Classification of Aggressive B-Cell Lymphomas Using Gene Expression Profiling

James Godwin[1], Sharon Barrans[2], John Davies[3], Simon Crouch[1], Daniel Painter[1], Alex Smith[1], Cathy Burton[2], Eve Roman[1] and David Westhead[3]

BACKGROUND

- Lymphoma classification has expanded; requiring the integration of tissue morphology, immunohistochemical markers and genetic features by an expert haemopathologist to form a diagnosis.
- This process becomes challenging when patients display a range of features from different subtypes – so called 'edge cases'. It is hoped gene expression can improve the diagnostic process.
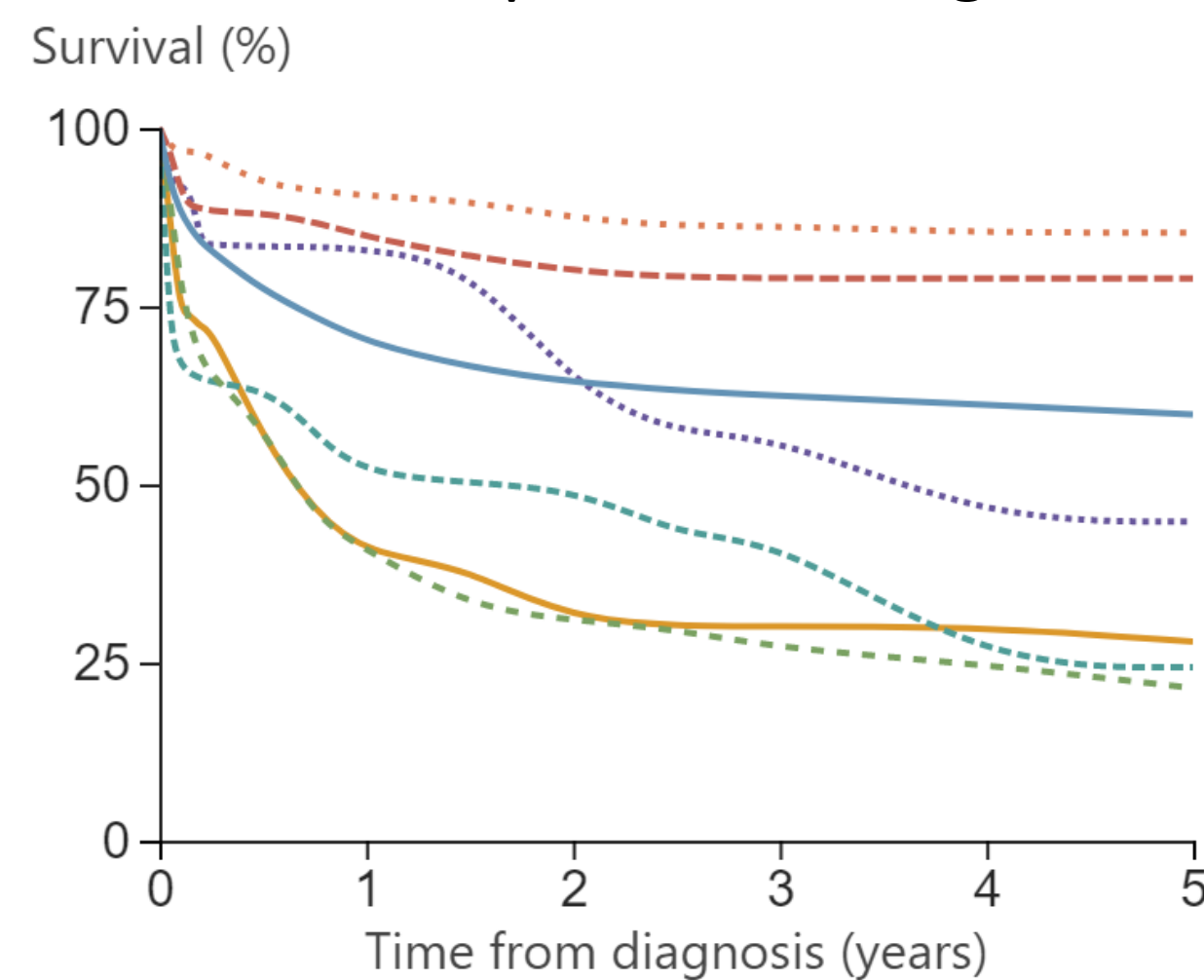

Figure 1: Relative 5-year survival for large B-cell lymphoma subtypes (hmrn.org)

Survival (%) / Time from diagnosis (years)

- Diffuse large B-cell lymphoma (DLBCL), NOS
- T-cell/histiocyte-rich large B-cell lymphoma
- Primary diffuse large B-cell lymphoma of the CNS
- Primary cutaneous DLBCL, leg type
- Primary mediastinal large B-cell lymphoma
- Intravascular large B-cell lymphoma
- Plasmablastic lymphoma

OBJECTIVES

- Investigate:
  - ❖ a 'pan aggressive B-cell lymphoma' classifier derived from a single gene expression panel
  - ❖ its accuracy against conventional diagnostic tools
  - ❖ its potential utility for classifying those challenging 'edge cases'

REFERENCES

Chapman J, Gentles AJ, Sujoy V, et al. Gene expression analysis of plasmablastic lymphoma identifies downregulation of B-cell receptor signaling and additional unique transcriptional programs. Leukemia. 2015;29(11):2270-2273. doi:10.1038/leu.2015.109
Dave et al (2006) Molecular Diagnosis of Burkitt's Lymphoma N Engl J Med 2006; 354:2431-2442
Ennishi D, Jiang A, Boyle M, et al. Double-Hit Gene Expression Signature Defines a Distinct Subgroup of Germinal Center B-Cell-Like Diffuse Large BCell Lymphoma. J Clin Oncol. 2019;37(3):190-201. doi:10.1200/JCO.18.01583
Hummel M, Bentink S, Berger H, et al. A biologic definition of Burkitt's lymphoma from transcriptional and genomic profiling. N Engl J Med. 2006;354(23):2419-2430. doi:10.1056/NEJMoa055351
Mottok A, Wright G, Rosenwald A, et al. Molecular classification of primary mediastinal large B-cell lymphoma using routinely available tissue specimens. Blood. 2018;132(22):2401-2405. doi:10.1182/blood-2018-05-851154
Painter, D., Barrans, S., Lacy, S., et al. (2019), Cell-of-origin in diffuse large B-cell lymphoma: findings from the UK's population-based Haematological Malignancy Research Network. Br J Haematol, 185: 781-784. doi:10.1111/bjh.15619
Rizzatti EG, Falcão RP, Panepucci RA, et al. Gene expression profiling of mantle cell lymphoma cells reveals aberrant expression of genes from the PI3K-AKT, WNT and TGFbeta signalling pathways. Br J Haematol. 2005;130(4):516-526. doi:10.1111/j.1365-2141.2005.05630.x
Salaverria I, Martin-Guerrero I, Wagener R, et al. A recurrent 11q aberration pattern characterizes a subset of MYC-negative high-grade B-cell lymphomas resembling Burkitt lymphoma. Blood. 2014;123(8):1187-1198. doi:10.1182/blood-2013-06-507996
Sha C, Barrans S, Cucco F, et al. Molecular High-Grade B-Cell Lymphoma: Defining a Poor-Risk Group That Requires Different Approaches to Therapy [published correction appears in J Clin Oncol. 2019 Apr 20;37(12):1035]. J Clin Oncol. 2019;37(3):202-212. doi:10.1200/JCO.18.01314
Wright G, Tan B, Rosenwald A, Hurt EH, Wiestner A, Staudt LM. A gene expression-based method to diagnose clinically distinct subgroups of diffuse large B cell lymphoma. Proc Natl Acad Sci U S A. 2003;100(17):9991-9996. doi:10.1073/pnas.1732008100

RESULTS

- Broad classification can discriminate between a variety of lymphomas with accuracy – including rare subtypes not commonly presented
- When no threshold is specified, 85-95% of samples are classified correctly when unseen.
- When a threshold is introduced, for example an 80% threshold, 92-97% of samples are classified correctly at the expense of 18-22% of samples not receiving a classification.
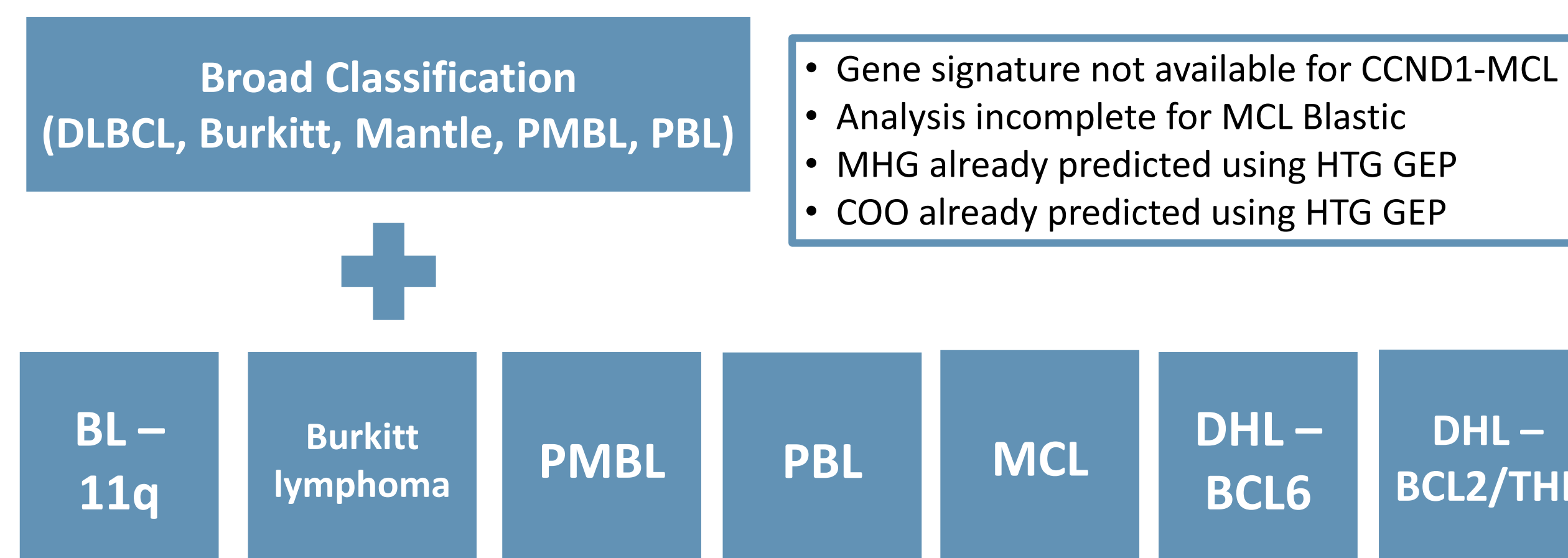
Broad Classification (DLBCL, Burkitt, Mantle, PMBL, PBL)

- Gene signature not available for CCND1-MCL
- Analysis incomplete for MCL Blastic
- MHG already predicted using HTG GEP
- COO already predicted using HTG GEP

BL – 11q | Burkitt lymphoma | PMBL | PBL | MCL | DHL – BCL6 | DHL – BCL2/THL

Figure 2: Pan aggressive B-cell lymphoma classifier components

Linear Predictor Score	Sensitivity	Specificity
Burkitt Like 11q	89% (77-100)	92% (90-95)
Burkitt lymphoma	89% (75-100)	90% (88-95)
Mantle Cell lymphoma	90% (88-100)	90% (88-92)
Primary Mediastinal B-cell lymphoma	82% (68-90)	95% (92-97)
Plasmablastic B-cell lymphoma	100% (90-100)	95% (92-97)
Double Hit lymphoma – BCL6	77% (60-86)	55% (52-60)
Double Hit lymphoma – BCL2/THL	86% (77-100)	68% (63-72)

Table 1 : Linear predictor score results: median accuracy % (LQ-UQ)

- From the LPS (Table 1): MCL, PBL, Burkitt lymphoma and BL-11q are well predicted. DHLs with BCL2/THL rearrangement were better predicted than DHL-BCL6.
- Classifications for 14/16 edge cases were given, when a minimum probability threshold of 50% was specified. All 16 cases' LPS are detailed in heatmap below.
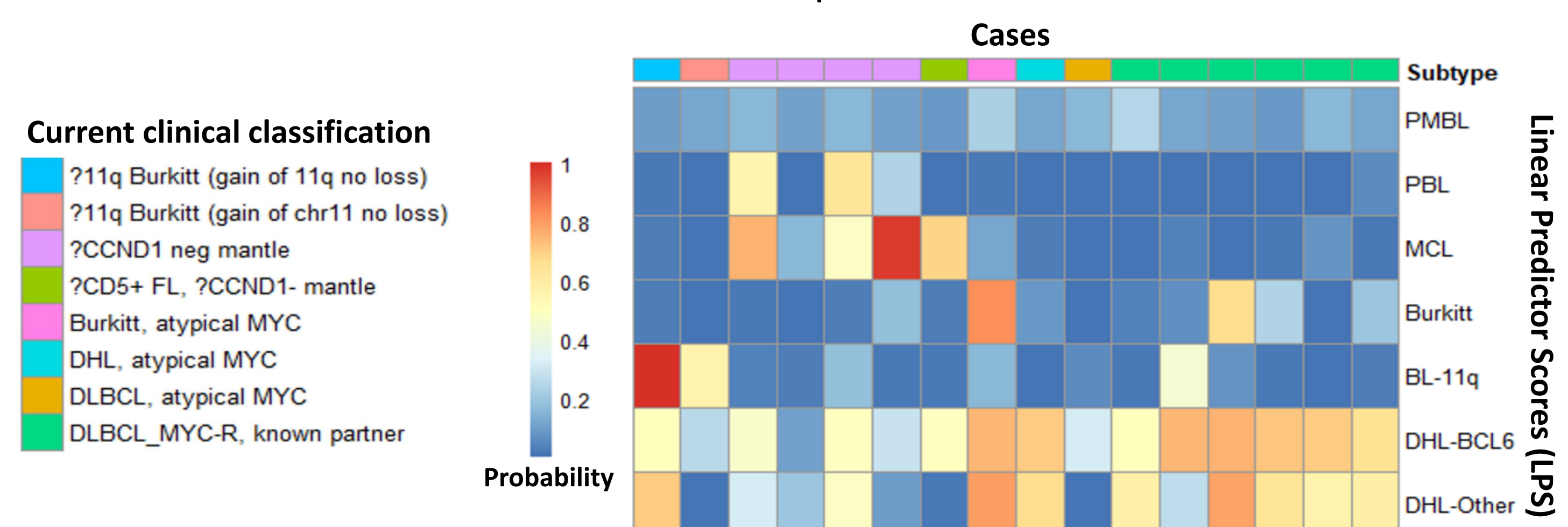

Current clinical classification
- ?11q Burkitt (gain of 11q no loss)
- ?11q Burkitt (gain of chr11 no loss)
- ?CCND1 neg mantle
- ?CD5+ FL, ?CCND1- mantle
- Burkitt, atypical MYC
- DHL, atypical MYC
- DLBCL, atypical MYC
- DLBCL_MYC-R, known partner

Figure 3: Linear predictor score classification of 'edge cases'

CONCLUSION

- A 'PanB' GEP classification algorithm can:
  - ❖ aid diagnosis of complex edge cases
  - ❖ provide a mechanism for triage of new cases
  - ❖ stratify patients in clinical trials

METHODS

- The HTG EdgeSeq Pan B-Cell Lymphoma Panel measured the expression of 298 genes, selected from published gene signatures for subtypes
- To train and validate the classifiers, 477 samples were used (Table 2). The majority came from the Haematological Malignancy Research Network (HMRN), a UK population-based patient cohort (www.hmrn.org). The remaining samples came from clinical trials and external research groups/hospitals.
- Formalin-Fixed Paraffin-Embedded (FFPE) samples were processed at the Haematological Malignancy Diagnostics Service in Leeds, UK.

Subtype	N
Diffuse Large B-Cell Lymphoma (DLBCL)	193
Molecular High Grade (MHG)	40
Double Hit Lymphoma (DHL)	42
MHG/DHL	18
Primary Mediastinal B-Cell Lymphoma (PMBL)	28

Subtype	N
Burkitt Lymphoma	28
Burkitt-like 11q (BL -11q)	23
Mantle Cell Lymphoma (MCL)	33
MCL Blastic	35
CCND1-neg MCL	12
Plasmablastic B-cell Lymphoma (PBL)	29

Table 2: Diagnostic makeup of cases used in the training and validation of algorithms

- A two-stage classification algorithm was developed; a 'broad' classification using machine learning and a subtype-specific classification using gene signatures.

- The broad classification classified samples into the pooled subtypes (Table 3). This used a Support Vector Machine (SVM) (Figure 4). Minority classes were oversampled using Synthetic Minority Oversampling Technique (SMOTE).
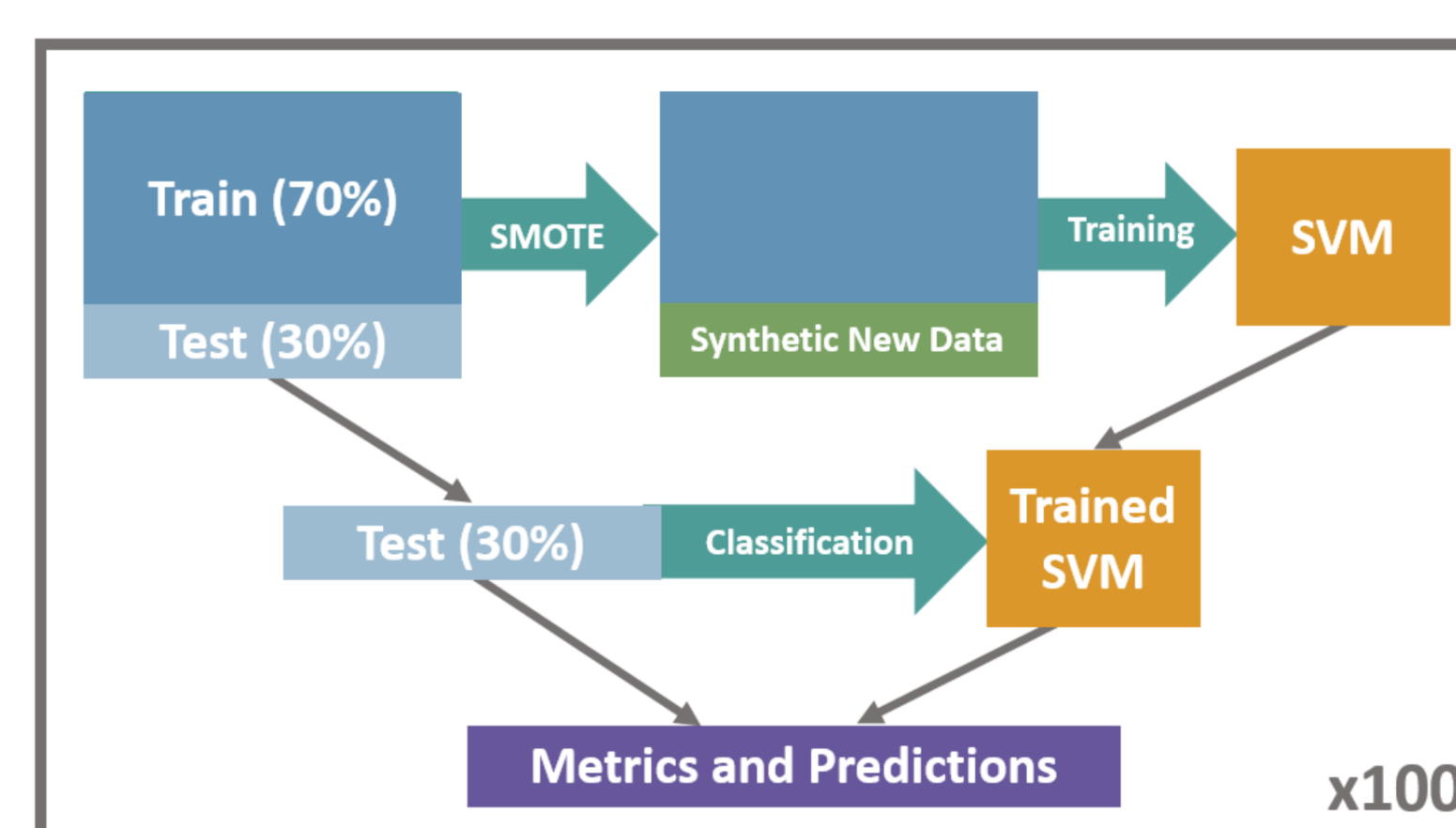

Figure 4: Schematic of broad classifier algorithm

Train (70%) → SMOTE → Training → SVM / Synthetic New Data
Test (30%) → Classification → Trained SVM → Metrics and Predictions x100

Pooled Subtype	Containing
DLBCL	DLBCL, MHG, DHL, MHG/DHL
Mantle	MCL, MCL Blastic, CCND1-neg MCL
Burkitt	Burkitt lymphoma, BL 11q
PMBL	PMBL
PBL	PBL

Table 3: Pooled classes for broad classification

- For uncertainty, user-defined thresholds can be introduced so classification is only given when the probability of the respective classification is above the threshold.
- For subtype specific prediction, linear predictor scores (LPS) were developed (Figure 3) using Bayes' rule which takes published gene signatures of subtypes and predicts a sample's subtype based on their expression of the signatures.

[1] Epidemiology & Cancer Statistics Group, Department of Health Sciences, University of York, York, UK.
[2] Haematological Malignancy Diagnostics Service, Leeds Cancer Centre, St James' University Hospital, Leeds, UK.
[3] Bioinformatics Group, Department of Biological Sciences, University of Leeds, Leeds, UK.

ECSG hmrn HMDS Haematological Malignancy Diagnostic Service CANCER RESEARCH UK Blood Cancer UK