSCLC) has been investigated in numerous studies with controversial results. We aimed to resolve these controversies by assessing the Ki-67 proliferation index (PI) in three large, independent NSCLC cohorts.

**Methods:** Proliferation index was retrospectively analysed by immunohistochemistry in a cohort of 1065 NSCLC and correlated with clinicopathological data including outcome and therapy. Results were validated in two independent cohorts of 233 squamous cell carcinomas (SQCC) and 184 adenocarcinomas (ADC).

**Results:** Proliferation index (overall mean: 40.7%) differed significantly according to histologic subtypes with SQCC showing a mean PI (52.8%) twice as high as ADC (25.8%). In ADC PI was tightly linked to growth patterns. In SQCC and ADC opposing effects of PI on overall (OS), disease-specific and disease-free survival were evident, in ADC high PI (optimised validated cut-off: 25%) was a stage-independent negative prognosticator (hazard ratio, HR OS: 1.56,  $P=0.004$ ). This prognostic effect was largely attenuated by adjuvant radio-/chemotherapy. In SQCC high PI (optimised validated cut-off: 50%) was associated with better survival (HR OS: 0.65,  $P=0.007$ ).

**Conclusions:** Our data demonstrate that PI is a clinically meaningful biomarker in NSCLC with entity-dependent cut-off values that allow reliable estimation of prognosis and may potentially stratify ADC patients for the need of adjuvant therapy.

Lung cancer accounts for almost one-fourth of total cancer deaths (Jemal *et al*, 2011). Approximately, 80% of all lung tumours belong to non-small cell lung cancer (NSCLC). Since most NSCLC are not resectable at diagnosis, chemotherapy remains the mainstay of treatment. Despite recent progress with individualised treatment, prognosis is still poor. To maximise the benefit of therapy,

prognostic assessment according to the tumour, node, metastasis system, histomorphological characteristics (Warth *et al*, 2012b), and predictive biomarker analyses, that is, for epidermal growth factor receptor (*EGFR*) mutations (Mok *et al*, 2009), and anaplastic lymphoma kinase (*ALK*) translocations (Kwak *et al*, 2010) are of utmost importance for patient stratification.

\*Correspondence: Dr A Warth; E-mail: arne.warth@med.uni-heidelberg.de or Professor Dr W Weichert; E-mail: wilko.weichert@med.uni-heidelberg.de

<sup>8</sup>Translational Lung Research Center Heidelberg (TLRC-H), Member of the German Center for Lung Research (DZL).

<sup>9</sup>These authors contributed equally to this work.

Received 31 March 2014; revised 16 June 2014; accepted 24 June 2014; published online 22 July 2014

© 2014 Cancer Research UK. All rights reserved 0007–0920/14

Assessment of tumour proliferation, as a morphology-based measure of tumour growth kinetics, has a long-standing history. To this end, analysis of proliferation-associated antigens, especially Ki-67, is frequently performed. Ki-67 is a DNA-binding nuclear protein expressed throughout the cell cycle in proliferating but not in quiescent (G0) cells. It is a well-known potent biomarker with significant prognostic and/or predictive value in leading cancer entities such as breast (Urruticoechea *et al*, 2005; Viale *et al*, 2008; Cuzick *et al*, 2011; Luporsi *et al*, 2011), prostate (Pollack *et al*, 2004; Khor *et al*, 2009) and colorectal cancer (Allegra *et al*, 2003). Consequently, the proliferation index (PI) influences clinical decision-making and choice of treatment in several tumour entities. This includes breast cancer where PI is used to separate luminal A and luminal B phenotypes (Leong and Zhuang, 2011), which impacts on therapy, and neuroendocrine tumours of the gastrointestinal tract where assessment of PI is the key component of tumour grading with strong prognostic as well as therapeutic implications (Rindi *et al*, 2012). First data also indicate a diagnostic and prognostic value of PI in pulmonary carcinoids (Zahel *et al*, 2012; Warth *et al*, 2013); the value of Ki-67 in neuroendocrine tumours of the lung will be further substantiated in the future (Pelosi *et al*, 2014). Beyond its use in formal classification systems, tumour cell proliferation is frequently analysed by pathologists to comprehensively describe the aggressiveness and the biological behaviour of a tumour which may provide additional 'non-formalised' clues for therapy selection.

In resected NSCLC, meta-analysis of numerous studies suggested that high PI is a negative prognosticator (Martin *et al*, 2004) and predicts recurrence (Poleri *et al*, 2003; Yamashita *et al*, 2011) as well as survival after lung tumour ablation (Sofocleous *et al*, 2013). However, the informative value of the vast majority of studies on this issue is limited due to the usually small sample numbers investigated. Hence, published meta-analyses are hampered by a lack of uniformity with respect to clinicopathological data and chosen cut-off values. Thus, Ki-67 assessment in NSCLC has not been successfully translated into routine diagnostics yet.

In order to clarify some of these issues and to elucidate the prognostic impact of PI in NSCLC we therefore retrospectively analysed three independent NSCLC cohorts including a total of almost 1500 cases in which we correlated PI with clinicopathological data including overall (OS), disease-specific (DSS), and disease-free survival (DFS).

## PATIENTS AND METHODS

**Patients.** For the test cohort only R0-resected NSCLC with clinicopathological data including sex, age, tumour size, lymph node and distant metastases, type of resection, histotype, applied adjuvant therapy, OS, DFS, and DSS were chosen. All tumours were resected between 2002 and 2008 at the Thoraxklinik Heidelberg, Germany (local ethics committee no. 206/2005). Tumours resected after neoadjuvant therapy were excluded. The ADC (adenocarcinoma) validation cohort consisted of tumours with the same dataset resected between 2009 and 2010 at the Thoraxklinik Heidelberg. The SQCC (squamous cell carcinoma) validation cohort consisted of tumours resected R0 between 1993 and 2002 at the University Hospital Zurich, Switzerland. Basic clinical (age, sex and stage) data, OS and DFS status as well as postoperative treatment data were available in this dataset, as well. Tumours were classified by AW, WW, and PAS (Heidelberg) or AS (Zurich) according to the criteria of the 2004 WHO classification and re-staged according to the seventh Edition International Union Against Cancer/American Joint Committee on Cancer Tumour, Node, and Metastasis (TNM) classification. Adenocarcinoma subtyping was performed according to the

International Association for the Study of Lung Cancer (IASLC)/American Thoracic Society (ATS)/European Respiratory Society (ERS) proposal (Travis *et al*, 2011).

**Clinicopathological characteristics.** The test cohort included 1065 NSCLC specimens, 482 ADC (45.3%), 437 SQCC (41%), 42 adenosquamous carcinomas (ASC; 3.9%), 57 large cell carcinomas (LC; 5.4%), 11 large cell neuroendocrine carcinomas (LCNEC; 1%), 28 sarcomatoid carcinomas (SC; 2.6%), and 8 combined tumours (0.8%). Mean age of patients was 63.2 years, 773 patients were male (72.6%), 292 patients were female (27.4%). A total of 17 patients were treated by wedge resection (1.6%), 8 patients received segmentectomy (0.8%), 779 (73.1%) lobectomy, 37 (3.5%) bilobectomy, and 224 (21%) pneumonectomy, accompanied by systematic lymph node dissection. Adjuvant chemotherapy or mediastinal irradiation was administered according to the guidelines in effect at the time of resection. Mean follow-up of patients alive at the endpoint of analysis was 45.6 (ADC: 43.5; SQCC: 49.3) months. Mean follow-up of patients alive at the endpoint of the analysis or who died of other causes was 42.1 (ADC: 40.5; SQCC: 44.8) months. Mean follow-up of patients alive without recurrence was 40.5 (ADC: 37.8; SQCC: 44.1) months. Clinicopathological characteristics of the test cohort are provided in Table 1.

The ADC validation cohort comprised 184 ADC. Mean age of patients was 63.1 years. Overall, 96 (52.2%) patients were male, 88 (47.8%) were female; 73 (39.7%) tumours were classified as stage I, 43 (23.4%) as stage II, 62 (33.6%) as stage III and 6 (3.3%) as stage IV. Mean follow-up time of patients alive at the endpoint of analysis was 27.9 months. Mean follow-up time of patients who were alive at the endpoint of analysis or died of non-cancer-related causes was 26.7 months. Mean recurrence-free follow-up time was 25.1 months.

The SQCC validation cohort comprised 233 SQCC. Mean age of patients was 65.6 years, 187 (80.3%) were male, 46 (19.7%) were female; 86 patients (36.9%) were classified as stage I, 88 (37.8%) as stage II, 46 (19.7%) as stage III and 13 (5.6%) as stage IV. Mean follow-up of patients alive at the endpoint of analysis was 89.1 months. Mean recurrence-free follow-up was 91.7 months.

**Tissue microarray construction, immunohistochemistry, and PI analysis.** Tissue microarray (TMA) construction for the test cohort and the ADC validation cohort was previously described in detail (Warth *et al*, 2012a). In brief, a TMA machine (AlphaMetrix Biotech, Rödermark, Germany) was used to extract two 1.0-mm cylindrical core samples from each tumour block. Immunohistochemistry (IHC) was performed on both cohorts using an antibody against Ki-67 (MIB1, 1:400; DAKO, Hamburg, Germany) according to a quality controlled and accredited (DAkKS) protocol. Tissue microarray slides were deparaffinised, pre-treated with an antigen retrieval buffer (pH 9.0; DAKO) and stained using an automated device (DAKO Techmate 500plus). For the SQCC validation cohort, TMA construction was performed with a custom-made, semiautomatic tissue arrayer (Beecher Instruments, Sun Prairie, WI, USA) extracting two 0.6-mm core samples from each donor block. A rabbit monoclonal anti-Ki-67 antibody clone 30-9 (Confirm kit, prediluted, Ventana-Roche, Basel, Switzerland) was used on a Benchmark automated IHC platform (Ventana-Roche) using protocol CC1m and ultraView-HRP detection.

Ki-67-positive tumour cells were counted per 100 tumour cells separately in every core by focusing on representative areas avoiding obvious hot spots as well as areas with specifically low indices and areas with necrosis or haemorrhage. The mean of the results of all cores, which could be evaluated for a given case, was assigned as final Ki-67 index to this case. In a certain number of cases, only one core was available for evaluation due to inadequate sampling of tumour tissue in one of the cores or loss of the second core during TMA processing. In the test cohort, for example, in 913 (ADC: 404; SQCC: 386; others: 123) cases both cores could be

**Table 1.** Association of proliferative activity with staging parameters for the whole test cohort of non-small cell lung cancers (NSCLC), the test cohort adenocarcinoma subgroup (ADC), and the test cohort squamous cell carcinoma subgroup (SQCC)

Characteristics	All NSCLC	Ki-67 mean	Ki-67 s.d.	P-value	ADC	Ki-67 mean	Ki-67 s.d.	P-value	SQCC	Ki-67 mean	Ki-67 s.d.	P-value
<b>All cases</b>												
	1065	40.7	27.5		482	25.8	24.6		437	52.8	24.3	
<b>UICC stage</b>												
I	389	36.6	28.7	0.001	198	21.5	22.2	<0.001	161	52.6	25.4	0.888
II	337	44.5	25.7		103	34.1	25.3		172	52.2	23.4	
III	311	41.4	27.1		165	31.2	25.0		98	53.7	24.2	
IV	28	42.9	28.1		16	34.7	27.3		6	59.0	19.9	
<b>Tumour stage</b>												
pT1	198	39.5	29.4	0.010	92	23.0	23.0	0.123	84	53.9	26.2	0.374
pT2	672	40.0	26.9		306	28.9	24.8		278	51.6	24.1	
pT3	174	46.2	26.7		70	31.1	25.4		69	56.8	21.6	
pT4	21	29.5	26.1		14	23.6	22.2		6	46.7	29.9	
<b>Nodal status<sup>a</sup></b>												
pN0	545	38.4	28.1	0.008	258	24.5	23.4	0.005	224	51.9	25.7	0.807
pN1	261	45.5	25.4		83	34.5	25.4		136	53.4	21.8	
pN2	247	40.0	27.5		138	30.5	25.4		71	54.1	24.6	
pN3	5	38.0	29.4		3	22.5	16.4		1	37.5		
<b>Distant metastasis</b>												
M0	1038	40.6	27.5	0.814	467	27.8	24.5	0.484	431	52.7	24.3	0.527
M1	27	41.9	28.1		15	32.3	26.5		6	59.0	19.9	

Abbreviations: s.d. = standard deviation; UICC = Union Internationale Contre le Cancer.

<sup>a</sup>For seven cases data of nodal status were not available (limited resection).

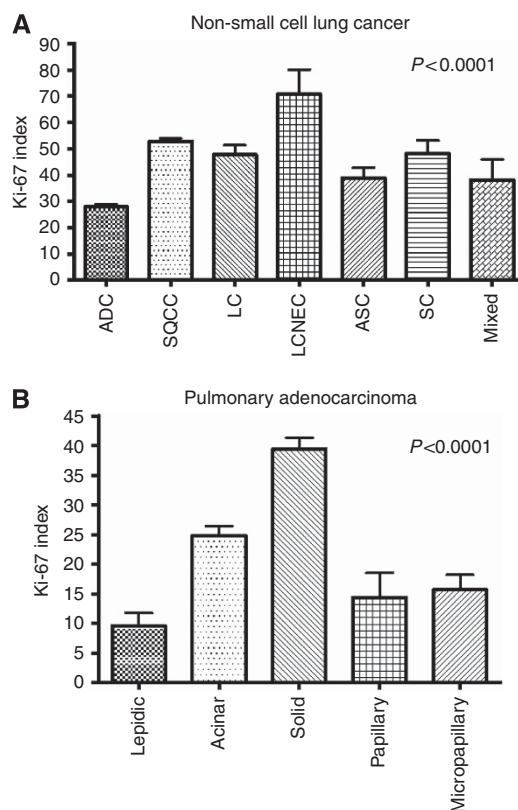
evaluated while in 152 (ADC: 78; SQCC: 51; others: 23) cases (14.3%) only one core was available for Ki-67 evaluation.

**Statistics.** Correlation of PI with clinicopathological characteristics was done by unpaired *t*-test and analysis of variance. Comparison of tumour diameters with PI was tested with Pearson's correlation coefficient. Overall survival, DSS, and DFS were analysed for all possible cut-offs using the Kaplan–Meier method, with a log-rank test to probe for significance. The testing for continuous hazard ratios (HRs) was done with a web-based software tool (cut-off finder) based on Cox regression (Budczies *et al*, 2012). Multivariate survival analysis was done by the Cox proportional hazard model. All tests were done two sided and performed using SPSS Statistics 19 (IBM, Ehningen, Germany), Prism 4.03 (GraphPad, La Jolla, CA, USA), and R (www.r-project.org). *P*-values <0.05 were considered significant.

## RESULTS

**NSCLC proliferative activity depends on the histotype.** Mean PI in the test cohort was 40.7%. Large cell neuroendocrine carcinomas showed the highest PI (mean: 70.7%) followed by SQCC (52.8%), SC (48.1%), LC (47.7%), mixed tumours (38.1%), and ADC (25.8%; Supplementary Figure 1, Figure 1). Within ADC subtypes solid predominant ADC showed the highest PI (39.4%), followed by acinar (21.6%), micropapillary (15.7%), papillary (14.4%), and lepidic predominant ADC (9.5%; Figure 1).

**Proliferative activity of NSCLC is associated with tumour stage and nodal status.** In the test cohort PI was associated with tumour size ( $r = 0.121$ ,  $P < 0.001$ ; Supplementary Figure 2), due to



**Figure 1.** Histology and proliferative activity. (A) Proliferative activity in dependence of histological tumour type. (B) Association of pulmonary adenocarcinoma growth patterns with proliferation.

an association of PI and size in ADC ( $r=0.118$ ). This association was not observed for SQCC. UICC stage, tumour stage, and nodal status were also associated with PI (Table 1). This could be demonstrated for the whole NSCLC cohort and for the ADC subgroup, but not for SQCC. The most striking differences of proliferation with respect to UICC stage were evident between stages I and II for ADC with a >10% higher PI in stage II. Adenocarcinomas with lymph node metastasis had higher PI compared to node-negative tumours (Table 1).

**Proliferative activity of NSCLC and survival: opposing effects of low and high Ki-67 indices.** In most published studies mean or median PI has been used for stratification of cases according to proliferative activity. Applying these categories to our test cohort did not yield significant differences with respect to OS, DSS, and DFS ( $P \geq 0.1$  for all comparisons; not shown). To investigate this further, we continuously plotted HRs for all possible cut-offs for OS, DSS, and DFS (Figure 2A, Supplementary Figure 3) using an automated software tool (Budczies *et al*, 2012). Cut-offs in the low PI range resulted in HR higher than (OS and DSS) or just ~1 (DFS), indicating a survival disadvantage, while cut-offs in the higher PI range showed HRs < 1, indicating a better prognosis for OS, DSS, and DFS. This is exemplified by Kaplan–Meier curves for the cut-offs of 10% and 55% as shown in Supplementary Figure 4. To dissect these surprising findings further, we then investigated ADC and SQCC separately.

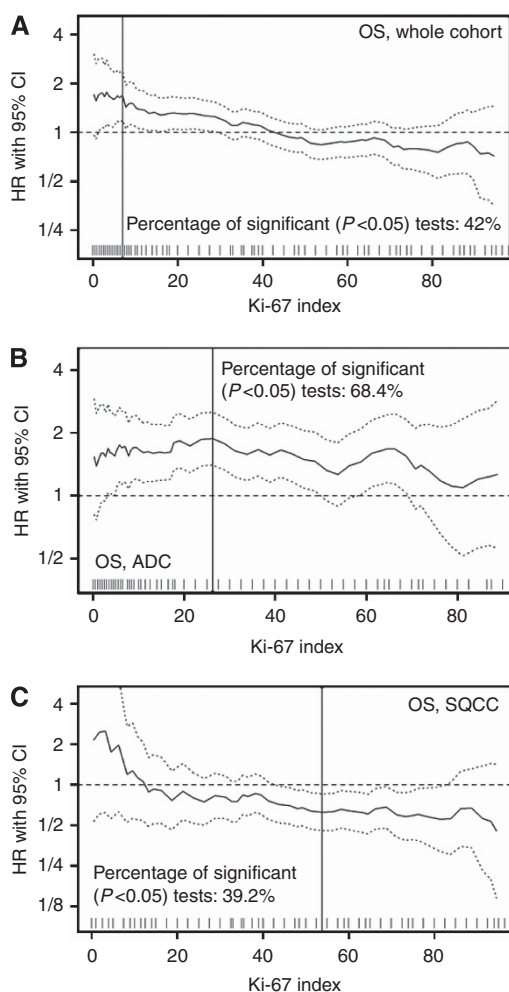


Figure 2. Association of proliferative activity with survival in non-small cell lung cancer. All possible cut-off values for proliferative activity and their impact on overall survival (OS) are depicted for the whole test cohort (A), and in the subgroups of pulmonary ADC (B) and SQCC (C).

**Strong proliferative activity in ADC predicts poor survival: definition of cut-off values and validation.** In ADC, high PI was significantly associated with a worse prognosis almost over the whole range of potential cut-off values for OS, DSS, and DFS (Figure 2B, Supplementary Figure 5). The statistically optimal cut-off for the separation of a good and a poor prognostic ADC group across all survival parameters was at 25%. Application of this cut-off led to a mean DFS of 60.3 months (OS: 71.3 months; DSS: 80.3 months) for patients with a PI < 25% vs 54.3 months (OS: 56.4 months; DSS: 65.2 months) for patients with a PI  $\geq$  25% (Figure 3) in the test cohort.

Stage and growth pattern of ADC were also predictors of OS, DSS, and DFS in univariate analyses (not shown). In multivariate survival analysis under inclusion of stage, pattern and PI, PI was an independent predictor of OS (HR 1.56,  $P=0.004$ ; Supplementary Table 1) and DSS (HR = 1.80,  $P=0.002$ ) but not DFS (HR = 1.11,  $P=0.442$ ) in the ADC test cohort.

Analysis of the independent ADC validation cohort confirmed the impact of the predefined cut-off of 25% on survival in ADC. Survival of patients with high PI ( $\geq 25\%$ ) was considerably shorter, which was statistically significant for DFS but not for OS and DSS (Figure 3). Mean DFS for patients with low PI (< 25%) was 35.4 months (OS: 39.3 months; DSS: 41.2 months), while DFS for patients with a PI  $\geq 25\%$  was only 28.8 months (OS: 37.5 months; DSS: 40.0 months). This survival difference was borderline significant in a multivariate Cox proportional hazard model for DFS (HR: 1.63,  $P=0.055$ ) under inclusion of the above mentioned parameters.

**The prognostic impact of proliferation in ADC patients is abolished by adjuvant treatment.** We addressed if adjuvant treatment regimens alter the prognostic impact of PI in ADC. The rate of potential cut-offs that produced a HR significantly > 1 for DSS decreased from 68.5% in ADC patients without adjuvant chemotherapy to 13.9% in patients with adjuvant chemotherapy (Supplementary Figure 6). This effect is exemplified for the cut-off at 25% (Supplementary Figure 7). A comparable drop in the frequency of significant results (48.1% to 6.7%) was also evident for adjuvant radiotherapy (Supplementary Figure 6). The same was true for OS and DFS (Supplementary Table 2). To test whether adjuvant therapy interfered with the overall independent prognostic impact of PI in ADC in general, we included this information into an additional multivariate analysis. In the respective multivariate analysis under inclusion of PI, stage, pattern as well as adjuvant chemotherapy and adjuvant radiotherapy, adjuvant chemotherapy (HR: 0.63, CI: 0.44–0.88;  $P=0.008$ ) but not adjuvant radiotherapy (HR: 0.88, CI: 0.59–1.31;  $P=0.536$ ) proved to be an additional independent factor for better OS in this setting. The independent impact of PI on OS, however, remained unchanged in this setting (HR: 1.57, CI: 1.16–2.12;  $P=0.004$ ).

**High proliferation in SQCC is associated with better survival.** In SQCC of the test cohort only cut-offs in the low range (< 15%) resulted in a HR > 1. The majority of potential cut-offs over the medium to high PI range resulted in a HR < 1 (Figure 2C, Supplementary Figure 8), indicating better survival of SQCC patients with higher PI. The most robust differentiation for survival was found at a PI cut-off of 50%. Patients with a PI < 50% had a mean OS of 63.2 months (DFS: 65.9 months), while patients with a PI of  $\geq 50\%$  had a mean OS of 73.7 months (DFS: 74.4 months; Figure 4). The survival effect was stage independent for OS (HR: 0.65, 95% CI: 0.48–0.89;  $P=0.007$ ; Supplementary Table 3) and DFS (not shown). We validated these data in an independent cohort of 233 SQCC. In our SQCC validation cohort patients with a PI < 50% also had a worse survival (mean OS: 58.7 months; DFS 54.8 months) compared to patients with a PI of  $\geq 50\%$  (mean OS: 74.8 months; mean DFS 67.1 months; Figure 4). Again, the effect of

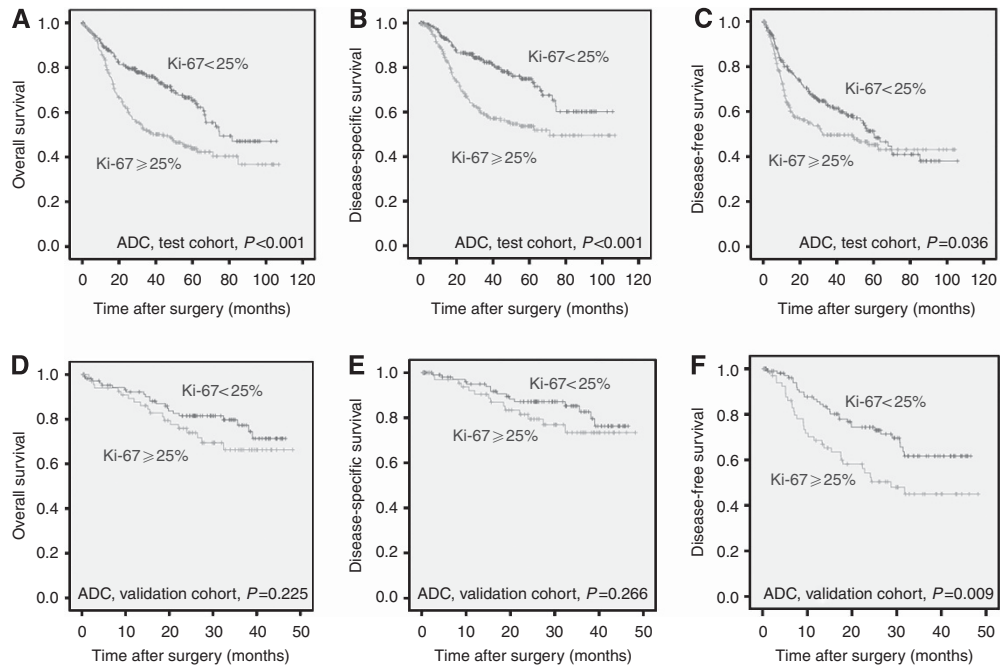


Figure 3. Adenocarcinoma survival probabilities in dependence of proliferative activity. Distribution of survival probabilities for (A and D) overall survival (OS), (B and E) disease-specific survival (DSS), and (C and F) disease-free survival (DFS) in pulmonary adenocarcinomas stratified for proliferative activity (cut-off 25%) in the test cohort (A–C) as well as in the independent validation cohort (D–F).

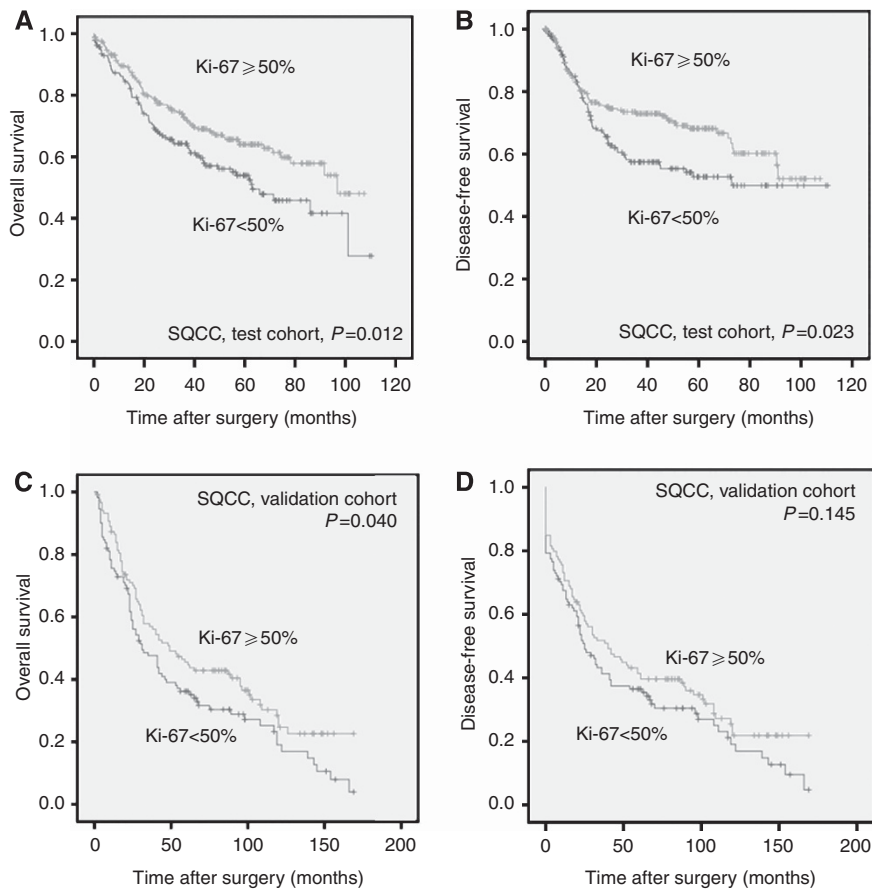


Figure 4. Squamous cell carcinoma survival probabilities in dependence of proliferative activity. Kaplan–Meier curves depicting survival differences for OS (A and C) and DFS (B and D) in dependence of proliferative activity in squamous cell carcinomas (SQCC) with a 50% cut-off in the test cohort (A and B) as well as in the independent validation cohort (C and D).

PI on OS was stage independent in multivariate survival analysis (HR: 0.63, 95% CI 0.46–0.87;  $P=0.005$ ). In contrast to ADC, adjuvant treatment did not significantly impact on the association of PI with survival in SQCC (data not shown).

#### Impact of proliferation on survival in other NSCLC subtypes.

To delineate whether the prognosis of rare NSCLC subgroups is influenced by PI we tested the prognostic impact of our optimised cut-offs for ADC (25%) and SQCC (50%) in LC, LCNEC, SC, and ASC. The 50% cut-off did not result in statistically significant survival differences in any of the groups. In contrast, patients with ASC were prognostically stratified in the same way as ADC when the cut-off of 25% was applied (Supplementary Figure 9; Supplementary Table 2). In the other subgroups no significant survival differences were noted.

## DISCUSSION

In this large-scale study on the prognostic impact of proliferation in NSCLC we demonstrate that PI is a highly significant and independent predictor of survival in ADC and ASC. In accordance with our data, Kadota *et al* (2012) recently showed that mitotic count in combination with architectural grade is an independent recurrence predictor in stage I ADC. Ki-67 has been described as a potent biomarker with significant prognostic value in cancer entities such as breast (Cuzick *et al*, 2011), prostate (Khor *et al*, 2009) and colorectal cancer (Allegra *et al*, 2003). In breast cancer and gastrointestinal neuroendocrine tumours it already entered clinical decision-making. The yet largest meta-analysis on PI in NSCLC indicated a prognostic value as well (Martin *et al*, 2004); however, one recent study investigating PI in 778 NSCLC failed to predict a benefit from adjuvant cisplatin-based chemotherapy (Filipits *et al*, 2007). Optimal and widely applicable Ki-67 cut-offs allowing clinically relevant patient stratification in NSCLC, however, have not been determined yet. Many single studies provided heterogeneous results; this has led to confusion and hampered the use of Ki-67 in NSCLC routine diagnostics. There are several reasons for this heterogeneity. First, many studies on this topic are considerably underpowered with regard to the number of analysed cases. Second, our data demonstrate that the prognostic impact of PI depends on the NSCLC subtype analysed. Hence, analysing a mixed NSCLC cohort will not lead to meaningful results. Third, different and sometimes randomly chosen cut-off values are likely confounders resulting in heterogeneous results (Lee *et al*, 1995). We could demonstrate that the determination of suitable cut-off values for NSCLC subtypes is one of the keys to obtain meaningful results. Evidence-based cut-off values and standardised assessment will be a prerequisite to establish proliferation as a clinically useful parameter in the routine diagnostic setting.

The usage of means/medians is generally accepted as one approach to define biomarker-related cut-off values (Altman *et al*, 1994), this, however, did not turn out to be useful for PI in NSCLC. Continuous evaluation of all possible cut-offs over several survival parameters with the subsequent choice of optimal discriminators and independent validation of these, suggested clinically meaningful PI cut-off values of 25% for ADC and 50% for SQCC. Nevertheless, it is evident that dichotomisation of continuous parameters may pose a problem for diagnostic categorisation in some instances (Royston *et al*, 2006); a sharp PI cut-off of 25% implies that patients with 24% and 26% PI have a significantly different outcome, this potentially leads to different treatment decisions. For clinical use it has therefore been suggested to apply two cut-off values that delineate a central 'grey' area in which other patient-related factors might be considered to select appropriate treatment regimens (Goldhirsch *et al*, 2009). Although we were not

able to define such a grey area in our dataset (not shown), large-scale validation studies in well-characterised NSCLC cohorts are warranted to ultimately determine respective cut-off values and potentially delineate such a 'grey' zone.

The survival differences between ADC with high and low PI were especially prominent in patients who did not receive adjuvant therapy and diminished in patients who received adjuvant treatment, implying that patients with strongly proliferating ADC may benefit to a higher extent from adjuvant treatment compared to patients with slowly proliferating ADC. Thus, PI may be a useful parameter for stratification of patients for adjuvant therapy. This should be tested in the prospective setting.

This study is the first to correlate PI assessed by Ki-67 staining with the different ADC subtypes defined by the IASLC/ATS/ERS (Travis *et al*, 2011). Of note, although micropapillary predominant ADC have the worst prognosis and are significantly associated with the presence of metastasis at the time of diagnosis (Warth *et al*, 2012b), this extremely aggressive biological behaviour is not reflected by a high PI, pointing to other malignancy-associated mechanisms, for example, altered expression of adhesion molecules (Kuner *et al*, 2009) as crucial factors.

In contrast to the situation in ADC we found a complex association of PI with prognosis in SQCC. An increased PI had a negative prognostic impact in SQCC only in the low PI range (<15%), while higher PI cut-offs resulted in an inverse correlation of PI with patient survival. Comparable findings concerning the prognostic impact of Ki-67 for SQCC (and also for ADC) have been reported previously in a few smaller studies (Hommura *et al*, 2000; Takahashi *et al*, 2002; Poleri *et al*, 2003). From a biological viewpoint these results are difficult to explain and presently one could only speculate on potential underlying mechanisms of this observation. One aspect might be that fast uncontrolled tumour growth could lead to an inadequate blood supply with a higher propensity of the tumour to develop necrosis, which in turn might induce a stronger antitumour immune response. However, functional data on such associations are missing yet. Although we do not see an association of adjuvant treatment with survival differences predicted by PI in SQCC, another explanation could be that chemotherapeutic treatment administered later in the disease course might have stronger effects in patients with rapidly proliferating tumours.

Limitations of our study include its retrospective nature as well as the use of TMA cores and not whole slide sections to determine the proliferative activity of tumours. Although the use of TMA may result in a bias, PI in NSCLC biopsies (which are comparable to TMA cores with respect to size) corresponds to PI in the respective resection specimens in >80% of cases (Meert *et al*, 2004). Another limitation in our study might be that staining protocols and antibodies used in the SQCC test and validation subgroup were not identical. However, this also indicates that the determination of PI as a biomarker is not strictly assay dependent. Nevertheless, international standardisation of analysis and assessment of any potential biomarker is an important aspect for a successful translation into the routine setting.

Although the findings concerning the prognostic impact of PI in our test cohort could largely be recapitulated in our ADC validation cohort with respect to DFS, we could not ultimately confirm the prognostic effect of PI for OS and DSS. This is most likely due to the fact that the validation cohort was smaller and follow-up was less mature when compared to the test cohort. This again points out that large and well-characterised cohorts are essential for biomarker studies which might explain why the results of many previous smaller studies on PI do not show conclusive results. Of course, the results of our study should be validated and further refined in other large, independent and well-characterised cohorts. Furthermore, it has to be investigated if PI is also of prognostic value in the palliative setting.

Perspectively, Ki-67 might be used in concert with other clinicopathological parameters to prognostically stratify patients. In preceding studies, we could already demonstrate that combined assessment of clinical, morphological, immunohistochemical, molecular, and imaging characteristics is helpful in this regard (Lederlin *et al*, 2013; Warth *et al*, 2014). Inclusion of a variety of different markers in multidimensional prognostic models will ultimately pave the way for an individualised tumour treatment based on evidence-based prognostic stratification. The combined assessment of clinical and (molecular) pathological risk factors has already been proven to be a helpful tool in other major cancer entities. Good examples for such clinically useful combined risk scores are nomograms, which have been established for a variety of different cancer entities, including, among others, prostate and breast cancers (Toi *et al*, 2012; Adamis and Varkarakis, 2014).

Taken together, we demonstrate that proliferation in ADC is an independent negative prognostic factor for survival, while in SQCC high proliferative activity is associated with better outcome. If these results can be confirmed in future prospective studies, standardised assessment of proliferation might become a useful biomarker in the daily routine diagnostic setting in NSCLC.

## ACKNOWLEDGEMENTS

The tissue bank of the National Center for Tumour Diseases, Heidelberg (NCT) is acknowledged for technical support.

## REFERENCES

- Adamis S, Varkarakis IM (2014) Defining prostate cancer risk after radical prostatectomy. *Eur J Surg Oncol* **40**(5): 496–504.
- Allegra CJ, Paik S, Colangelo LH, Parr AL, Kirsch I, Kim G, Klein P, Johnston PG, Wolmark N, Wieand HS (2003) Prognostic value of thymidylate synthase, Ki-67, and p53 in patients with Dukes' B and C colon cancer: a National Cancer Institute-National Surgical Adjuvant Breast and Bowel Project collaborative study. *J Clin Oncol* **21**(2): 241–250.
- Altman DG, Lausen B, Sauerbrei W, Schumacher M (1994) Dangers of using "optimal" cutpoints in the evaluation of prognostic factors. *J Natl Cancer Inst* **86**(11): 829–835.
- Budczies J, Klauschen F, Sinn BV, Gyorffy B, Schmitt WD, Darb-Esfahani S, Denkert C (2012) Cutoff Finder: a comprehensive and straightforward Web application enabling rapid biomarker cutoff optimization. *PLoS One* **7**(12): e51862.
- Cuzick J, Dowsett M, Pineda S, Wale C, Salter J, Quinn E, Zabaglo L, Mallon E, Green AR, Ellis IO, Howell A, Buzdar AU, Forbes JF (2011) Prognostic value of a combined estrogen receptor, progesterone receptor, Ki-67, and human epidermal growth factor receptor 2 immunohistochemical score and comparison with the Genomic Health recurrence score in early breast cancer. *J Clin Oncol* **29**(32): 4273–4278.
- Filipits M, Pirker R, Dunant A, Lantuejoul S, Schmid K, Huynh A, Haddad V, Andre F, Stahel R, Pignon JP, Soria JC, Popper HH, Le Chevalier T, Brambilla E (2007) Cell cycle regulators and outcome of adjuvant cisplatin-based chemotherapy in completely resected non-small-cell lung cancer: the International Adjuvant Lung Cancer Trial Biologic Program. *J Clin Oncol* **25**(19): 2735–2740.
- Goldhirsch A, Ingle JN, Gelber RD, Coates AS, Thurlimann B, Senn HJ (2009) Thresholds for therapies: highlights of the St Gallen International Expert Consensus on the primary therapy of early breast cancer 2009. *Ann Oncol* **20**(8): 1319–1329.
- Hommura F, Dosaka-Akita H, Mishina T, Nishi M, Kojima T, Hiroumi H, Ogura S, Shimizu M, Katoh H, Kawakami Y (2000) Prognostic significance of p27KIP1 protein and ki-67 growth fraction in non-small cell lung cancers. *Clin Cancer Res* **6**(10): 4073–4081.
- Jemal A, Bray F, Center MM, Ferlay J, Ward E, Forman D (2011) Global cancer statistics. *CA Cancer J Clin* **61**(2): 69–90.
- Kadota K, Suzuki K, Kachala SS, Zabor EC, Sima CS, Moreira AL, Yoshizawa A, Riely GJ, Rusch VW, Adusumilli PS, Travis WD (2012) A grading system combining architectural features and mitotic count predicts recurrence in stage I lung adenocarcinoma. *Mod Pathol* **25**(8): 1117–1127.
- Khor LY, Bae K, Paulus R, Al-Saleem T, Hammond ME, Grignon DJ, Che M, Venkatesan V, Byhardt RW, Rotman M, Hanks GE, Sandler HM, Pollack A (2009) MDM2 and Ki-67 predict for distant metastasis and mortality in men treated with radiotherapy and androgen deprivation for prostate cancer: RTOG 92-02. *J Clin Oncol* **27**(19): 3177–3184.
- Kuner R, Muley T, Meister M, Ruschhaupt M, Buness A, Xu EC, Schnabel P, Warth A, Poustka A, Sultmann H, Hoffmann H (2009) Global gene expression analysis reveals specific patterns of cell junctions in non-small cell lung cancer subtypes. *Lung Cancer* **63**(1): 32–38.
- Kwak EL, Bang YJ, Camidge DR, Shaw AT, Solomon B, Maki RG, Ou SH, Dezube BJ, Janne PA, Costa DB, Varella-Garcia M, Kim WH, Lynch TJ, Fidias P, Stubbs H, Engelman JA, Sequist LV, Tan W, Gandhi L, Mino-Kenudson M, Wei GC, Shreeve SM, Ratain MJ, Settleman J, Christensen JG, Haber DA, Wilner K, Salgia R, Shapiro GI, Clark JW, Iafrate AJ (2010) Anaplastic lymphoma kinase inhibition in non-small-cell lung cancer. *N Engl J Med* **363**(18): 1693–1703.
- Lederlin M, Puderbach M, Muley T, Schnabel PA, Stenzinger A, Kauczor HU, Heussel CP, Herth FJ, Hoffmann H, Dienemann H, Weichert W, Warth A (2013) Correlation of radio- and histomorphological pattern of pulmonary adenocarcinoma. *Eur Respir J* **41**(4): 943–951.
- Lee JS, Yoon A, Kalapurakal SK, Ro JY, Lee JJ, Tu N, Hittelman WN, Hong WK (1995) Expression of p53 oncoprotein in non-small-cell lung cancer: a favorable prognostic factor. *J Clin Oncol* **13**(8): 1893–1903.
- Leong AS, Zhuang Z (2011) The changing role of pathology in breast cancer diagnosis and treatment. *Pathobiology* **78**(2): 99–114.
- Luporsi E, Andre F, Spyrtos F, Martin PM, Jacquemier J, Penault-Llorca F, Tubiana-Mathieu N, Sigal-Zafrani B, Arnould L, Gompel A, Egele C, Poulet B, Clough KB, Crouet H, Fourquet A, Lefranc JP, Mathelin C, Rouyer N, Serin D, Spielmann M, Haugh M, Chenard MP, Brain E, de Cremoux P, Bellocq JP (2011) Ki-67: level of evidence and methodological considerations for its role in the clinical management of breast cancer: analytical and critical review. *Breast Cancer Res Treat* **132**(3): 895–915.
- Martin B, Paesmans M, Mascaux C, Berghmans T, Lothaire P, Meert AP, Lafitte JB, Sculier JP (2004) Ki-67 expression and patients survival in lung cancer: systematic review of the literature with meta-analysis. *Br J Cancer* **91**(12): 2018–2025.
- Meert AP, Martin B, Verdebout JM, Paesmans M, Berghmans T, Ninane V, Sculier JP (2004) Correlation of different markers (p53, EGF-R, c-erbB-2, Ki-67) expression in the diagnostic biopsies and the corresponding resected tumors in non-small cell lung cancer. *Lung Cancer* **44**(3): 295–301.
- Mok TS, Wu YL, Thongprasert S, Yang CH, Chu DT, Saijo N, Sunpaweravong P, Han B, Margono B, Ichinose Y, Nishiwaki Y, Ohe Y, Yang JJ, Chewaskulyong B, Jiang H, Duffield EL, Watkins CL, Armour AA, Fukuoka M (2009) Gefitinib or carboplatin-paclitaxel in pulmonary adenocarcinoma. *N Engl J Med* **361**(10): 947–957.
- Pelosi G, Rindi G, Travis WD, Papotti M (2014) Ki-67 antigen in lung neuroendocrine tumors: unraveling a role in clinical practice. *J Thorac Oncol* **9**(3): 273–284.
- Poleri C, Morero JL, Nieva B, Vazquez MF, Rodriguez C, de Titto E, Rosenberg M (2003) Risk of recurrence in patients with surgically resected stage I non-small cell lung carcinoma: histopathologic and immunohistochemical analysis. *Chest* **123**(6): 1858–1867.
- Pollack A, DeSilvio M, Khor LY, Li R, Al-Saleem TI, Hammond ME, Venkatesan V, Lawton CA, Roach 3rd M, Shipley WU, Hanks GE, Sandler HM (2004) Ki-67 staining is a strong predictor of distant metastasis and mortality for men with prostate cancer treated with radiotherapy plus androgen deprivation: Radiation Therapy Oncology Group Trial 92-02. *J Clin Oncol* **22**(11): 2133–2140.
- Rindi G, Falconi M, Klersy C, Albarello L, Boninsegna L, Buchler MW, Capella C, Caplin M, Couvelard A, Doglioni C, Delle Fave G, Fischer L, Fusai G, de Herder WW, Jann H, Komminoth P, de Krijger RR, La Rosa S, Luong TV, Pape U, Perren A, Ruzsiewicz P, Scarpa A, Schmitt A, Solcia E, Wiedenmann B (2012) TNM staging of neoplasms of the endocrine pancreas: results from a large international cohort study. *J Natl Cancer Inst* **104**(10): 764–777.
- Royston P, Altman DG, Sauerbrei W (2006) Dichotomizing continuous predictors in multiple regression: a bad idea. *Stat Med* **25**(1): 127–141.

- Sofocleous CT, Garg SK, Cohen P, Petre EN, Gonen M, Erinjeri JP, Downey RJ, Travis WD, Solomon SB (2013) Ki 67 is an independent predictive biomarker of cancer specific and local recurrence-free survival after lung tumor ablation. *Ann Surg Oncol* **20**(Suppl 3): S676–S683.
- Takahashi S, Kamata Y, Tamo W, Koyanagi M, Hatanaka R, Yamada Y, Tsushima T, Takaya S, Fukuda I (2002) Relationship between postoperative recurrence and expression of cyclin E, p27, and Ki-67 in non-small cell lung cancer without lymph node metastases. *Int J Clin Oncol* **7**(6): 349–355.
- Toi M, Benson JR, Winer EP, Forbes JF, von Minckwitz G, Golshan M, Robertson JF, Sasano H, Cole BF, Chow LW, Pegram MD, Han W, Huang CS, Ikeda T, Kanao S, Lee ES, Noguchi S, Ohno S, Partridge AH, Rouzier R, Tozaki M, Sugie T, Yamauchi A, Inamoto T (2012) Preoperative systemic therapy in locoregional management of early breast cancer: highlights from the Kyoto Breast Cancer Consensus Conference. *Breast Cancer Res Treat* **136**(3): 919–926.
- Travis WD, Brambilla E, Noguchi M, Nicholson AG, Geisinger KR, Yatabe Y, Beer DG, Powell CA, Riely GJ, Van Schil PE, Garg K, Austin JH, Asamura H, Rusch VW, Hirsch FR, Scagliotti G, Mitsudomi T, Huber RM, Ishikawa Y, Jett J, Sanchez-Cespedes M, Sculier JP, Takahashi T, Tsuboi M, Vansteenkiste J, Wistuba I, Yang PC, Aberle D, Brambilla C, Flieder D, Franklin W, Gazdar A, Gould M, Hasleton P, Henderson D, Johnson B, Johnson D, Kerr K, Kuriyama K, Lee JS, Miller VA, Petersen I, Roggli V, Rosell R, Saijo N, Thunnissen E, Tsao M, Yankelewitz D (2011) International association for the study of lung cancer/American Thoracic Society/European Respiratory Society international multi-disciplinary classification of lung adenocarcinoma. *J Thorac Oncol* **6**(2): 244–285.
- Urruticoechea A, Smith IE, Dowsett M (2005) Proliferation marker Ki-67 in early breast cancer. *J Clin Oncol* **23**(28): 7212–7220.
- Viale G, Giobbie-Hurder A, Regan MM, Coates AS, Mastropasqua MG, Dell'Orto P, Maiorano E, MacGrogan G, Braye SG, Ohlschlegel C, Neven P, Orosz Z, Olszewski WP, Knox F, Thurlimann B, Price KN, Castiglione-Gertsch M, Gelber RD, Gusterson BA, Goldhirsch A (2008) Prognostic and predictive value of centrally reviewed Ki-67 labeling index in postmenopausal women with endocrine-responsive breast cancer: results from Breast International Group Trial 1-98 comparing adjuvant tamoxifen with letrozole. *J Clin Oncol* **26**(34): 5569–5575.
- Warth A, Fink L, Fisseler-Eckhoff A, Jonigk D, Keller M, Ott G, Rieker RJ, Sinn P, Soder S, Soltermann A, Willenbrock K, Weichert W (2013) Interobserver agreement of proliferation index (Ki-67) outperforms mitotic count in pulmonary carcinoids. *Virchows Arch* **462**(5): 507–513.
- Warth A, Muley T, Herpel E, Meister M, Herth FJ, Schirmacher P, Weichert W, Hoffmann H, Schnabel PA (2012a) Large-scale comparative analyses of immunomarkers for diagnostic subtyping of non-small-cell lung cancer biopsies. *Histopathology* **61**(6): 1017–1025.
- Warth A, Muley T, Meister M, Stenzinger A, Thomas M, Schirmacher P, Schnabel PA, Budczies J, Hoffmann H, Weichert W (2012b) 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* **30**(13): 1438–1446.
- Warth A, Penzel R, Lindenmaier H, Brandt R, Stenzinger A, Herpel E, Goeppert B, Thomas M, Herth FJ, Dienemann H, Schnabel PA, Schirmacher P, Hoffmann H, Muley T, Weichert W (2014) EGFR, KRAS, BRAF and ALK gene alterations in lung adenocarcinomas: patient outcome, interplay with morphology and immunophenotype. *Eur Respir J* **43**(3): 872–883.
- Yamashita S, Moroga T, Tokuishi K, Miyawaki M, Chujo M, Yamamoto S, Kawahara K (2011) Ki-67 labeling index is associated with recurrence after segmentectomy under video-assisted thoracoscopic surgery in stage I non-small cell lung cancer. *Ann Thorac Cardiovasc Surg* **17**(4): 341–346.
- Zahel T, Krysa S, Herpel E, Stenzinger A, Goeppert B, Schirmacher P, Hoffmann H, Schnabel PA, Warth A (2012) Phenotyping of pulmonary carcinoids and a Ki-67-based grading approach. *Virchows Arch* **460**(3): 299–308.

This work is published under the standard license to publish agreement. After 12 months the work will become freely available and the license terms will switch to a Creative Commons Attribution-NonCommercial-Share Alike 3.0 Unported License.

Supplementary Information accompanies this paper on British Journal of Cancer website (<http://www.nature.com/bjc>)