# A Novel Deep Learning Model for Automated Classification of Thymic Epithelial Tumors

Rishi Sharma[1], Matteo Sacco[1], Erica Pietroluongo[1,2], Anna Di Lello[1], Mirella Marino[3], Alessandra Esposito[1], Aliya N. Husain[4], Qudsia Arif[4], James M. Dolezal[5], and Marina Chiara Garassino[1]

[1]Department of Medicine, Section of Hematology/Oncology, University of Chicago, Chicago, IL, USA
[2]Department of Clinical Medicine and Surgery, University of Naples Federico II, Naples, Italy
[3]Department of Pathology, IRCCS Regina Elena National Cancer Institute, Rome, Italy
[4]Department of Pathology, University of Chicago, Chicago, IL, USA
[5]Geisinger Cancer Institute, Danville, PA, USA

**Background:** Thymic epithelial tumors (TETs) subtypes, which include thymoma A, AB, B1, B2, B3, and thymic carcinoma (TC), present major classification challenges due to overlapping histomorphologies of the subtypes and their rarity. Inter-observer variability remains high, with up to 56% of cases reclassified upon second opinion expert review, altering management in 40% of cases. Automated, reproducible classification that uses inexpensive computational resources could dramatically improve diagnostic consistency and patient care, even in resource-limited settings.
**Methods:** We developed a deep learning model trained on all 119 hematoxylin and eosin (H&E)-stained whole slide images (WSIs) of TETs from The Cancer Genome Atlas (TCGA), which provides the subtype label decided by a panel of expert pathologists. Slides were partitioned into 224×224 pixel patches. High-dimensional features from each patch were extracted using UNI, a foundation model pre-trained on ¿100M histopathology images across 20 tissue types. An attention-based multiple instance learning (abMIL) framework aggregated patch-level features for slide-level predictions. The key innovation in our approach is the implementation of a biologically informed hierarchical loss function which incorporates multiclass (group-level: thymoma A vs. thymoma B vs. TC), binary (thymoma A subtypes: A vs. AB), and ordinal (thymoma B subtypes: B1–B3 continuum) tasks. The ordinal nature of the B1-B3 subtypes was encoded using a binary scheme where each class receives one more "1" bit than the previous class (B1: [0,0], B2: [1,0], B3: [1,1]), allowing the model to understand the lymphocyte-to-epithelial cell ratios that define each B subtype. The model was run on standard laptop hardware.
**Results:** Performance was evaluated with a 3-fold cross-validation. The model achieved a subtype-level accuracy of 59.4% (95% CI: 55.3–63.5) and Cohen's kappa of 0.485 (0.437–0.534). Grouped category classification (A/AB vs. B1–B3 vs. TC) reached 81.0% accuracy (76.5–85.5) with a kappa of 0.678 (0.598–0.757). Performance for TC was particularly strong, with 95.8% accuracy (92.5–99.0), 94.4% sensitivity (86.0–100), and 96.0% specificity (93.1–98.9).
**Conclusions:** Our study addresses a longstanding challenge in thymic tumour pathology by introducing an automated, hierarchical deep learning approach for TET classification using digital WSIs. Our results demonstrate the model's ability to reliably distinguish TET subtypes, with particularly strong performance in identifying TC, the most aggressive of the six subtypes. Given the notable inter-observer variability reported in the literature and the scarcity of expert reviewers, our model offers a valuable adjunct to traditional pathology, supporting more consistent and objective diagnoses. This advancement has the potential not only to improve individual patient care but also to facilitate research and standardize practice across healthcare facilities with limited access to pathologists who are experts in this field.