

# Deep learning tertiary lymphoid structures detection on HES/H&E slides and association to survival outcome in sarcoma

ESMO | Oct 20-24, 2024 | EPN #1739P

Lucile Vanhersecke<sup>1</sup>, Yann Rodriguez<sup>2</sup>, Paul Jacob<sup>2</sup>, Lucie Gillet<sup>2</sup>, Jean-Eudes Le Douget<sup>2</sup>, Jean Michel Coindre<sup>3</sup>, Gaetan MacGrogan<sup>4</sup>, Katharina von Loga<sup>5</sup>, Kathryn Schutte<sup>6</sup>, Antoine Italiano<sup>7</sup>

<sup>1</sup>Institut Bergonié, Bordeaux, France; <sup>2</sup>Owkin Inc., New York, USA. Disclosure statement: Lucile Vanhersecke has no conflict of interest to declare.

## Background

- Sarcomas represent a group of rare and heterogeneous cancers presenting significant challenges in diagnosis and treatment.
- Tertiary lymphoid structures (TLS)** are organized T and B immune aggregates which have been associated with improved survival, and mature TLS has emerged as predictive biomarker for immunotherapy response in various cancers, including sarcomas [1].
- Current **TLS detection methods** vary across centers, employing different techniques such as hematoxylin and eosin (H&E)/hematoxylin-eosin-saffron (HES), immunohistochemistry (IHC) or immunofluorescence [2].
- Pathological screening is both laborious and time-consuming, and often requires additional paraffin sections, leading to consumption of tumour material.

## Objectives

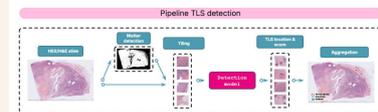
- Develop a **deep-learning model to accurately detect and localize TLS** on H&E/HES-stained whole slide images (WSI) of sarcoma tumors.
- Validate the **prognostic value** of this automatic TLS quantification in sarcomas.

## Data

- Training cohort:**
- PERISARC, sarcomas (N=212 patients), HES slides
    - Patient-level annotations: 162 TLS+, 50 TLS-
    - Local annotations: 2,999 TLS annotations
    - TLS annotated using HES and CD20 staining by IHC.
- Validation cohorts:**
- PEMBROSARC, sarcomas (N=234 patients), HES slides
    - Patient-level annotations: 46 TLS+, 188 TLS-
    - TLS annotated using HES, CD3 and CD20 staining by IHC.
  - TCGA-SARC, sarcomas (N=254 patients), H&E slides
    - Overall survival data.
  - TCGA-PDAC, adenocarcinomas (N=161 patients, 182 slides), H&E slides
    - Slide-level annotations: 128 TLS+, 54 TLS-
    - Local annotations: 1140 TLS annotations
    - TLS annotated using H&E.

## Method

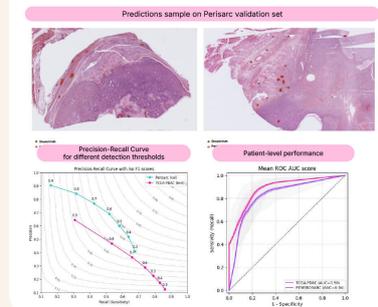
- Input slides are partitioned in smaller tiles of size 1768x1768 px at magnification x10 (~1 μm/pix).
- Our model is an object detection deep learning architecture with **Phikon<sup>®</sup>** backbone (Owkin's self-supervised feature extractor pretrained on 40M slides).
- For each patient, the model detects TLS by providing a bounding box around each structure as well as a confidence score.
- The model is trained on 16k tiles of a 70% split of the patients on the training cohort. We also used positive sampling and histology-specific augmentation methods [3]. The two last layers of the backbone were unfrozen.



## Results: Detection performance

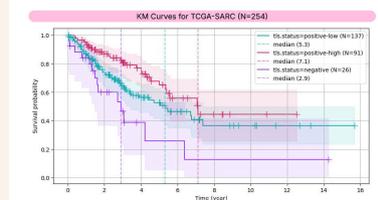
- We evaluate the model performance on local detection and on patient global TLS status prediction, depending on the granularity of annotations of each dataset.
- For local detection performance, we report the average precision (AP) as well as the recall for a threshold of 0.5 on the confidence scores.
- To aggregate local predictions to a patient level TLS status prediction, we select the maximum value of detections confidence scores for each slide.
- The performance reported on TCGA-PDAC shows the model's ability to generalize to other indications than the one seen in training.

Cohort	Local Detection		Patient status prediction	
	Average Precision (AP)	Recall @ 0.5	ROC-AUC	
Perisarc test-set	0.54 [0.484, 0.602]	0.6 [0.571, 0.641]	0.87 [0.832, 0.913]	
PembroSarc	NA	NA	0.84 [0.828, 0.844]	
TCGA-PDAC	0.46 [0.406, 0.524]	0.8 [0.764, 0.854]	0.90 [0.891, 0.901]	



## Results: Prognosis value on external cohort

- To validate the prognosis value of our model on a sarcoma cohort, we tested our model on the TCGA-SARC cohort and performed a survival analysis on the overall survival of patients.
- We defined 3 groups based on the median number of predicted TLS per positive slide:
  - TLS negative (no predicted TLS)
  - TLS positive-low (less than 8 TLS predicted)
  - TLS positive-high (more than 9 TLS predicted)



- Predicted TLS presence on TCGA-SARC is significantly associated with a **better overall survival** (HR=0.392, 95% CI [0.180, 0.854], p=0.016).
- Among positives, TLS-high slides are significantly associated with a better overall survival than TLS-low slides [HR=0.612, 95% CI [0.402, 0.931], p=0.022].

## Conclusions

- Our deep learning model is able to **accurately detect TLS** on H&E/HES stained histology slides of sarcoma patients.
- Our model provides **robust TLS assessment**, enabling more **uniform clinical studies** and potential integration into **routine diagnostics** while **preserving tumor material**.
- Using an external dataset of sarcoma patients, we showed that our model's prediction holds **significant prognostic value** in sarcomas.

## References & acknowledgements

- Petignat, F. et al. B cells are associated with survival and immunotherapy response in sarcoma. *Nature* 572, 558-560 (2020).
- Vanhersecke L et al. Mature tertiary lymphoid structures predict immune checkpoint inhibitor efficacy in solid tumors independently of PD-L1 expression. *Nature cancer* 2, 784-802 (2021).
- Vanhersecke L, et al. Standardized pathology screening of mature tertiary lymphoid structures in cancers. *Laboratory Investigation* 103.5, 10063 (2023).
- Filfil A, et al. Scaling self-supervised learning for histopathology with masked image modeling. *medRxiv* 2023-07 (2023).
- Shen Y, et al. StainStainNet: Learning stain-agnostic features from histology slides by bridging stain augmentation and normalization. *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Cham: Springer Nature Switzerland (2022).