

Editor's note:

In the era of personalized medicine, a critical appraisal new developments and controversies are essential in order to derived tailored approaches. In addition to its educative aspect, we expect these discussions to help younger researchers to refine their own research strategies.

Controversies on Lung Cancer: Pros and Cons

What is better/reliable, mitosis counting or Ki67/MIB1 staining?

Mark Kriegsmann, Arne Warth

Institute of Pathology, Heidelberg University, Heidelberg, Germany

Correspondence to: Arne Warth, MD. Institute of Pathology, Heidelberg University, INF 224, D-69120 Heidelberg, Germany.

Email: arne.warth@med.uni-heidelberg.de.

Submitted May 17, 2016. Accepted for publication Oct 13, 2016.

doi: 10.21037/tlcr.2016.10.11

View this article at: <http://dx.doi.org/10.21037/tlcr.2016.10.11>

Uncontrolled proliferation is a hallmark of cancers and tumor cell division is a prerequisite for the effectiveness of most conventional anticancer treatments. Thus, by either counting mitotic figures (mitotic index; MI) or immunohistochemical analysis of proliferation associated antigens like Ki-67 (proliferation index; PI), proliferation in cancers was assessed in a multitude of scientific studies. PubMed research for the terms “cancer” AND “proliferation” as well as “lung cancer” AND “proliferation” results in >190,000 and >16,000 hits, respectively. However, despite this exhausting amount of literature the translation of proliferation assessment into daily routine has largely failed. It has its role in some grading systems, e.g., Elston-Ellis grading for breast cancer or the French Federation of Cancer Centers Sarcoma Group (FNCLCC) grading for sarcomas, but has currently no diagnostic meaning for thoracic tumors despite for neuroendocrine tumors, where MI is used to separate typical from atypical carcinoids or large cell neuroendocrine carcinomas. Why was this promising biomarker lost in translation?

Two large meta-analyses investigated the clinical impact of PI on lung cancer (1,2). The first meta-analysis included 37 studies of non-small cell lung cancer (NSCLC) published between 1991 and 2002 (1). The second meta-analysis included 28 articles of NSCLC published between 2000 and 2012 (2). Both studies reveal that the prognostic and predictive impact of the PI is debatable. The first meta-analysis concluded that high PI correlates with poor

prognosis in lung cancer (1). However, only 10 out of 29 studies performed on NSCLC demonstrated a statistically significant (P value: <0.05) negative prognostic effect for the PI. Moreover, the more recent meta-analysis revealed that there is no consensus on the prognostic impact neither in uni- nor in multivariate analysis (2). There were no studies included in the meta-analysis that could provide evidence for a predictive value of the PI. However, the majority of the studies analyzed suffered from a lack of standardized methodology including preanalytics such as the type and time of fixation, means of storage, but also analytics as the usage of different antibodies, staining protocols, and finally post-analytics as different cut-offs and methodologies for PI assessment were applied. The cut-offs applied ranged from 1% to 60% of and many articles did not even mention the number of analyzed tumor cells. Thus, comparison of the studies is difficult due to a lack of standardization. To this end it is well known that different preanalytic and analytic conditions have a significant impact on the percentage of immunoreactive nuclei (3). Since NSCLC comprise different entities such as adenocarcinomas (ADC) or squamous cell carcinomas (SqCC) this may account for differences in the reported PIs as well. Indeed, it has been shown that large cell neuroendocrine carcinomas demonstrate the highest PI followed by SqCC, sarcomatoid carcinomas, large cell carcinomas, and ADC and there is even a significant difference of PI among ADC subtypes (4,5). In a recent study on 1,056 NSCLC we could

demonstrate that PI is of independent prognostic value for ADC and adenosquamous carcinomas applying an optimized cut-off of 25% (5). In accordance with our data, Kadota and colleagues showed that MI in combination with architectural grade is an independent recurrence predictor in stage I ADC (6), another study even demonstrated that MI trumps T stage in early tumor stages (7).

However, by plotting the hazard ratios of all potential cut-offs we could clearly demonstrate that ADC with a higher PI always have a worse outcome compared to those with a lower PI, independent of the cut-off chosen (5). This makes it very difficult to establish a clinically meaningful cut-off since dichotomization of a continuous prognostic variable is always arbitrary. Although 25% PI resulted in the best prognostic stratification in our cohort, ADC with 24% and 26% PI do likely not behave different from a tumorbiological point of view and this cut-off would be a very weak argument for different clinical decisions with respect to therapy. These findings argue for a central “grey zone” as proposed for breast cancer in the most recent St.-Gallen consensus (8). In SqCC high PI was surprisingly found to be associated with better survival in some studies (5,9). The underlying mechanisms of this finding are not yet clear. It may be speculated that fast uncontrolled tumor growth could lead to an inadequate blood supply with a higher propensity of the tumor to develop necrosis, which in turn might induce a stronger antitumor immune response. Another explanation could be that chemotherapeutic treatment administered later in the disease course might have stronger effects in patients with rapidly proliferating tumors (5). Independent of the reason for these observations this finding clearly shows that PI assessment is not trivial and must be performed highly standardized and separately for each tumor entity.

In pulmonary neuroendocrine tumors MI is firmly implemented in routine diagnostics for diagnostic and prognostic purposes but assessment of PI is still not officially recommended. Whereas both MI and PI seem to have its role in this setting, it is interesting to have a closer look on the reproducibility of both parameters. Diagnostic criteria must necessarily achieve a high inter-observer agreement in order to allow for a reliable diagnosis in the majority of cases independent of the evaluating pathologist. Of note, there is increasing evidence that assessment of PI in pulmonary carcinoids results in a higher inter-observer agreement compared to MI (10,11). The superiority of PI compared to MI has also been demonstrated in other entities such as soft tissue sarcomas (12). What are potential

reasons for this? It is notable that many pathological parameters are becoming more and more standardized, however, there are no internationally accepted guidelines or criteria for a standardized and robust assessment of MI, yet (13). Indeed, although there are mitotic figures which most pathologists will agree on, there is undoubtedly a grey zone in the differentiation to apoptotic bodies or even tissue artifacts in which it depends on the single pathologist whether he or she interprets a given finding as mitotic figure or not (13). This likely explains why other inter-observer studies on MI also show a poor inter-observer agreement in soft tissue tumors (14), brain tumors (15), prostate cancer (16), and breast cancer (17). As a solution of this issue the application of strict rules for the identification of mitotic figures (18) or usage of mitotic immunomarkers such as phosphohistoneH3 (PHH3) (19) have been proposed, which might improve the inter-observer agreement. Regardless of all methodological shortcomings (see below) PI assessment has the advantage that positively stained cells are specifically delineated and the quick and easy selection of hotspot areas even at low power scanning magnification might increase the chance that different pathologists assess the PI in the same region of the slide whereas this might not be necessarily the case for MI, where the selection of areas with high proliferation might be more random. However, although PI assessment results in a higher inter-observer agreement according to yet available data, quantification of immunohistochemical stains is also affected by several issues. Different fixation standards, staining protocols, and varying methods of interpretation of Ki-67 may potentially lead to high inter-laboratory variability (20). To reduce this center-to-center variability, reference standards for Ki-67 staining and evaluation need to be established.

Besides morphology driven approaches such as mitosis counting and the determination of PI by Ki-67 staining, it is important to note that new molecular methods such as proliferation-based gene expression signatures and derived cell cycle progression- and molecular prognostic scores may provide objective, quantitative, and reproducible stratification for prognosis and response to therapy (21). Furthermore, the emerging field of digital image analysis will substantially increase the chance for a high standardization, at least with respect to PI assessment.

Taken together, assessment of proliferation is still a promising tissue-derived biomarker with high prognostic and potentially also predictive implications in many tumor entities. In the lungs there is an established role for

neuroendocrine tumors and perspective also for a better stratification of the large group of ADC with intermediate prognosis (acinar and papillary predominant ADC). The current major limitation is the lack of standardization. Next, the establishment of clinically meaningful cut-offs requires detailed analysis of large and well characterized cohorts and needs to be done separately for all entities. With the advent of novel technologies and international harmonization of methodologies, for example as demonstrated for ALK immunohistochemistry, proliferation assessment might emerge as a relevant parameter which, in principal, can be easily assessed in a time- and cost-effective manner by pathologists worldwide.

Acknowledgements

None.

Footnote

Conflicts of Interest: The authors have no conflicts of interest to declare.

Related to: Chirieac LR. Ki-67 expression in pulmonary tumors. *Transl Lung Cancer Res* 2016;5:547-51.
Kriegsmann M, Warth A. Ki-67 expression in pulmonary tumors—reply. *Transl Lung Cancer Res* 2016;5:552-3.
Chirieac LR. Tumor cell proliferation, proliferative index and mitotic count in lung cancer. *Transl Lung Cancer Res* 2016;5:554-6.

References

- Martin B, Paesmans M, Mascaux C, et al. Ki-67 expression and patients survival in lung cancer: systematic review of the literature with meta-analysis. *Br J Cancer* 2004;91:2018-25.
- Jakobsen JN, Sørensen JB. Clinical impact of ki-67 labeling index in non-small cell lung cancer. *Lung Cancer* 2013;79:1-7.
- Grillo F, Pigozzi S, Ceriolo P, et al. Factors affecting immunoreactivity in long-term storage of formalin-fixed paraffin-embedded tissue sections. *Histochem Cell Biol* 2015;144:93-9.
- Tsubochi H, Sato N, Hiyama M, et al. Combined analysis of cyclooxygenase-2 expression with p53 and Ki-67 in nonsmall cell lung cancer. *Ann Thorac Surg* 2006;82:1198-204.
- Warth A, Cortis J, Soltermann A, et al. Tumour cell proliferation (Ki-67) in non-small cell lung cancer: a critical reappraisal of its prognostic role. *Br J Cancer* 2014;111:1222-9.
- Kadota K, Suzuki K, Kachala SS, et al. A grading system combining architectural features and mitotic count predicts recurrence in stage I lung adenocarcinoma. *Mod Pathol* 2012;25:1117-27.
- Duhig EE, Dettrick A, Godbolt DB, et al. Mitosis trumps T stage and proposed international association for the study of lung cancer/american thoracic society/european respiratory society classification for prognostic value in resected stage 1 lung adenocarcinoma. *J Thorac Oncol* 2015;10:673-81.
- Jackisch C, Harbeck N, Huober J, et al. 14th St. Gallen International Breast Cancer Conference 2015: Evidence, Controversies, Consensus - Primary Therapy of Early Breast Cancer: Opinions Expressed by German Experts. *Breast Care (Basel)* 2015;10:211-9.
- Poleri C, Morero JL, Nieva B, et al. Risk of recurrence in patients with surgically resected stage I non-small cell lung carcinoma: histopathologic and immunohistochemical analysis. *Chest* 2003;123:1858-67.
- Warth A, Fink L, Fisseler-Eckhoff A, et al. Interobserver agreement of proliferation index (Ki-67) outperforms mitotic count in pulmonary carcinoids. *Virchows Arch* 2013;462:507-13.
- Swarts DR, van Suylen RJ, den Bakker MA, et al. Interobserver variability for the WHO classification of pulmonary carcinoids. *Am J Surg Pathol* 2014;38:1429-36.
- Hasegawa T, Yamamoto S, Nojima T, et al. Validity and reproducibility of histologic diagnosis and grading for adult soft-tissue sarcomas. *Hum Pathol* 2002;33:111-5.
- Barry M, Sinha SK, Leader MB, et al. Poor agreement in recognition of abnormal mitoses: requirement for standardized and robust definitions. *Histopathology* 2001;38:68-72.
- Coindre JM, Trojani M, Contesso G, et al. Reproducibility of a histopathologic grading system for adult soft tissue sarcoma. *Cancer* 1986;58:306-9.
- Gilles FH, Tavaré CJ, Becker LE, et al. Pathologist interobserver variability of histologic features in childhood brain tumors: results from the CCG-945 study. *Pediatr Dev Pathol* 2008;11:108-17.
- Svanholm H, Starklint H, Barlebo H, et al. Histological evaluation of prostatic cancer. (II): Reproducibility of a histological grading system. *APMIS* 1990;98:229-36.
- Tsuda H, Akiyama F, Kurosumi M, et al. Evaluation of the interobserver agreement in the number of mitotic

- figures of breast carcinoma as simulation of quality monitoring in the Japan National Surgical Adjuvant Study of Breast Cancer (NSAS-BC) protocol. *Jpn J Cancer Res* 2000;91:451-7.
18. Yadav KS, Gonuguntla S, Ealla KK, et al. Assessment of interobserver variability in mitotic figure counting in different histological grades of oral squamous cell carcinoma. *J Contemp Dent Pract* 2012;13:339-44.
 19. Schimming TT, Grabelus F, Roner M, et al. pHH3 immunostaining improves interobserver agreement of mitotic index in thin melanomas. *Am J Dermatopathol* 2012;34:266-9.
 20. Hsu CY, Ho DM, Yang CF, et al. Interobserver reproducibility of MIB-1 labeling index in astrocytic tumors using different counting methods. *Mod Pathol* 2003;16:951-7.
 21. Eguchi T, Kadota K, Chaft J, et al. Cell cycle progression score is a marker for five-year lung cancer-specific mortality risk in patients with resected stage I lung adenocarcinoma. *Oncotarget* 2016;7:35241-56.

Cite this article as: Kriegsmann M, Warth A. What is better/reliable, mitosis counting or Ki67/MIB1 staining? *Transl Lung Cancer Res* 2016;5(5):543-546. doi: 10.21037/tlcr.2016.10.11