Predicting cell of origin in patients with diffuse large B-cell lymphoma using an explainable feature-based model

Ping-Chang Lin,[1] Nazim Shaikh,[1] Prasanna Porwal,[1] Srinath Jayachandran,[1] Qiangqiang Gu,[1] Xiao Li,[2] Konstanty Korski,[3] Yao Nie[1]

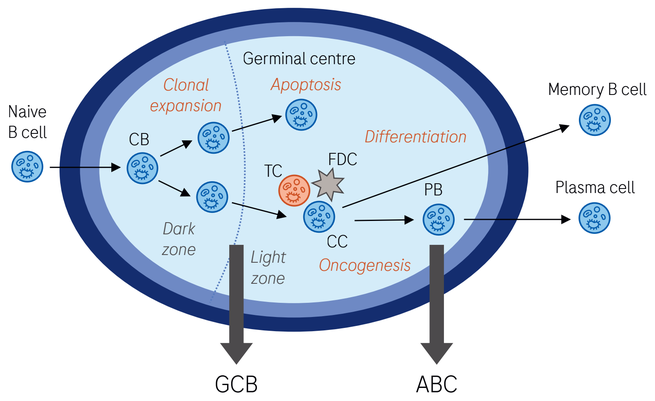
[1]Computational Science and Informatics, Roche Diagnostic Solutions, Pathology Lab Solutions, Santa Clara, CA, USA; [2]Department of Personalized Healthcare, Data, Analytics and Imaging Group, Genentech, Inc., San Francisco, CA, USA; [3]Department of Personalized Healthcare, Data, Analytics and Imaging Group, F. Hoffmann-La Roche Ltd, Basel, Switzerland

Poster 012

Introduction

- Diffuse large B-cell lymphoma (DLBCL) is the most commonly diagnosed form of non-Hodgkin lymphoma and is often characterised by aggressive tumour growth in lymph nodes or extranodal sites.[1]
- DLBCL can be classified by cell of origin (COO) into two principal subtypes: activated B-cell-like (ABC) or germinal centre B-cell-like (GCB) tumours (**Figure 1**).[1] COO classification can have prognostic value because patients with ABC tumours may experience poorer treatment outcomes with rituximab plus cyclophosphamide, doxorubicin, vincristine and prednisone (R-CHOP) immunochemotherapy than those with GCB tumours.[2–4]
- Among current methods for determining COO some can be expensive, time-consuming, weakly reproducible among pathology labs, and may poorly reflect the underlying tumour biology.[4,5]
- Deep-learning models that classify DLBCL by COO using whole-slide images (WSIs) stained with haematoxylin and eosin (H&E) offer an opportunity to automate and standardise COO classification.
- Random forest (RF) models,[6] which perform classifications using a set of simple decision trees, have greater explainability and are less computationally intensive than previously proposed attention-based multiple instance learning (A-MIL) models, which use deep networks.[7]

Figure 1. COO in DLBCL



ABC, activated B-cell-like; CB, centroblast; CC, centrocyte; COO, cell of origin; DLBCL, diffuse large B-cell lymphoma; FDC, follicular dendritic cell; GCB, germinal centre B-cell-like; PB, plasmablast; TC, T cell.

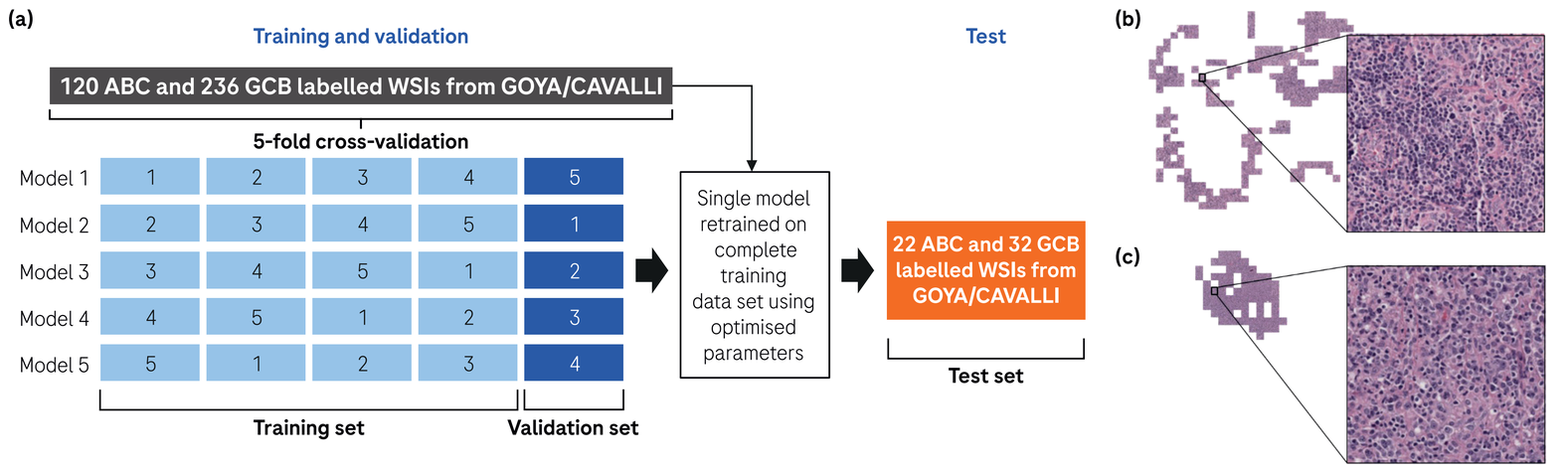
Aim

- To develop an RF model and compare its performance in COO classification versus an A-MIL model, and to evaluate the importance of cellular features that the RF model uses to perform COO classification.

Methods

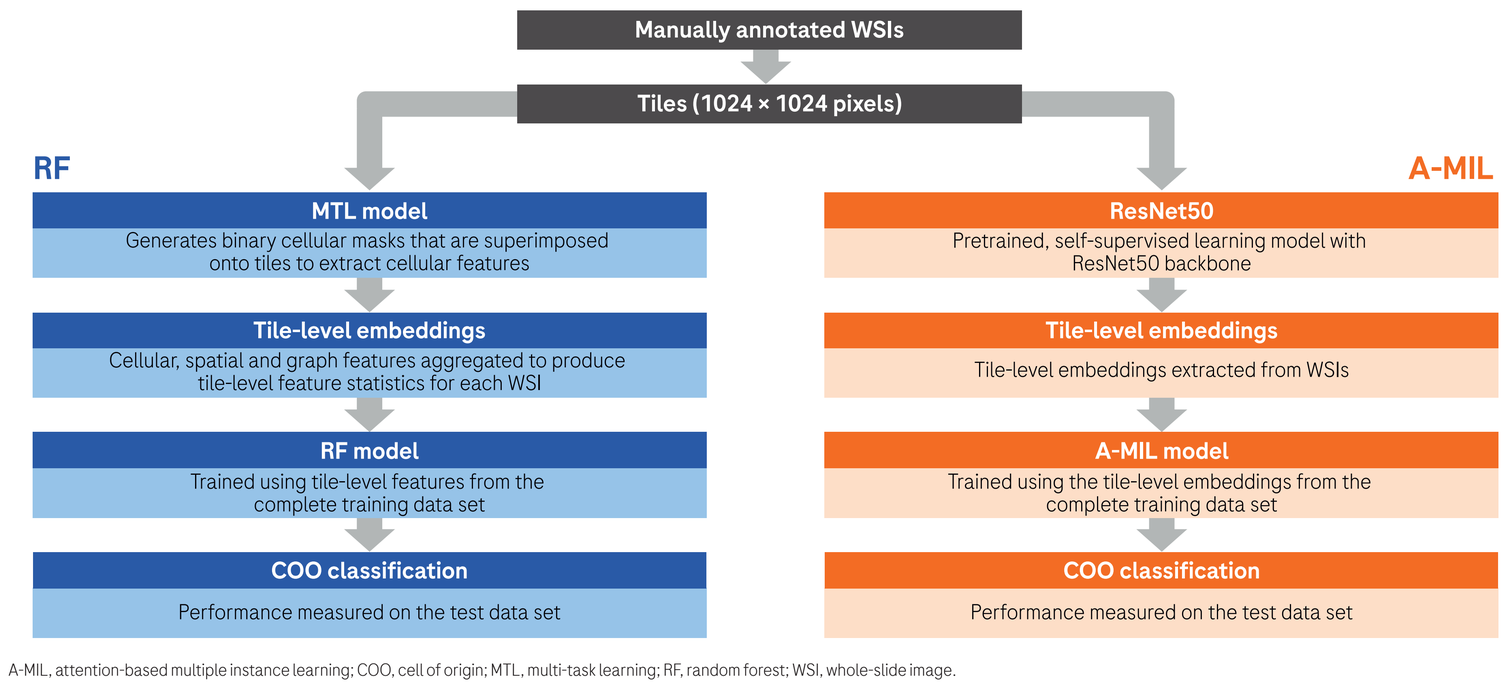
- Algorithms were trained, validated and tested using data from the phase 2 CAVALLI (ClinicalTrials.gov identifier: NCT02055820) and phase 3 GOYA (ClinicalTrials.gov identifier: NCT01287741) trials.[8,9]
- H&E-stained WSIs (40 × magnification) from 410 patients with DLBCL were used. The training set contained 120 ABC-labelled and 236 GCB-labelled WSIs; the test set contained 22 ABC-labelled and 32 GCB-labelled WSIs (**Figure 2**).
- Tumour regions on each WSI were manually annotated and a maximum of 30 tiles (1024 × 1024 pixels) were extracted from annotated regions for each WSI (**Figure 2**).
- Gene expression profiling was used to confirm the ground truth COO classification.
- **RF model**
  - The workflow for RF model training and COO classification is shown in **Figure 3**.
  - Tiles extracted from annotated tumour regions were superimposed with binary cellular masks to extract cellular features.
  - Cell-level features were aggregated to produce tile-level statistical profiles for each WSI.
  - Tile-level feature arrays and WSI-level ground truth COO labels from the training data set were used to train an RF classifier model with 5-fold cross-validation (**Figure 2**).
  - RF hyper-parameters optimised through cross-validation were used to retrain the RF model on the full training data set; model performance was tested on the test data set.
  - Model explainability was assessed by computing the contribution of each cellular feature to the outcome of the COO classification using SHapley Additive exPlanations (SHAP).[10]
- **A-MIL model**
  - The workflow for A-MIL model training and COO classification is shown in **Figure 3**.
  - A pretrained, self-supervised learning model with a ResNet50 backbone was used to generate tile-level embeddings from the same tiles used to train the RF model.
  - COO classification was performed using an A-MIL network to calculate attention weights for each tile and predict the WSI label based on the weighted sum of tile-level predictions. The model was trained on the training data set with 5-fold cross-validation (**Figure 2**).
  - A-MIL hyper-parameters optimised through cross-validation were used to retrain the A-MIL model on the full training data set; model performance was tested on the test data set.
- The performance of the RF and A-MIL models was measured using the area under the receiver operating characteristic (AUROC) curve.

Figure 2. Description of (a) the data sets used to train, validate and test the RF and A-MIL models and representative tiles extracted from H&E-stained WSIs taken from patients with (b) GCB and (c) ABC DLBCL



ABC, activated B-cell-like; A-MIL, attention-based multiple instance learning; DLBCL, diffuse large B-cell lymphoma; GCB, germinal centre B-cell-like; H&E, haematoxylin and eosin; RF, random forest; WSI, whole-slide image.

Figure 3. Workflows for image feature extraction, model training and performance testing for the RF and A-MIL models



A-MIL, attention-based multiple instance learning; COO, cell of origin; MTL, multi-task learning; RF, random forest; WSI, whole-slide image.

Results

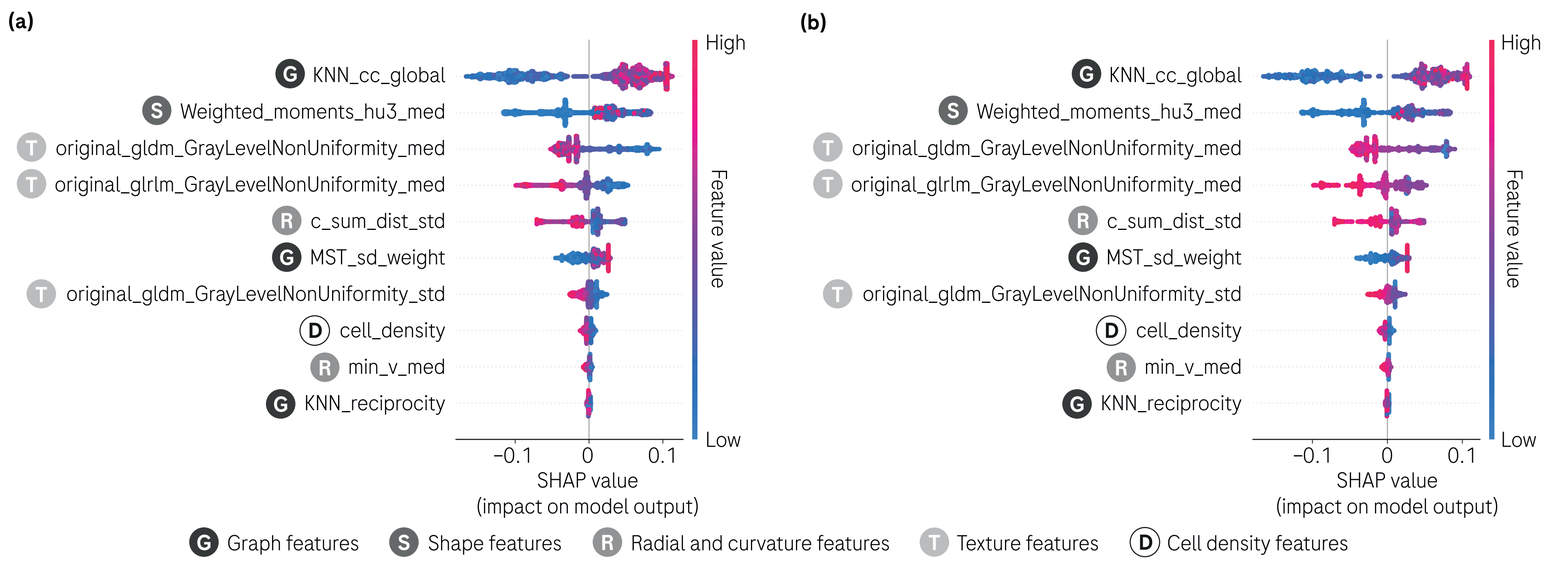
- The COO classification performance of the RF and A-MIL models is shown in **Table 1**. In the validation and test data sets, the A-MIL model had slightly better performance than the RF model.
- SHAP analysis of the RF model performance on the training and test sets revealed the 10 cellular features that had the greatest effect on COO classification (**Figure 4**). These included: graph features that characterised tumour cell spatial distribution; shape features that characterised the nucleus shape; radial and curvature features that characterised tumour cell nuclear boundaries; texture features that characterised tumour cell chromatin pattern; and cell density features.

Table 1. COO classification performance for RF and A-MIL models

Model type	Training set[a]	Validation set[a]	Test set[b]
	AUROC, mean ± SD	AUROC, mean ± SD	AUROC
RF model	0.771 ± 0.004	0.675 ± 0.045	0.715
A-MIL model	0.713 ± 0.020	0.687 ± 0.026	0.737

[a]Cross-validated mean and SD values are shown for the training and validation data sets. [b]Performance of the single optimised model is shown for the test data set.
A-MIL, attention-based multiple instance learning; AUROC, area under the receiver operating characteristic curve; COO, cell of origin; RF, random forest; SD, standard deviation.

Figure 4. SHAP analysis of the RF model for (a) the training data set and (b) the test data set



RF, random forest; SHAP, SHapley Additive exPlanations.

Conclusions

- **Using H&E-stained WSIs from patients with DLBCL, an RF model achieved similar COO classification performance to that of an A-MIL model.**
- **In contrast to A-MIL models that are explainable by locating high-attention regions in WSIs, the RF model was able to identify specific cellular features that have a high impact on the output of the COO classification.**
- **The RF model provides insightful information that may contribute to better understanding of disease biology in DLBCL and improve model credibility.**

References

1. Susanibar-Adaniya S, Barta SK. *Am J Hematol* 2021;96:617–29. **2.** Scott DW *et al. J Clin Oncol* 2015;33:2848–56. **3.** Rosenwald A *et al. N Engl J Med* 2002;346:1937–47. **4.** Hunter E *et al. Transl Med Commun* 2020;5:5. **5.** de Jong D *et al. J Clin Oncol* 2007;25:805–12. **6.** Pedregosa F *et al. J Mach Learn Res* 2011;12:2825–30. **7.** Ilse M *et al.* Attention-based deep multiple instance learning. arXiv:1802.04712 [Preprint; 16 pp.] 2018. Available from: https://arxiv.org/abs/1802.04712 (Accessed 14 August 2023). **8.** Morschhauser F *et al. Blood* 2021;137:600–9. **9.** Sehn LH *et al. J Hematol Oncol* 2020;13:71. **10.** Lundberg SM *et al. Nat Mach Intell* 2020;2:56–67.

Acknowledgements

This study was funded by F. Hoffmann-La Roche Ltd, Basel, Switzerland. Medical writing support was provided by Adam Errington PhD of PharmaGenesis Cardiff, Cardiff, UK and funded by F. Hoffmann-La Roche Ltd in accordance with Good Publication Practice (GPP 2022) guidelines (www.ismpp.org/gpp-2022).

Disclosures

PL, NS, PP, SJ, QG, XL, KK and YN are employees of, have received research funding and expenses from, and have intellectual property interests with F. Hoffmann-La Roche AG. NS, SJ, QG, XL, KK and YN hold stock in F. Hoffmann-La Roche AG.

Presented at the 35th European Congress on Pathology (ECP), 9–13 September 2023, Dublin, Ireland