



Published in final edited form as:

Virchows Arch. 2020 August ; 477(2): 259–267. doi:10.1007/s00428-020-02750-7.

Reproducibility of histologic prognostic parameters for mantle cell lymphoma: cytology, Ki67, p53 and SOX11

Giorgio A. Croci^{1,2}, Eva Hoster^{3,4}, Sílvia Beà^{5,6}, Guillem Clot^{5,6}, Anna Enjuanes^{5,6}, David W. Scott⁷, José Cabeçadas⁸, Luis Veloza⁹, Elias Campo^{5,6,9}, Erik Clasen-Linde¹⁰, Rashmi S. Goswami¹¹, Lars Helgeland¹², Stefano Pileri¹³, Grzegorz Rymkiewicz¹⁴, Sarah Reinke¹, Martin Dreyling⁴, Wolfram Klapper¹

¹Department of Pathology, Haematopathology Section and Lymph Node Registry, University of Kiel, Kiel, Germany ²Pathology Unit, Department of Pathophysiology and Transplantation, University of Milan and Fondazione IRCCS, Ca' Granda - Maggiore Policlinico, Milan, Italy ³Institute for Medical Information Processing, Biometry and Epidemiology (IBE), Ludwig-Maximilians-Universität Munich, Munich, Germany ⁴Department of Medicine III, University Hospital of the Ludwig Maximilians University Munich, Munich, Germany ⁵Institute for Biomedical Research August Pi i Sunyer (IDIBAPS), Barcelona, Spain ⁶Centro de Investigación Biomédica en Red de Cáncer (CIBERONC), Spain ⁷Centre for Lymphoid Cancer, British Columbia Cancer, Vancouver, Canada ⁸Department of Pathology, Instituto Português de Oncologia de Lisboa Francisco Gentil, E.P.E., Lisbon, Portugal ⁹Hematopathology Unit-Laboratory of Pathology, Hospital Clínic of Barcelona, University of Barcelona, Barcelona, Spain ¹⁰Department of Pathology, Rigshospitalet, Copenhagen, Denmark ¹¹Sunnybrook Health Sciences Centre, Dept. of Laboratory Medicine and Molecular Diagnostics, and Department of Laboratory Medicine and Pathobiology, University of Toronto, Toronto, Canada ¹²Department of Pathology, Haukeland University Hospital, Bergen, Norway ¹³Division of Diagnostic Haematopathology, European Institute of Oncology, IRCCS, Milan, Italy ¹⁴Department of Pathology and Laboratory Diagnostics, Maria Skłodowska-Curie Institute-Oncology Centre, Warsaw, Poland

Abstract

Mantle cell lymphoma (MCL) shows a clinical aggressiveness that varies from patient to patient. Despite major advances in outcomes with current immunochemotherapy, the future development

Terms of use and reuse: academic research for non-commercial purposes, see here for full terms. <http://www.springer.com/gb/open-access/authors-rights/aam-terms-v1>

Corresponding Author: Giorgio Alberto Croci, UOC di Anatomia Patologica, IRCCS Fondazione Ca' Granda – Ospedale Maggiore Policlinico, Via Francesco Sforza, 35 - 20122 Milano, giorgio.croci@unimi.it.

Contributions: all authors contributed to the study conception and design. Giorgio A. Croci and Wolfram Klapper designed the research; Giorgio A. Croci, Wolfram Klapper, José Cabeçadas, Elias Campo, Luis Veloza, Erik Clasen-Linde., Rashmi S. Goswami, Lars Helgeland., Stefano Pileri and Grzegorz Rymkiewicz performed histopathological assessment; Giorgio A. Croci and Sarah Reinke performed image analysis; Sílvia Beà, David W. Scott, Guillem Clot and Anna Enjuanes performed molecular and bioinformatic analyses; Eva Hoster planned and performed statistical analyses; Martin Dreyling provided medical support; Giorgio A. Croci, Wolfram Klapper, Sílvia Beà and Eva Hoster wrote the manuscript; all authors approved the final version of the manuscript.

Compliance with Ethical Standards

Conflict of interest: the authors declare that they have no conflict of interest.

Publisher's Disclaimer: This Author Accepted Manuscript is a PDF file of a an unedited peer-reviewed manuscript that has been accepted for publication but has not been copyedited or corrected. The official version of record that is published in the journal is kept up to date and so may therefore differ from this version.

of therapies requires risk stratification to tailor therapy intensity. Within the group of reference pathologists for the ongoing trials of the European MCL Network, we performed a round robin test on a tissue microarray, to evaluate the reproducibility in assessing the biomarkers of outcome in MCL. Cytological subtype, Ki67-index and expression of p53 and SOX11 were evaluated on 20 diagnostic tumour samples by eight participating labs independently. We demonstrate that the assessment of the proliferation index by counting the Ki67 positive cells as well as assessment of SOX11 and p53 expression status are reproducible between labs. For the most established prognostic biomarker, Ki67, the intra-class correlation coefficient was very good when assessed as a continuous parameter (0.87). The agreement was lower when the values were analysed in a dichotomized way applying the commonly used cut-off of 30% [$\kappa=0.65$, complete concordance of all labs in 13/20 (65%)]. Cases with discrepant results between labs in the dichotomized analysis showed mean values close to the cut-off of 30%. Centralized scoring and digital image analysis revealed results in line with the scores from individual labs. All cases in our cohort were additionally assessed for gene expression signatures and of *TP53* gene alterations. Given the good reproducibility when guidelines of assessment are applied, the biomarker studied in this inter-laboratory test present potential candidates, to be enhanced for risk-stratification in the future clinical trials.

Keywords

mantle cell lymphoma; Ki67; SOX11; p53; cytology

Introduction

Mantle cell lymphoma (MCL) has been considered the paradigm of a clinically aggressive neoplasm with dismal prognosis in the past, but it is now regarded as a disease with inter-tumour heterogeneity in response to treatment and survival, but also in its molecular features. The evolution of therapy regimens for this relatively rare lymphoid neoplasm has improved outcomes over the last decade [1]. However, despite the heterogeneity in outcome, MCL patients are currently not stratified by risk factors other than patient age before start of therapy, with the exception that the leukemic non-nodal variant of MCL is frequently managed differently than the “common” nodal MCL. However, the future development of MCL therapeutic regimens requires patient-specific tailoring of therapy (e.g. by biomarkers).

Currently, MCL is the sole lymphoid neoplasm for which a pathologic parameter is integrated within a prognostic index and retains its significance [2,3]. Indeed, risk assessment in MCL is best provided by the combined MCL International Prognostic Index (MIPI-c), developed within the European MCL Network [4], which aggregates the clinical parameters from previous MIPI with the evaluation of the antigen Ki67 by immunohistochemistry (IHC) as a measure of cell proliferation.

A prerequisite for the implementation of a biomarker into patient stratification is its reproducibility among laboratories. To address this issue, we conducted an inter-laboratory test among reference pathology labs of the European MCL Network, which presents a potential framework for future trials using risk adapted patient stratification. We aimed to

analyse prognostic biomarkers of MCL that have been well established in retrospective analysis. Our test included pathologic parameters, which have been recently proved helpful on retrospective series from clinical trials, such as Ki67, p53 and SOX11 protein expression and cytological variants [5–7] and which are potentially widely available in pathology labs. Our analysis was completed by assessment of the prognostic gene expression assay MCL35 [8,9], a signature for leukemic non-nodal MCL (L-MCL16 assay) [10] and for *TP53* gene alterations (mutations and deletions).

Methods

Samples and stainings

A series was provided by selecting 20 cases from the files of the Lymph Node Registry in Kiel, including all cytological and phenotypic features observed in MCL (table 1), regardless of clinical presentation. From each formalin-fixed, paraffin-embedded (FFPE) tissue, two cores of 0.6 mm in diameter were transferred into a tissue microarray (TMA). The TMA was cut into 3µm slides and immediately sent to the 8 participating pathology centres (Barcelona, Bergen, Copenhagen, Kiel, Lisbon, Milan, Toronto and Warsaw) within 24 hours. Each laboratory was requested to stain the slides for Hematoxylin-Eosin (H&E), Ki67, p53 and SOX11 according to their local protocols for immunohistochemistry (IHC, supplementary table 1).

Definition of variables

Cytological variant was assessed as small cell, classic, blastoid or pleomorphic [11]. For additional analysis, cytology variants were lumped into low grade (classic or small cell) and high grade (blastoid or pleomorphic).

Ki67 index was assessed as quantitative value, reflecting the percentage of Ki67 positive tumour cells according to the published guidelines [12]. For the analysis, the average of two manual counts of 100 cells each in separated representative areas (i.e. not containing residual germinal centers or hot spots of proliferation [12]) was calculated. Subsequently, Ki67 index was classified as low (<30%) or high (≥ 30%). p53 and SOX11 IHC were assessed by counting positive cells, regardless of intensity of staining, in a semi-quantitative manner at 10% intervals, as previously described [7]. Accordingly, for p53 a four-tiered score was assigned as negative (0% of lymphoma cells strongly positive), low (1–10%), intermediate (11–50%) and high (>50%). For SOX11, a 3-tiered score was assigned as negative (0%), low (1–10%) and positive (>10%). Scores were further transformed into two-tiered systems: non-high (< 50%) vs high (>50%) for p53 and negative (< 10%) vs positive (>10%) for SOX11.

Centralized evaluation and image analysis

After assessment in each pathology centre, the slides were collected and a centralized score was performed by an independent pathologist (GC), both manually and by quantitative image analysis. The latter was carried out via TissueStudio 64 software (Definiens AG, 80636 Munich, Germany). Scans from each section were acquired (Hamamatsu Nanozoomer; Hamamatsu Photonics, Herrsching am Ammersee, Germany) and processed.

Prior to analysis, the software requires staining thresholds to be set for nuclei and staining recognition, a task which was performed separately for Ki67, p53 and SOX11, arbitrarily on the slides from Kiel. For nucleus detection, a typical nucleus size was defined as $25 \mu\text{m}^2$, to allow the software to separate packed nuclei, and a size filter was applied to remove stained artefacts with a size $< 4 \mu\text{m}^2$. Finally, a chromogen threshold was set at 0.3 for assessment of positive cells. Prior to analysis, each single core from digital scans was reviewed and region of interest were manually selected, to exclude from the analysis hotspots of proliferation and areas with gross technical artefacts.

Molecular analysis

For each case, DNA and RNA were extracted from whole FFPE sections taken from the corresponding, paraffin embedded blocks using QIAGEN AllPrep DNA/RNA FFPE Kit (QIAGEN, Hilden, Germany) after deparaffinization. The NanoString platform (NanoString Technologies, Seattle, WA) was used to measure the proliferation signature (MCL35 assay) [8,9], and L-MCL16 assay [10]. Three samples did not pass the quality controls and were removed from subsequent analyses. *TP53* mutations in the coding region were assessed by targeted next generation sequencing using Nextera Flex for enrichment protocol (Illumina) combined with custom DNA-probes from IDT (xGen Lockdown_probes) and run in a MiSeq sequencer. Determination of 17p deletion was performed with fluorescence in situ hybridization and a probe that includes a centromeric control (Vysis Abbott, USA).

Statistical analysis

The data originated from the assessment of 4 markers (cytology, SOX11, p53, Ki67 index) by 8 scorers (Labs A-H) on 20 samples. Samples with missing values were excluded for the respective assessment of agreement. For quantitative and ordered categorical data, intra-class correlation coefficient (ICC) was calculated using two-way ANOVA (considering samples and labs as random samples of larger populations), estimating agreement and using single scorer values. For binary and unordered categorical data, Fleiss Kappa (k value) was evaluated. ICC and k reflect the ratios of observed agreement in relation to agreement by chance and vary between -1 (complete disagreement of observers), 0 (agreement by chance) and 1 (all observers agree). As additional sensitivity analyses, each lab was omitted once from every analysis, in order to check the stability of the results. BlandAltman-Plots were produced for Ki67 index by manual counting at each lab in comparison to Kiel lab as reference.

The statistical analysis was performed using the package irr in R version 3.5.1 (www.r-project.org). Spearman's correlation was used to measure the association between the MCL35 score and Ki67 staining. The mean L-MCL16 score between SOX11 negative and positive cases was compared with a Welch's t-test.

Results

Detailed ICC and kappa values are listed in supplementary table 2.

Cytology

Agreement between observers in cytology assessment was moderate when analysing all categories (small-cell, classical, pleomorphic, blastoid; kappa=0.39; supplementary figure 1). Omitting one lab resulted in ICC values ranging between 0.36 and 0.46. Complete concordance among all eight centres was reached in only 6/20 (30%) of cases. A slightly better agreement was achieved when cytology subtypes were combined into low grade (classical+small cell) and high grade (pleomorphic and blastoid) categories, respectively (kappa=0.57, with full concordance in 11/20 [55%] cases). Omitting one lab resulted in kappa values ranging between 0.53 and 0.68.

Ki67

The agreement for Ki67 assessed as a continuous parameter was very good (ICC 0.87, Figure 1, range 0.86–0.89 after omitting one lab at a time; Bland-Altman plot, supplementary figure 2). The agreement was lower when the values were analysed in a dichotomized way applying the commonly used cut-off of 30% (kappa=0.65, complete concordance of all labs in 13/20 (65%), Figure 1; range kappa 0.62–0.72 after omitting one lab at a time). Cases with discrepant results between labs in the dichotomized analysis showed mean values close to the cut-off of 30% (Figure 1). The MCL35 assay was evaluable in 17 of the 20 cases and assigned a high-risk in 5 cases, standard-risk in 5 cases and low-risk in 7 cases, based on the proliferation (Figure 3). The MCL35 signature (as a continuous variable) and the Ki67 staining showed a high correlation with the reference (Kiel) score (Spearman rho = 0.82, IC95% = [0.46, 0.97], p-value < 0.001; table 1).

In a second step, all stainings from individual labs were collected centrally and manually counted by one observer. The concordance between the central manual counts and the individual counts of the labs was very good (range: 0.88–0.98). Relevant discrepancies between the individual lab scores and the central score were occasionally due to heterogeneity in staining intensity and/or to a variable content of proliferation hotspots and/or residual germinal centres within the slides stained and scored by different labs (Figure 2).

Centrally performed digital image analysis achieved a very good agreement with the counts derived from the individual labs (ICC range: 0.86–0.95) as well as with the centrally generated manual counts (ICC range: 0.88–0.95).

p53

The agreement for p53 expression analysis was good using the four-tiered system (ICC=0.71; range 0.67–0.75 after omitting one lab at a time), even though complete concordance among all centres was reached in only 3/20 (15%) cases (supplementary figure 3). Interestingly all cases with complete agreement showed high level of p53 expression. When the p53 values were analysed in a dichotomized way as high and non-high cases, kappa value was high (0.95, range 0.90–1 after omitting one lab, full concordance in 18/20 (90%) cases, supplementary figure 3). *TP53* gene was mutated in 3 cases (all p53 high and SOX11-negative), one of which also carrying 17p deletion, whereas a further one had a deletion (identified both by NGS and FISH), but no mutation (Table 1). Notably, p53 high

expression was observed only in the SOX11-low subset and was characterized by positivity of the vast majority of nuclei in a strong and homogeneous pattern (Figure 2).

SOX11

IHC for SOX11 expression displayed a very good concordance between individual labs using a threetiered system (ICC=0.95, range 0.95–0.96 when one lab omitted), with full concordance between labs in 17/20 (85%; supplementary figure 4) and comparable agreement between labs when dichotomized as positive and non-positive (kappa=0.91, range 0.90–0.93 with one lab omitted, full concordance between labs 18/20, 90%). Representative examples of stainings are illustrated in Figure 2. SOX11 mRNA level displayed a good concordance with SOX11 staining, with the notable exception of case #20. On the other hand, the L-MCL16 score was higher in the SOX11-positive cases (785.7 versus 607.3, p-value = 0.001; table 1) and unequivocally separated the two subtypes (Figure 3).

Discussion

The purpose of the present study was to assess the performance of experienced hematopathologic laboratories in the assessment of prognostic parameters of MCL and to provide informative hints to general pathologists. We avoided the bias of the centralization for staining and processed the slides at each separate lab, as our intention was to mimic the real world workup of diagnostic cases. It is important to consider that our test was not designed to evaluate the prognostic power of the biomarkers, nor to correlate biomarkers of prognosis with each other. Our analysis included molecular analysis for imbalances and gene expression signatures, which in the clinical setting are surrogated by the respective immunohistochemical stainings. Collectively, despite the low number of cases tested, the results from morphologic molecular analyses largely overlap, but with some caveats discussed below.

Our data showed only a moderate agreement among labs in cytology assessment, thus it may be argued that cytology should not be used as a prognostic marker in clinical practice. As well, the poor prognostic significance historically attributed to blastoid and pleomorphic variants seems replaceable by the Ki67 index and TP53 status [3,13]. It is our recommendation to include cytologic variants in the pathology reports, as it represents a helpful diagnostic feature, but not to use it as a risk factor for patient stratification.

Immunostaining for p53 is a promising tool for prognostic purpose, as it acts as a surrogate marker for TP53 mutation and/or 17p deletion, thus of p53 pathway deregulation. When complying with the cutoffs assessed in the larger clinical series [7], p53 scoring yielded a good overall concordance, with most of the disagreement in cases within the negative to low (0–10%) range, which seems clinically a less important subgroup. Agreement was excellent when the task was restricted to sorting out high (>50%) from non-high p53 cases. Dichotomized scores miss full negative cases, for which there is some existing evidence on their association with inferior outcome (i.e., either for absent or nonfunctional protein) [7]. To this regard, our observation of case #16, carrying a 17p deletion, but with complete p53 negativity, stresses the point that, besides its promising prognostic power [6], p53 immunohistochemistry is not a valid substitute of molecular analysis. The pattern of p53

positivity may vary in respect to staining intensity. Notably, a homogeneous pattern of reactivity (either low or intense, or fully negative) was most commonly observed in truly positive cases (p53-high), in which dysregulated protein expression can be related to a clonal anomaly, whereas a variable intensity expression intensity, with occurrence of negative nuclei, is encountered in p53-non-high cases, most likely reflecting the expression of a functional protein. From a technical perspective, it should be finally noted that all laboratories applied the same DO-7 clone (albeit from different producers).

SOX11 staining is regarded as a strong diagnostic tool to identify the non-nodal variant of MCL, which is typically SOX11-negative, as compared to the more common nodal type of MCL, usually SOX11-positive. However, absence of SOX11 expression seems to impact prognosis in a bi-modal fashion, as SOX11 negativity is mostly encountered in leukemic MCL, which harbours a more protracted course, but is also associated with development of TP53 aberrations, particularly in the terminal phases of the disease and in cases of nodal-type MCL [5,6,14]. Recently, a study on 365 patients enrolled within the European MCL Network (only 1% consisting of leukemic non-nodal MCL) demonstrated superior survival curves in cases harbouring >10% SOX11+ cells, with the poorest outcomes in completely negative cases [7]; however, analysis and correlation with *TP53* imbalances was not performed. From the technical point of view, it should be noted that, besides the high standardization of available clones (overlap was observed in 6/8 laboratories), variability in staining results is acknowledged in the literature [15]. Our data show that SOX11 status can reliably be assessed by IHC: given the uncertain relevance in nodal MCL, either as a prognostic parameter or as a diagnostic tool to identify clinical subgroups, the future application of this biomarker needs to be explored.

Lymphoma cell proliferation is the best established prognostic biomarker in MCL [2]. In our study, Ki67 assessment achieved a very good concordance as a continuous variable among different scorers, with comparable ICC values when the counts were performed centrally and highest agreement when they were performed manually by a single, independent pathologist. On the other hand, the consistency was lower when the scores were translated into the “non-high” vs “high” system, since cases with values close to the cut-off switched groups occasionally. Practically, in concordance with the previously published data [12], exact count of Ki67-positive cells is an essential requisite to assess the Ki67 index. In other words, it is technically required to generate Ki67 as a continuous parameter, but post-hoc transformation of this continuous parameter into prognostic groups leads to a reduction in inter-observer agreement for cases in which the true biological proliferation rate lies in the range of the cut-offs. Thus, our data stress the requirement of counting for Ki67 and the value of the established guidelines [12]. So far it remains uncertain whether the Ki67 index will finally be applied as a continuous variable or in a dichotomized way for pre-treatment stratification of patients.

MCL carries specificities in its histology, which impact the evaluation of the proliferative index. First and most important, the selection of the best area to count requires careful exclusion of foci of non-neoplastic cell proliferation and areas appearing as exceptionally high proliferation of lymphoma cells (so called “hotspots”). Our results are based on the recommendation previously developed by the group, which includes a selection of fields

within homogeneous proliferation free of hot-spots and high levels of non-neoplastic cells. It needs to be considered that our test included TMA and not whole slides [12], thus the issue of proper area selection is not reflecting the real-life diagnostic scenario, a potential insufficiency of our study. However, TMA represents an effective tool for assessment of pathologic parameters [16,17] and our results show that a good degree of concordance in counting can be achieved both by manual and automated approaches, although the need to set optical thresholds for image analysis and the necessity of accurate selection of regions of interest still requires the skill of a trained pathologist [18].

To date, the issue of reproducibility of Ki67 assessment has been extensively addressed, with most of the literature focusing on the pathology of breast cancer and neuroendocrine tumors [19,20]. It is acknowledged that Ki67 scoring can achieve a high inter-rater agreement once proper guidelines are developed [21]. However, besides the advantage of the widespread availability of Ki67 staining in diagnostic pathology labs and the attempts to establish standardized methods for its assessment, IHC remains a technology largely influenced by pre-analytical variability of tissue processing (most important, type and time of fixation) and by the wide array of primary antibodies clones and of staining procedures available on the market [19,22,23]. In our series, 8 participating labs applied 4 different Ki67 clones to 6 different staining systems, a heterogeneity which cannot be reasonably overcome in the daily life setting, as choice of laboratory procedures are bound to local policies. The good reproducibility of the counts, when pairing the original assessment with the central reevaluation of each slide, both manually and by digital image analysis (which still depends on an arbitrary selection of thresholds) further strengthens the point that laboratory standards effectively impact the results, but still allowing an overall good performance of the test.

Collectively, these observations confirm the applicability of Ki67 staining as prognostic factor in clinical trials. A centralized revision of stained slides may be applied for cases with Ki67 values close to the cut-off and/or cases not diagnosed in a specialized reference pathology centre like the ones participating in the current test. Future inter-laboratory testing should thus include the use of digital slides (whole slide scans) which are most likely used for a second opinion scoring within next clinical trials. Finally, our analysis supports the concept of molecular signatures as valid substitutes of current standards, also for assessment of proliferative index, particularly as they improve the standardization and reproducibility of procedures [9].

In summary, we report a good inter-observer agreement of scoring for Ki67 index, expression of p53 and expression of SOX11 in MCL. Given the well-established prognostic relevance of Ki67 and p53 expression in the common nodal form of MCL, these biomarkers qualify for a future application in pre-treatment risk stratification of patients, even though it should be advisable to manage Ki67 as a continuous parameter. Of note, our results are generated by highly specialized pathologist under conditions that do not represent daily diagnostic practice. Finally, we need to stress, that these biomarkers in MCL have not yet been proven to improve outcome of patients with MCL if applied for treatment selection. This will be a major task for future clinical trials.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

the authors thank Dana Germer, Charlotte Botz von Drathen and Reine ZühlkeJenisch for excellent technical support.

Funding: the work was supported in part by the NIH, grant number 1P01CA229100.

References

- Martin P, Ghione P, Dreyling M (2017) Mantle cell lymphoma - Current standards of care and future directions. *Cancer Treat Rev* 58:51–60. doi: 10.1016/j.ctrv.2017.05.008. [PubMed: 28651117]
- Dreyling M, Ferrero S, Vogt N, Klapper W; European Mantle Cell Lymphoma Network (2014) New paradigms in mantle cell lymphoma: is it time to risk-stratify treatment based on the proliferative signature? *Clin Cancer Res* 20(20):5194–206. doi: 10.1158/1078-0432.CCR-14-0836. [PubMed: 25320369]
- Dreyling M, Campo E, Hermine O et al. (2017) Newly diagnosed and relapsed mantle cell lymphoma: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann Oncol* 28(suppl_4):iv62–iv71. doi: 10.1093/annonc/mdx223.
- Hoster E, Rosenwald A, Berger F et al. (2016) Prognostic value of Ki-67 index, cytology, and growth pattern in mantle-cell lymphoma: results from randomized trials of the European Mantle Cell Lymphoma Network. *J Clin Oncol* 34(12):1386–94. doi: 10.1200/JCO.2015.63.8387. [PubMed: 26926679]
- Nordström L, Sembo S, Eden P et al. (2014) SOX11 and TP53 add prognostic information to MIPI in a homogenously treated cohort of mantle cell lymphoma--a Nordic Lymphoma Group study. *Br J Haematol* 166(1):98–108. doi: 10.1111/bjh.12854 [PubMed: 24684350]
- Nygren L, Baumgartner Wennerholm S, Klimkowska M et al. (2012) Prognostic role of SOX11 in a population-based cohort of mantle cell lymphoma. *Blood* 3;119(18):4215–23. doi: 10.1182/blood2011-12-400580. [PubMed: 22431568]
- Aukema SM, Hoster E, Rosenwald A et al. (2018) Expression of TP53 is associated with the outcome of MCL independent of MIPI and Ki-67 in trials of the European MCL Network. *Blood* 131(4):417–420. doi: 10.1182/blood-2017-07-797019. [PubMed: 29196411]
- Scott DW, Abrisqueta P, Wright GW et al. (2017) New molecular assay for the proliferation signature in mantle cell lymphoma applicable to formalin-fixed paraffin-embedded biopsies. *J Clin Oncol* 35(15):1668–1677. doi: 10.1200/JCO.2016.70.7901. [PubMed: 28291392]
- Rauert-Wunderlich H, Mottok A, Scott DW et al. (2019) Validation of the MCL35 gene expression proliferation assay in randomized trials of the European Mantle Cell Lymphoma Network. *Br J Haematol* 184(4):616–624. doi: 10.1111/bjh.15519.
- Clot G, Jares P, Giné E et al. (2018) A gene signature that distinguishes conventional and leukemic nonnodal mantle cell lymphoma helps predict outcome. *Blood* 132(4):413–422. doi: 10.1182/blood2018-03-838136. [PubMed: 29769262]
- Tiemann M, Schrader C, Klapper W et al. (2005) Histopathology, cell proliferation indices and clinical outcome in 304 patients with mantle cell lymphoma (MCL): a clinicopathological study from the European MCL Network. *Br J Haematol* 131(1):29–38. [PubMed: 16173960]
- Klapper W, Hoster E, Determann O et al. (2009) Ki-67 as a prognostic marker in mantle cell lymphoma-consensus guidelines of the pathology panel of the European MCL Network. *J Hematop* 2(2):103–11. doi: 10.1007/s12308-009-0036-x. [PubMed: 19669190]
- Dreyling M, Klapper W, Rule S (2018) Blastoid and pleomorphic mantle cell lymphoma: still a diagnostic and therapeutic challenge! *Blood* 132(26):2722–2729. doi: 10.1182/blood-2017-08737502. [PubMed: 30385481]

14. Federmann B, Frauenfeld L, Pertsch H et al. (2019) Highly sensitive and specific in situ hybridization assay for quantification of SOX11 mRNA in mantle cell lymphoma reveals association of TP53 mutations with negative and low SOX11 expression. *Haematologica* doi: 10.3324/haematol.2019.219543
15. Soldini D, Valera A, Solé C et al. (2014) Assessment of SOX11 expression in routine lymphoma tissue sections: characterization of new monoclonal antibodies for diagnosis of mantle cell lymphoma. *Am J Surg Pathol* 38(1):86–93. doi: 10.1097/PAS.0b013e3182a43996. [PubMed: 24145648]
16. de Jong D, Rosenwald A, Chhanabhai M et al. (2007) Immunohistochemical prognostic markers in diffuse large B-cell lymphoma: validation of tissue microarray as a prerequisite for broad clinical applications--a study from the Lunenburg Lymphoma Biomarker Consortium. *J Clin Oncol*; 25(7):805–12. doi: 10.1200/JCO.2006.09.4490 [PubMed: 17327602]
17. Reinke S, Richter J, Fend F et al. (2018) Round-robin test for the cell-of-origin classification of diffuse large B-cell lymphoma-a feasibility study using full slide staining. *Virchows Arch* 473(3):341349. doi: 10.1007/s00428-018-2367-4.
18. Christgen M, von Ahsen S, Christgen H et al. (2015) The region-of-interest size impacts on Ki67 quantification by computer-assisted image analysis in breast cancer. *Hum Pathol* 46(9):1341–9. doi: 10.1016/j.humpath.2015.05.016. [PubMed: 26206765]
19. Polley MY, Leung SC, McShane LM et al. (2013) An international Ki67 reproducibility study. *J Natl Cancer Inst* 105(24):1897–906. doi: 10.1093/jnci/djt306. [PubMed: 24203987]
20. Klöppel G, La Rosa S (2018) Ki67 labeling index: assessment and prognostic role in gastroenteropancreatic neuroendocrine neoplasms. *Virchows Arch* 472(3):341–349. doi: 10.1007/s00428-017-2258-0. [PubMed: 29134440]
21. Raap M, Ließem S, Rüschoff J et al. (2017) Quality assurance trials for Ki67 assessment in pathology. *Virchows Arch* 471(4):501–508. doi: 10.1007/s00428-017-2142-y. [PubMed: 28497316]
22. Ács B, Kulka J, Kovács KA et al. (2017) Comparison of 5 Ki-67 antibodies regarding reproducibility and capacity to predict prognosis in breast cancer: does the antibody matter? *Hum Pathol* 65:31–40. doi: 10.1016/j.humpath.2017.01.011. [PubMed: 28188752]
23. Mengel M, von Wasielewski R, Wiese B et al. (2002) Inter-laboratory and inter-observer reproducibility of immunohistochemical assessment of the Ki-67 labelling index in a large multicentre trial. *J Pathol* 198(3):292–9. [PubMed: 12375261]

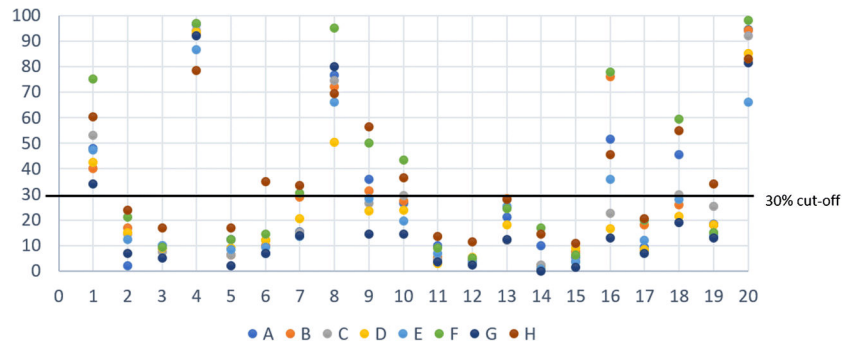


Figure 1. Ki67 scores assessed at individual labs (A-H) are plotted for each case (1–20). The y-axis shows the Ki67 value as percent of positive cells. The 30% cut-off commonly used in previous publications is indicated by a grey line.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

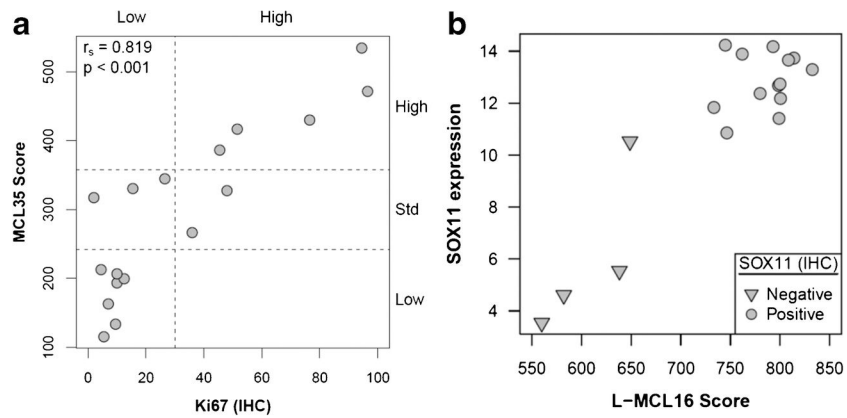


Figure 2.

Ki67 assessment: panels a) and b) depict two matching fields from case 18, with major discrepancy in Ki67 count among different labs (respectively, 59.5% from lab F vs 19% from lab G), reasonably impacted by a diversity in staining intensity and interpretation of weaker signals (centralized manual count was, respectively, 46.5% and 31.5%).

p53 assessment: panels c) and d) exemplify the pattern of positivity among two «p53 high» cases from lab F, respectively, case 10 (*TP53* mutated) and case 20 (*TP53* mutated plus 17p deleted); intensity is variable but consistent, or «clonal», as compared e) with the variable intensity in case 8 (*TP53* wild type, del17p negative), scored intermediate at lab C.

SOX11 assessment: though heterogeneous (f, g; respectively, case 9 and 1, lab C), the consistency of SOX11 staining among each given case allows an easy assessment of SOX11 positivity, as compared to «low positive» (h, case 18, lab C) and true negative cases, which at times harbour only isolated positive cells in an homogeneously negative background.

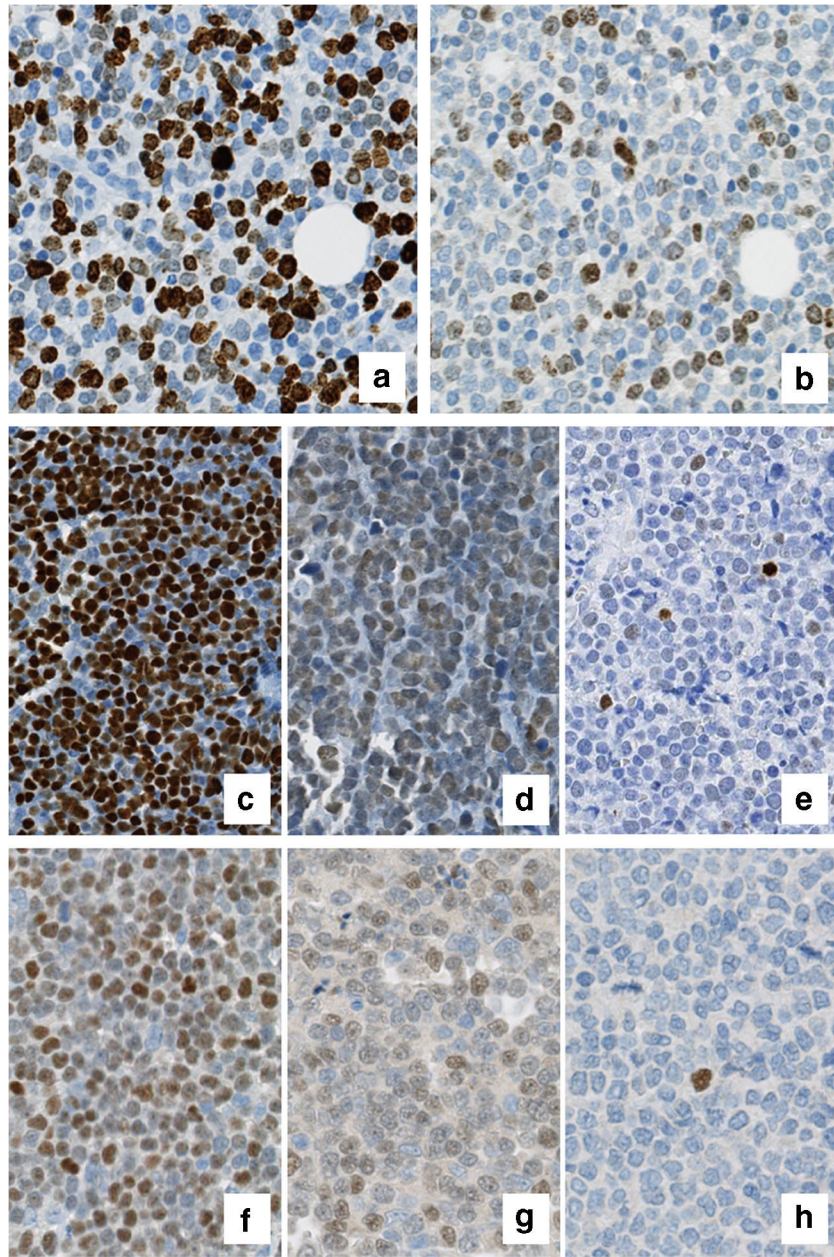


Figure 3.
 a) scatter plot, Ki67 count (Kiel assessment as reference value) vs MCL35 score; b) box plot, SOX11 mRNA by Nanostring vs L-MCL16 score, featuring one case (#20) carrying SOX11 negativity on IHC and a low L-MCL16 score, but with high mRNA level.

Table 1 –

Summary of histologic (Kiel scores, for reference) and molecular assessment

ID	TP53				Proliferation				SOX11			
	Cytology	IHC count	IHC score*	Mutation	del17p FISH	Ki67, IHC	MCL35	Ki67, score	MCL35 group	IHC count	IHC score#	L-MCL16
1	classical	5	1		neg	48	327.6	High	Std	80	2	798.8
2	classical	3	1		neg	2	317.0	Low	Std	40	2	832.9
3	classical	5	1		neg	10	193.9	Low	Low	20	2	793.0
4	pleomorphic	50	2		neg	96.5	472.1	High	High	0	1	582.1
5	pleomorphic	3	1	NE	neg	12.5	199.4	Low	Low	90	2	814.4
6	classical	3	1		neg	7	163.4	Low	Low	80	2	780.0
7	classical	5	1		neg	15.5	330.5	Low	Std	60	2	744.9
8	blastoid	20	1		neg	76.5	430.3	High	High	70	2	800.6
9	classical	5	1		neg	36	267.1	High	Std	90	2	798.9
10	small cell	80	2	p.R273C	neg	26.5	344.8	Low	Std	0	1	559.9
11	classical	2	1		neg	10	206.7	Low	Low	80	2	808.6
12	pleomorphic	2	1		x3-4	4.5	213.1	Low	Low	60	2	733.2
13	classical	5	1		neg	21	NE	Low	NE	80	2	NE
14	classical	5	1	NE	NE	10	NE	Low	NE	80	2	NE
15	classical	5	1		neg	5.5	115.3	Low	Low	80	2	746.5
16	classical	0	1		del	51.5	416.5	High	High	20	2	762.2
17	classical	20	1	NE	neg	9.5	133.8	Low	Low	80	2	799.9
18	pleomorphic	90	2	p.H179	neg	45.5	386.7	High	High	0	1	638.2
19	classical	20	1		neg	18.5	NE	Low	NE	50	2	NE
20	blastoid	90	2	p.R342*	del + x4-5#	94.5	534.8	High	High	0	1	648.8

Legend: IHC, immunohistochemistry; NE, not evaluable; del, deletion.

complex pattern observed, amplified locus with presence of uncoupled centromeres (17p deletion).

* p53 scores: 1-non-high (50%), 2-high (>50%)

SOX11 scores: 1-negative (10%), 2-positive (>50%)