Clinical Epidemiology of T-cell Acute Lymphoblastic Leukaemia

Ozcan Gulbey

A thesis submitted in part requirement for the degree of Doctor of Philosophy

Leukaemia Research Cytogenetics Group,
Wolfson Childhood Cancer Research Centre,
Newcastle University Centre for Cancer,
Translational and Clinical Research Institute,
Faculty of Medical Sciences,
Newcastle University

January 2025

Abstract

T-cell acute lymphoblastic leukaemia (T-cell ALL) is a rare and aggressive disease in children and adults. Although overall survival rates reach ~85% with contemporary treatment and stratification, relapse occurs in ~20% of cases, and these patients still have poor survival. Demographic, clinical, and genetic risk and prognostic factors have been proposed, but they are rarely validated. Minimal/measurable residual disease (MRD) is widely used for risk stratification; however, it is not helpful with relapsed patients and has limited accuracy. Therefore, there is a need to identify additional prognostic factors in T-cell ALL.

This thesis aimed to investigate clinical factors (sex, age, white blood cell count at diagnosis, central nervous system involvement, organomegalies, early marrow response, MRD response, and genetics) in paediatric/young adult UKALL trials (UKALL VIII to UKALL 2011) for the improvement of current MRD stratification.

Treatment response variables (early marrow response and MRD response) were the only significant prognostic factors among other variables in T-cell ALL. D8BM% model, containing only bone marrow percentage measured at day 8, was effective in separating risk groups at very early stage of treatment in UKALL 97/99, UKALL 2003, and UKALL 2011. More importantly, this model improved MRD prediction for clinical events. Dynamic prognostic models with Nonlinear Mixed-Effects (NLME) and Area Under the Curve (AUC) methods were successfully applied to data with treatment response variables at multiple time points; however, they did not have a superior advantage over single-time point measurements. Additionally, T-cell ALL subtypes detected by laboratory techniques and unsupervised clustering algorithm did not have prognostic importance.

In summary, treatment response variables are crucial to stratifying T-cell ALL patients other than genetic abnormalities/subtypes. Early treatment response adds prognostic value to MRD stratification at the end of induction, so it can be efficient to stratify T-cell ALL patients at very early stage of treatment.

Acknowledgments

Firstly, I would like to thank my sponsor, which is the Ministry of National Education-Republic of Türkiye (International Graduate Education Scholarship (YLSY)), for giving me this great chance to do my PhD in the UK.

Additionally, I would most like to thank my supervisors, Prof Anthony V. Moorman and Dr Frederik W. van Delft, as well as Dr Amir Enshaei, for their invaluable academic guidance and moral support during my PhD project.

I would also like to thank Dr Debbie Hicks and Dr Simon Bomken for helpful comments in the annual progress report (APR) meetings.

I also wish to thank my friends/colleagues Melvin Joy, Diyanath S. Ranasinghe, and Dr Ruth Cranston for their help with statistical tests and coding in bioinformatics.

Moreover, special thanks to my friends Mankaran Singh, Shun Chang, and Jude Gibson for their amazing friendships.

Finally, many thanks to my family, as I could never be successful in finishing my PhD without the unconditional love and endless support from my parents, İlknur/Ercan Gülbey, and my brother, Sercan Gülbey.

“ To my mother,

who has not been given an opportunity in education

to show her real potential and value in the society... ”

List of Abbreviations

ALL	Acute lymphoblastic leukaemia
AUC	Area Under the Curve
AUROC	Area Under the Receiver Operating Characteristic
B-cell ALL	B-cell acute lymphoblastic leukaemia
BM	Bone marrow
CD	Cluster of differentiation
CDF	Cumulative distribution function
CI	Confidence interval
CIMS	Cytogenetics Information Management System
C-index	Concordance index
CNS	Central nervous system
COG	Children's Oncology Group
D8BM%	Day 8 bone marrow percentage
DN	Double negative
DNA	Deoxyribonucleic acid
DP	Double positive
EFS	Event-free survival
EGIL	European Group for the Immunological Characterisation of Leukaemias
EOC	End of consolidation
EOI	End of induction
ETP	Early T-cell precursor
FAB	French-American-British
GRAALL	Group for Research on Adult Acute Lymphoblastic Leukemia
HDMTX	High-dose methotrexate
HR	High-risk
HSC	Haemopoietic stem cell
INDEL	Insertion-deletion
ITMTX	Intrathecal methotrexate
KNN	K^{th}-Nearest Neighbours
LASSO	Least Absolute Shrinkage and Selection Operator

LLP	Liverpool Lung Project
LRCG	Leukaemia Research Cytogenetics Group
MAD	Median absolute deviation
MAR	Missing at random
MCAR	Missing completely at random
MNAR	Missing not at random
MRD	Minimal/measurable residual disease
NCI-Rome	National Cancer Institute-Rome
NHS	National Health Service
NK	Natural killer
NLME	Nonlinear Mixed-Effects
OR	Odds ratio
OS	Overall survival
PC1	The first principal component
PC2	The second principal component
PCA	Principal component analysis
RER	Rapid early response
RFS	Relapse-free survival
RNA	Ribonucleic acid
RNA-seq	Ribonucleic acid-sequencing
RR	Relapse rate
SCT	Stem cell transplantation
SER	Slow early response
SR	Standard-risk
T-cell ALL	T-cell acute lymphoblastic leukaemia
TCR	T-cell receptor
UKALL	UK Acute Lymphoblastic Leukaemia
vHR	Very high-risk
WBC	White blood cell count
WHO	World Health Organisation
WTS	Whole transcriptome sequencing

Table of Contents

Abstract i

Acknowledgments ii

List of Abbreviations iv

Table of Contents vi

List of Figures xi

List of Tables xv

Chapter 1. Introduction 1

1.1. Haematopoiesis 2

1.2. Overview of Acute Lymphoblastic Leukaemia (ALL) 4

1.2.1. Epidemiology and Aetiology 4

1.2.2. Diagnosis 4

1.2.3. Classification 5

1.2.4. Genetics 7

1.2.5. Treatment 7

1.2.6. Clinical Outcome 8

1.3. Overview of T-cell Acute Lymphoblastic Leukaemia 10

1.3.1. Epidemiology 10

1.3.2. Classification 10

1.3.3. Cytogenetics/Genomics of T-cell ALL 11

1.3.3.1. Type-A Abnormalities 11

1.3.3.2. Type-B Abnormalities 14

1.3.4. Treatment 16

1.3.5. Clinical Outcome 16

1.4. Evolution of Risk Stratifications in Acute Lymphoblastic Leukaemia 18

1.4.1. Immunophenotype 18

1.4.2. National Cancer Institute-Rome (NCI-Rome) Criteria 18

1.4.3. Early Treatment Response 19
1.4.3.1. Marrow/Blood Response 19
1.4.3.2. Minimal/Measurable Residual Disease 20
1.4.4. Genetics 21
1.4.4.1. Genetic Risk Classifications in T-cell ALL 22
1.4.4.2. Novel Prognostic Models Developed and Validated for ALL 23
1.5. Prognostic Models 24
1.5.1. Overview of Methods Commonly Used in Medicine 24
1.5.1.1. Logistic Regression 24
1.5.1.2. Proportional Hazards Regression (Cox Regression) 25
1.5.1.3. Classification Tree 25
1.5.2. Examples of Prognostic Models in Cancer 26
1.6. Aims and Objectives 28
Chapter 2. Data and Methods 29
2.1. Data Management 30
2.1.1. Data Collection and Manipulation 30
2.1.2. Information about the UKALL Trials 32
2.1.3. Manipulation of Genetic Data 35
2.2. Data Analysis 37
2.2.1. Methodology of Model Development and Validation 37
2.2.1.1. Transformation of Continuous Variables 37
2.2.1.2. Variable Selection Strategy 37
2.2.1.3. Developing Prognostic Indexes and Identifying Optimal Thresholds 38
2.2.1.4. Model Validation, Calibration and Evaluation Metrics 39
2.2.1.5. Missing Values and Imputation Methods 39
2.3. Bulk RNA-seq Analysis 41
2.4. Comparison Tests and Clinical Endpoints 42

Chapter 3. Evaluation of Clinical Risk and Prognostic Factors in Paediatric/Young Adult UKALL Trials 43
3.1. Introduction 44
3.2. Data and Methods 46
3.2.1. Definition of Subgroups 46
3.2.2. Statistical Analysis 47
3.3. Results 48
3.3.1. Overview of Variables in Paediatric and Young Adult UKALL Trials 48
3.3.2. Clinical Outcomes of Patients in UKALL Trials 48
3.3.3. Prognostic Values of Traditional Variables in T-cell ALL 51
3.3.4. Genetic Subtypes and Their Prognostic Impacts on T-cell ALL 53
3.4. Discussion 59
Chapter 4. Integration of Early Treatment Response Variables into Risk Stratification in T-cell Acute Lymphoblastic Leukaemia 62
4.1. Introduction 63
4.2. Data and Methods 64
4.3. Results 66
4.3.1. Distribution and Handling of Missing Values in UKALL X+XI and UKALL 97/99 66
4.3.2. Model Development and Validation Using UKALL X+XI and UKALL 97/99 70
4.3.2.1. Identification of Candidate Prognostic Variables in UKALL X+XI 70
4.3.2.2. Developing and Validating Prognostic Models in UKALL X+XI and UKALL 97/99 71
4.3.3. Model Development and Validation in UKALL 97/99, UKALL 2003 and UKALL 2011 75
4.3.3.1. Identification of Candidate Predictors in UKALL 97/99 76
4.3.3.2. Developing Prognostic Models in UKALL 97/99 77
4.3.3.3. Identifying Risk Groups Using Optimal Threshold 80

4.3.3.4. Clinical Characteristics and Outcomes of the Risk Groups Defined by the D8BM% Model in UKALL 97/99, UKALL 2003 and UKALL 2011 81
4.3.3.5. Model Calibration in UKALL 2003 and UKALL 2011 86
4.3.4. Genetic Abnormalities within the Risk Groups in UKALL 97/99 and UKALL 2003 86
4.3.5. The D8BM% Model and MRD Interaction in UKALL 2003 and UKALL 2011 ... 88
4.3.6. Robustness of the D8BM% Model in Different Patients Subgroups 94
4.3.7. The D8BM% Model and Traditional Marrow Classification 95
4.3.8. The D8BM% Model in TARGET-ALL Dataset 97
4.4. Discussion 99
Chapter 5. Dynamic Prognostic Models for Risk Stratification in T-cell Acute Lymphoblastic Leukaemia 103
5.1. Introduction 104
5.2. Data and Methods 105
5.3. Results 108
5.3.1. Development and Validation of the Dynamic Models for Early Response in UKALL 2003 and UKALL 2011 108
5.3.1.1. Distribution Patterns of the NLME and AUC Trend Scores 108
5.3.1.2. Accuracy of Predicted Values from the NLME Model 108
5.3.1.3. Identification of Optimal Thresholds for the Dynamic Models in UKALL 2003 110
5.3.1.4. Clinical Characteristics and Outcomes of the Risk Groups in UKALL 2003 111
5.3.1.5. Clinical Characteristics and Outcomes of the Risk Groups in UKALL 2011 115
5.3.1.6. Comparison of the Dynamic Models with the D8BM% Model and EOI MRD 119
5.3.2. Dynamics of EOI-EOC MRD Values in UKALL 2011 122
5.4. Discussion 127
Chapter 6. Investigation of T-cell Acute Lymphoblastic Leukaemia Genetics Using Bioinformatics 130

6.1. Introduction 131
6.2. Data and Methods 135
6.2.1. Processing FASTQ Files 135
6.2.2. Genetic Classifier (TALLSorts) 135
6.2.3. Differential Expression Analysis (DESeq2) 136
6.2.4. Unsupervised Clustering Analysis (ConsensusClusterPlus) 138
6.3. Results 139
6.3.1. Comparing Tested and Non-tested Cohorts for RNA-seq Analyses 139
6.3.2. The Genetic Classifier (TALLSorts) 140
6.3.2.1. The Characterised Genetic Subgroup 143
6.3.2.2. The Uncharacterised Genetic Subgroup 146
6.3.2.3. Distribution of the Predicted Subtypes within the Risk Groups 147
6.3.3. Identification of Differentially Expressed Genes between the Risk Groups 148
6.3.4. The Genes Overlapping between Two Different Strategies 151
6.3.5. Unsupervised Clustering Analysis 153
6.3.5.1. Identification of Optimal Cluster Number 153
6.3.5.2. Distribution of Clinical Characteristics, Risk Groups and Predicted Subtypes in the Clusters 155
6.4. Discussion 157
Chapter 7. General Discussion 159
7.1. Summary of Findings and Discussion 160
7.2. Strengths and Limitations 162
7.3. Options for Expanding This Work 162
Chapter 8. References 164

List of Figures

Figure 1.1. Mature blood cells stemming from two different lineages originated from stem cells 2
Figure 1.2. Differentiation stages of the T cells in different parts of the body 3
Figure 1.3. The incidence rate of ALL from 1993 to 2018 in the UK 4
Figure 1.4. Overall view of treatment phases in T-cell ALL 8
Figure 1.5. Four broader subgroups of T-cell ALL based on differentiation stages and genetic/cytogenetic abnormalities 12
Figure 1.6. Simplified overview of treatment stratification based on NCI-Rome classification and early marrow response in UKALL 99, UKALL 2003 and UKALL 2011 20
Figure 1.7. MRD-based risk stratification on AIEOP-BFM-ALL 2000 study 21
Figure 1.8. The formula and function of logistic regression 25
Figure 1.9. Cox regression function 25
Figure 1.10. Simplified steps of classification tree algorithm 26
Figure 2.1. Overall steps for data check and data manipulation in the UKALL datasets 32
Figure 2.2. Treatment phases and randomisations of UKALL VIII, UKALL X, UKALL XI, UKALL 97/99, UKALL 2003 and UKALL 2011 from 1980 to 2018 35
Figure 2.3. Classification of patients according to the presence/absence of known genetic abnormalities detected by different laboratory methods 36
Figure 2.4. Backward and forward selections with criteria/stopping rule 38
Figure 3.1. Kaplan-Meier graphs for event-free survival, relapse, and overall survival rates at five years in paediatric and young adult T-cell ALL trials 50
Figure 3.2. Univariable relapse-free survival (RFS) and overall survival (OS) for traditional risk/prognostic variables over the six trials 52
Figure 3.3. Prognostic impact of treatment response variables in univariable relapse-free survival (RFS) and overall survival (OS) 53
Figure 3.4. Event-free survival (EFS), relapse rate (RR) and overall survival (OS) at five years for the main genetic subtypes 58
Figure 4.1. Main steps of the model development and validation in phase I (UKALL X+XI and UKALL 97/99) and phase II (UKALL 97/99, UKALL 2003 and UKALL 2011) 65
Figure 4.2. Prognostic effect of clinical variables in univariable Cox regression in the combined UKALL X+XI trial 71
Figure 4.3. Evaluation of Cox regression assumption for the variables in baseline prognostic models developed in UKALL X+XI 73

Figure 4.4. Prognostic impact of available variables from univariable Cox regression analyses in UKALL 97/99 76
Figure 4.5. Evaluation of Cox regression assumption for the variables in baseline prognostic models developed in UKALL 97/99 78
Figure 4.6. Identification of optimal threshold for the D8BM% model 81
Figure 4.7. The proportion of T-cell ALL patients within the risk groups defined by the D8BM% model on UKALL 97/99, UKALL 2003 and UKALL 2011 82
Figure 4.8. The percentage of MRD cases at the end of induction in the risk groups in UKALL 2003 and UKALL 2011 82
Figure 4.9. Kaplan-Meier graphs for 5-year event-free survival, relapse rate and overall survival of the risk groups 85
Figure 4.10. Calibration graphs and scores of the D8BM% model in relapse-free survival analysis at five years in UKALL 2003 and UKALL 2011 86
Figure 4.11. Interactions between the risk groups and MRD levels at the end of induction in terms of 5-year relapse rate 89
Figure 4.12. Discriminative abilities of the D8BM% model and MRD stratification at the end of induction in terms of relapse status in UKALL 2003 and UKALL 2011 90
Figure 4.13. Interactions between the risk groups and MRD levels at the end of induction (<0.1%, 0.1%-5%, ≥5%) in terms of 5-year relapse rate 92
Figure 4.14. Discriminative abilities of the D8BM% model and MRD stratification at the end of induction in terms of relapse status in UKALL 2003 and UKALL 2011 93
Figure 4.15. Forest plot showing the hazard ratios of the subgroups in UKALL 97/99, UKALL 2003, and UKALL 2011 94
Figure 4.16. Prognostic significance of traditional marrow classification in UKALL 97/99, UKALL 2003, and UKALL 2011 96
Figure 4.17. Possible clinical usefulness of the D8BM% model with MRD stratification at the end of induction 102
Figure 5.1. Inputs (blue and orange colours) and outputs (black colour) of a Nonlinear Mixed-Effects (NLME) model for the identification of treatment response trend in UKALL 2003 and UKALL 2011 106
Figure 5.2. The calculation of the AUC model score using the trapezoid formula 107
Figure 5.3. The distributions of the NLME trend score (coefficient 3) and AUC trend score in UKALL 2003 and UKALL 2011 108
Figure 5.4. Correlation between the observed and predicted response values in UKALL 2003 and UKALL 2011 109

Figure 5.5. Threshold analysis for the NLME and AUC models in UKALL 2003 110
Figure 5.6. Treatment response trends of patients during induction phase in the standard-risk (SR) and high-risk (HR) groups defined by the NLME and AUC models in UKALL 2003. 111
Figure 5.7. Patients moving to different risk groups identified by the NLME and AUC models 113
Figure 5.8. Clinical endpoints of the risk groups defined by the dynamic prognostic models in UKALL 2003 114
Figure 5.9. Treatment responses of the risk groups identified by the NLME and AUC models in UKALL 2011 115
Figure 5.10. Patients shifting from the standard-risk group to the high-risk group between the dynamic models in UKALL 2011 117
Figure 5.11. Event-free survival (EFS), relapse rate (RR) and overall survival (OS) of the risk groups in UKALL 2011 118
Figure 5.12. Evaluation of the discriminative abilities of the static and dynamic models in UKALL 2003 and UKALL 2011 in terms of relapse status 122
Figure 5.13. Distribution of the score of EOI-EOC MRD model in UKALL 2011 123
Figure 5.14. Threshold analysis of EOI-EOC MRD model score to identify optimal cut-off in UKALL 2011 123
Figure 5.15. MRD responses in log scale in the risk groups defined by the EOI-EOC MRD model in UKALL 2011 124
Figure 5.16. Clinical outcomes of the risk groups defined by the EOI-EOC MRD model in UKALL2011 125
Figure 5.17. Discriminative abilities of the EOI-EOC MRD model and MRD stratification with the cut-off of 0.01% at the end of consolidation 126
Figure 6.1. Main steps of differential expression analysis using the DESeq2 algorithm 132
Figure 6.2. Overview of an unsupervised clustering method using the ConsensusClusterPlus algorithm 133
Figure 6.3. Processing RNA-seq data (FASTQ files) for further investigations using the TALLSorts, DESeq2 and ConsensusClusterPlus 135
Figure 6.4. PCA plots for the identification of the batch effect in RNA-seq data 137
Figure 6.5. Assumption of the DESeq2 model (negative binomial distribution) depending on mean and variance, and its function 137
Figure 6.6. Probability (predicted) scores of the subtypes estimated by the TALLSorts 140
Figure 6.7. Distribution of the probability (predicted) scores of the subtypes over patients (samples) 141

Figure 6.8. Distribution of probability (predicted) scores of the subtypes in the characterised genetic subgroup 143
Figure 6.9. Confusion matrix comparing the DNA subtypes and predicted subtypes in the characterised genetic subgroup 144
Figure 6.10. Distribution of probability (predicted) scores and confusion matrix for unmatched subtypes in the characterised genetic subgroup..... 145
Figure 6.11. The probability (predicted) scores of the samples without the known gene fusions 146
Figure 6.12. Proportion of the predicted subtypes in the risk groups defined by the D8BM% model..... 147
Figure 6.13. Dysregulated genes (blue) with adjusted p-value <0.05 and absolute log2 fold change >0 between the risk groups 148
Figure 6.14. Volcano plot for differentially expressed genes with their gene symbols in the high-risk group 149
Figure 6.15. Gene set enrichment analysis for the 156 genes..... 150
Figure 6.16. Methodology for identification of the overlapping genes between the DESeq2 model and TALLSorts algorithm 151
Figure 6.17. Normalised expression levels of the four highly dysregulated genes in the risk groups..... 152
Figure 6.18. Identification of optimal cluster number using the ConsensusClusterPlus 154
Figure 6.19. Proportion of the risk groups defined by the D8BM% model and the subtypes predicted by TALLSorts in five clusters..... 156

List of Tables

Table 1.1. Acute lymphoblastic leukaemia classification according to the latest WHO report . 6
Table 1.2. European Group for the Immunological Characterisation of Leukaemias (EGIL) classification for acute lymphoblastic leukaemia . 6
Table 1.3. Immunophenotypic classification of T-cell ALL at different maturation stages . 11
Table 1.4. NCI-Rome criteria based on age and WBC stratifying ALL patients into standard-risk and high-risk groups (WBC, white blood cell count, L, litre) . 19
Table 2.1. Summary information about available paediatric/young adult patients on the UKALL trials . 30
Table 2.2. Availability of the variables in paediatric/young adult UKALL trials for T-cell ALL . 31
Table 2.3. Percentage of missing values in the UKALL trials . 40
Table 3.1. Definition of early marrow response groups based on regimens and marrow measurement days . 46
Table 3.2. Trends of clinical characteristics of T-cell ALL patients over time . 48
Table 3.3. The number of tested type-A and type-B abnormalities in UKALL 97/99 and UKALL 2003 . 54
Table 3.4. Classification of genetic abnormalities in UKALL 97/99 and UKALL 2003 . 55
Table 3.5. Clinical characteristics of classified group (immature, TLX3, proliferative, TAL/LMO) and classified + unclassified group in UKALL 97/99 and UKALL 2003 . 56
Table 3.6. Clinical characteristics of patients with the main genetic subtypes . 57
Table 4.1. Distribution of missing values in marrow variables by treatment . 66
Table 4.2. Distribution of clinical characteristics between complete data (patient with complete values) and missing data (patients with missing values) in UKALL X+XI and UKALL 97/99 . 67
Table 4.3. The number of patients recruited to different hospitals for UKALL97/99 trial between 1997 and 2002 . 68
Table 4.4. Selection of optimal k threshold in the KNN imputation for marrow variables in UKALL X+XI and UKALL 97/99 . 69
Table 4.5. Comparison of clinical characteristics of patients with complete and complete+imputed data in UKALL X+XI and UKALL 97/99 . 70
Table 4.6. Selection methods for the model development in time-to-event analysis . 74
Table 4.7. Hazard ratios of normalised scores of the prognostic models developed in UKALL X+XI and validated in UKALL 97/99 . 75

Table 4.8. Backward and forward selection strategies to develop baseline prognostic models in UKALL 97/99 79
Table 4.9. Hazard ratios and C-indexes of normalised scores of the prognostic models developed in UKALL 97/99 80
Table 4.10. Hazard ratios of normalised index of RFS-model (D8BM% model) in UKALL 2003 and UKALL 2011 80
Table 4.11. Clinical characteristics of the risk groups defined by the D8BM% model..... 83
Table 4.12. Distribution of the genetic subtypes and type-B abnormalities within the risk groups in the combined dataset (UKALL 97/99 + UKALL 2003)..... 87
Table 4.13. Correlation between the D8BM% model and MRD levels 88
Table 4.14. Correlation between the D8BM% model and MRD levels in UKALL 2003 and UKALL 2011 91
Table 4.15. Clinical characteristics and outcomes of the patients treated with AALL0434 protocol in the TARGET-ALL cohort 98
Table 4.16. Clinical characteristics and outcomes of patients in the risk groups identified by the D8BM% model in TARGET-ALL cohort 98
Table 5.1. Clinical characteristics of patients in the risk groups from the NLME and AUC models in UKALL 2003 112
Table 5.2. Clinical characteristics of patients in the risk groups identified by the dynamic models in UKALL 2011 116
Table 5.3. The number of patients moving to the opposite risk groups between the dynamic models and D8BM% model in UKALL 2003 and UKALL 2011..... 119
Table 5.4. Correlation and interation between the risk groups defined by the dynamic models and MRD levels at the end of induction in UKALL 2003 and UKALL 2011 121
Table 5.5. Clinical characteristics of patients in the high-risk and very high-risk groups in UKALL 2011 124
Table 5.6. Distribution of patients in the risk groups defined by different prognostic models and MRD stratification at the end of consolidation in UKALL 2011 126
Table 6.1. Clinical characteristics and outcomes of patients in the tested cohort and total cohort (tested + non-tested cases) from UKALL 2003 and UKALL 2011 139
Table 6.2. The number and proportion of the predicted subtypes by the TALLSorts..... 142
Table 6.3. The accuracy of TALLSorts predicting correct subtypes (DNA subtypes) detected by the laboratory methods..... 144
Table 6.4. Distribution of sample quality in the characterised genetic subgroup..... 145

Table 6.5. Distribution of the predicted subtypes in the risk groups defined by the D8BM% model 147
Table 6.6. Summary information about 12 overlapping genes between the genetic classifier and differential expression analysis 151
Table 6.7. Clinical characteristics and outcomes of patients in the clusters defined by the consensus clustering analysis 155
Table 6.8. Distribution of the predicted subtypes within the clusters defined by the unsupervised analysis 156

Chapter 1. Introduction

1.1. Haematopoiesis

Haematopoiesis is defined by the formulation of cellular elements of blood and blood plasma, including maturation of stem cells into different types of blood cells. Mature cells in the blood have numerous functions from supplying oxygen to organs/tissues, to fighting pathogens/infections (Rieger and Schroeder, 2012).

The primitive haematopoietic cells (haemopoietic stem cells, HSCs) are highly self-renewing stem cells and have the ability to differentiate into other cells. They are produced and reside in the bone marrow. The HSCs generate progenitor cells that can differentiate into two main lineages: myeloid lineage and lymphoid lineage. As the cells become more mature according to lineage type, B-lymphocytes, T-lymphocytes, and natural killer cells are produced from the lymphoid lineage cells, and thrombocytes, basophils, neutrophils, eosinophils, and monocytes are generated from myeloid lineage cells (Gunsilius, Gastl and Petzer, 2001) (Figure 1.1.).

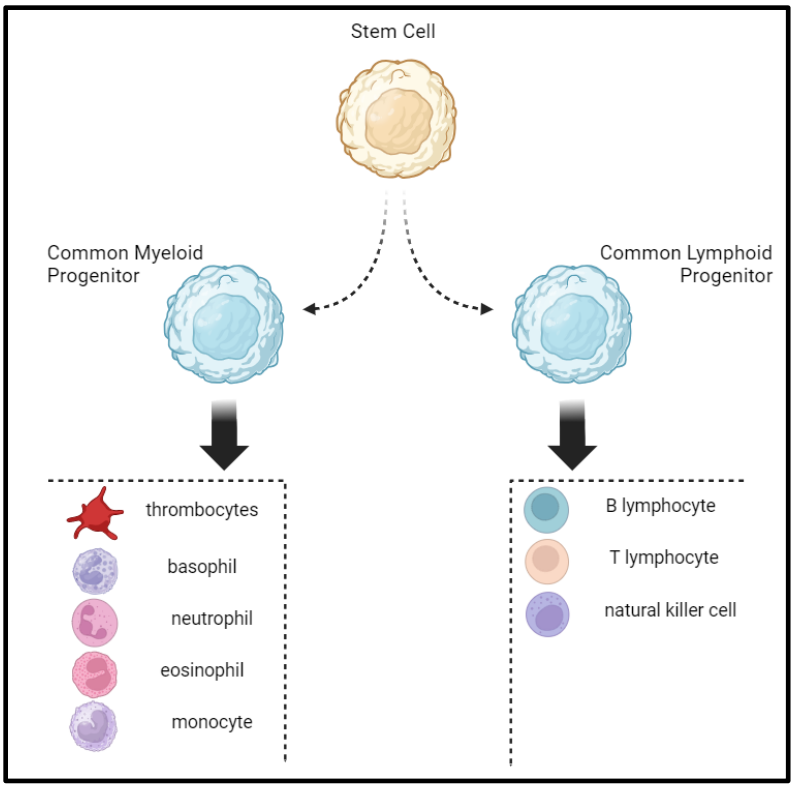

Figure 1.1. Mature blood cells stemming from two different lineages originated from stem cells (created with BioRender.com).

B-cells (B-lymphocytes) mature in the bone marrow, whereas T-cells (T-lymphocytes) differentiate to final form in the thymus. They are responsible for critical functions in the human immune system, such as initiating and mediating the adaptive immune response against foreign antigens, toxins, pathogens, or cancerous cells. B-cells help the immune response by producing antibodies, while T-cells contribute by directly eliminating targets and releasing cytotoxins (Todd, 2010).

T-cell development begins in the bone marrow with haemopoietic stem cells, and these progenitor cells (double negative (DN) pre-T cells with $CD4^-$ & $CD8^-$) migrate to the thymus. The pre-T cells undergo cellular changes within the thymus, and they become double positive (DP) T cells with $CD4^+$ & $CD8^+$. Double positive lymphocytes transform to single positive CD4 T cells with $CD4^+$ & $CD8^-$ or single positive CD8 T-cells with $CD4^-$ & $CD8^+$ via the mechanism of negative and positive selection. This selection process plays a critical role in eliminating cells with T-cell receptor (TCR) failures, and the fate of the cells depends on the level of antigen binding avidity. After the elimination process, $CD4^+$ and $CD8^+$ T cells mature in lymph nodes to become helper and cytotoxic T cells as the last step of the development (Kumar, Connors and Farber, 2018) (Figure 1.2.).

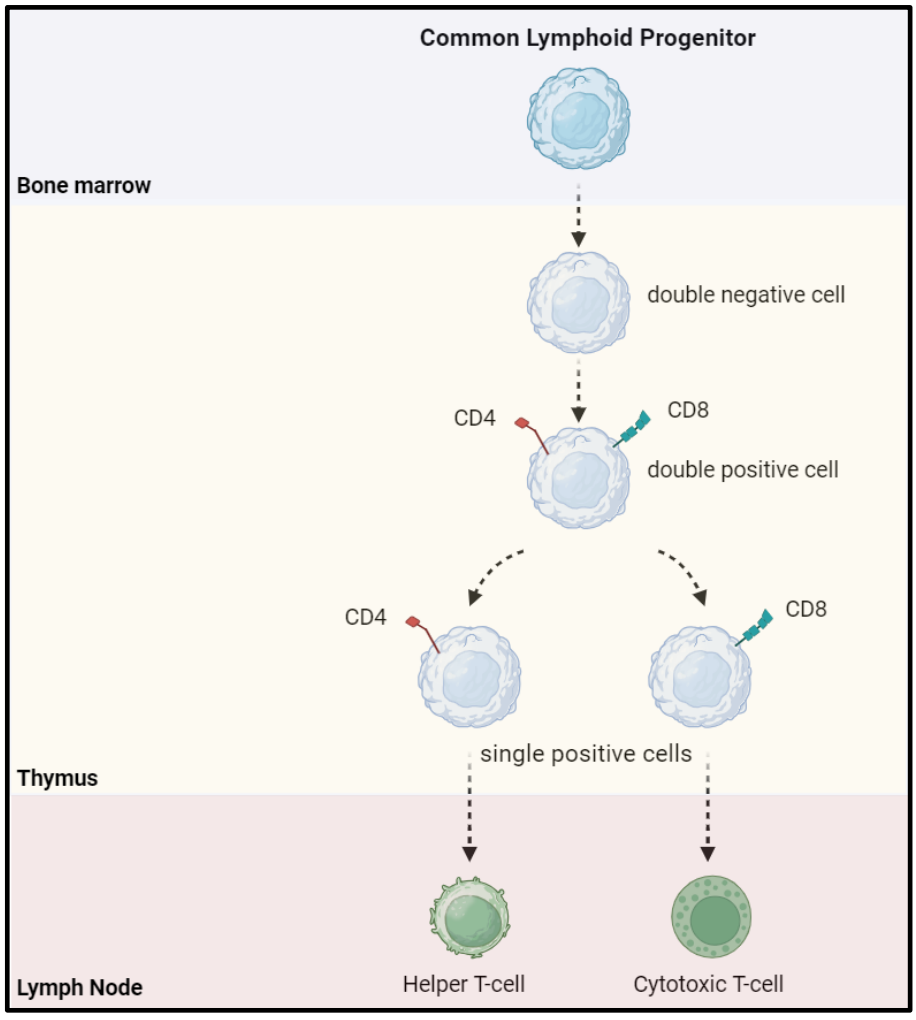

Figure 1.2. Differentiation stages of the T cells in different parts of the body (CD4, cluster of differentiation 4; CD8, cluster of differentiation 8) (Created with BioRender.com).

1.2. Overview of Acute Lymphoblastic Leukaemia (ALL)

1.2.1. Epidemiology and Aetiology

Acute lymphoblastic leukaemia (ALL) is associated with abnormal proliferation of lymphoid stem cells blocked at certain differentiation stages. Patients with this disease need immediate treatment because of its quick progression (Chang *et al.*, 2021). ALL is a rare cancer type, which consists of approximately 1% of the total cancers in United Kingdom, and the incidence rates of ALL are stable in the UK over time (Cancer-Research-UK, 2021) (Figure 1.3.). In 2025, the projected number of the disease in the UK will be 440-490 in children aged 0-19 and 300-330 in adults aged over 20 (Katz *et al.*, 2015).

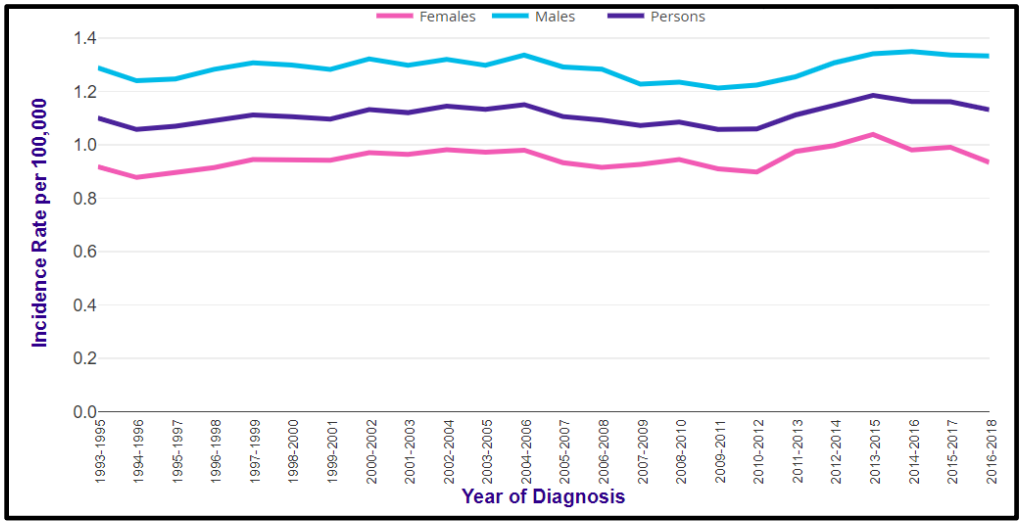

Figure 1.3. The incidence rate of ALL from 1993 to 2018 in the UK, taken from (Cancer-Research-UK, 2021).

Genetic susceptibility (Down syndrome, Fanconi anaemia, Ataxia telangiectasia and Bloom syndrome) and environmental causes (radiation exposure especially *in utero*, non-ionizing radiation, pesticides, benzene and smoking) can be risk factors in ALL. In addition to genetic susceptibility and environmental factors, *de novo* chromosomal abnormalities (specific translocations and gene rearrangements) can cause development of ALL (Redaelli *et al.*, 2005; Terwilliger and Abdul-Hay, 2017; Belson, Kingsley and Holmes, 2007). Viral infections (human T-cell lymphoma/leukaemia virus-1), age (more common in children and in adults aged over 50), and gender (more common in males) might be other risk factors.

1.2.2. Diagnosis

Clinical indications are the accumulation of immature lymphoid cells in bone marrow, blood and the sites of extramedullary haematopoiesis (liver, spleen, lymph nodes). Patients with ALL can have low levels of red blood cells, platelets and leukocytes, and the frequent symptoms are fever, bleeding, shortness of breath, enlargement of lymph nodes, spleen or

liver. Morphological evaluation of bone marrow samples is the first and necessary step to distinguish the myeloid and lymphoid lineages. Due to limitations of morphological assessment, flow cytometry is the gold standard method to identify lineage type and define subpopulations (Chiaretti, Zini and Bassan, 2014; Terwilliger and Abdul-Hay, 2017).

1.2.3. Classification

Historically, the earliest method of ALL classification is the French-American-British (FAB) system, which is based on morphological examination of the abnormal cells using cytochemistry techniques. The FAB system categorises ALL into three subgroups: L1 (small monotonous lymphocytes), L2 (mixed L1- and L3-type lymphocytes) and L3 (large homogeneous lymphocytes) (Bennett *et al.*, 1981). This classification system is not useful for leukaemia to distinguish subgroups of L1 and L2 in terms of predicting the clinical prognosis of patients; therefore, it is not used in the disease anymore.

ALL is categorised into two main subgroups: B-cell acute lymphoblastic leukaemia (B-cell ALL) with approximately 85% of childhood ALL cases, and T-cell acute lymphoblastic leukaemia (T-cell ALL) with about 15% of childhood ALL cases, depending on immunophenotypic features of the cells (Harrison and Johansson, 2015; Mroczek *et al.*, 2021). Moreover, they have been further classified into different subtypes within each subgroup in World Health Organisation (WHO) classification based on genetics and in EGIL classification based on immunophenotype.

The WHO published a report about classification of neoplastic diseases of the hematopoietic and lymphoid tissues in 1997. In this report, lymphoid neoplasms were categorised into B-cell neoplasms, T-cell and natural killer (NK)-cell neoplasms, and Hodgkin's disease (Harris *et al.*, 1999). In 2008, the revised WHO classification classified acute lymphoblastic leukaemia into B-cell lymphoblastic leukaemia with recurrent genetic abnormalities (*BCR::ABL1*, *KMT2A* (*MLL)* rearranged, *ETV6::RUNX1*, *IL3::IGH*, *TCF3::PBX1,* hyperdiploidy and hypodiploidy) and T-cell lymphoblastic leukaemia (Vardiman *et al.*, 2009). In the next revision (2016), the WHO added provisional entities to B-cell lymphoblastic leukaemia (*BCR/ABL1*–like and *iAMP21*) and to T-cell lymphoblastic leukaemia (Early T-cell precursor lymphoblastic leukaemia and NK-cell lymphoblastic leukaemia) (Arber *et al.*, 2016). The latest update of WHO classification has changed the subtype of “B-cell lymphoblastic leukaemia with hyperdiploidy” to “B-cell lymphoblastic leukaemia with high hyperdiploidy”. Also, B-cell lymphoblastic leukaemia with *ETV6/RUNX1*-like features and B-cell lymphoblastic leukaemia with *TCF3::HLF* fusion have been added into the revised classification. In

addition, NK-lymphoblastic leukaemia subtype has been removed in the section of precursor T-cell lymphoblastic leukaemia (Alaggio *et al.*, 2022) (Table 1.1.).

Acute Lymphoblastic Leukaemia
B-cell Acute Lymphoblastic Leukaemia
B-Lymphoblastic Leukaemia, NOS
B-Lymphoblastic Leukaemia with high hyperdiploidy
B-Lymphoblastic Leukaemia with hypodiploidy
B-Lymphoblastic Leukaemia with *iAMP21*
B-Lymphoblastic Leukaemia with *BCR::ABL1* fusion
B-Lymphoblastic Leukaemia *BCR::ABL1*-like features
B-Lymphoblastic Leukaemia with *KMT2A* rearrangement
B-Lymphoblastic Leukaemia with *ETV6::RUNX1* fusion
B-Lymphoblastic Leukaemia with *ETV6::RUNX1*-like features
B-Lymphoblastic Leukaemia with *TCF3::PBX1* fusion
B-Lymphoblastic Leukaemia with *IGH::IL3* fusion
B-Lymphoblastic Leukaemia with *TCF3::HLF* fusion
B-Lymphoblastic Leukaemia with other defined genetic abnormalities
T-cell Acute Lymphoblastic Leukaemia
T-Lymphoblastic Leukaemia, NOS
Early T-Precursor Lymphoblastic Leukaemia

Table 1.1. Acute lymphoblastic leukaemia classification according to the latest WHO report (NOS, not otherwise specified)

European Group for the Immunological Characterisation of Leukaemias (EGIL) classification is widely used in ALL and assesses leukemic cells according to maturation markers. B-cell ALL is grouped to Pro B-cell ALL, common B-cell ALL, pre B-cell ALL and mature B-cell ALL, while T-cell ALL is categorised into pro T-cell ALL, pre T-cell ALL, cortical/thymic T-cell ALL and mature T-cell ALL (Bene *et al.*, 1995) (Table 1.2.) (see more information about T-cell ALL classification in section 1.3.2.).

Acute Lymphoblastic Leukaemia
B-cell Acute Lymphoblastic Leukaemia
Pro B-cell Acute Lymphoblastic Leukaemia
Common B-cell Acute Lymphoblastic Leukaemia
Pre B-cell Acute Lymphoblastic Leukaemia
Mature B-cell Acute Lymphoblastic Leukaemia
T-cell Acute Lymphoblastic Leukaemia
Pro T-cell Acute Lymphoblastic Leukaemia
Pre T-cell Acute Lymphoblastic Leukaemia
Cortical/Thymic T-cell Acute Lymphoblastic Leukaemia
Mature T-cell Acute Lymphoblastic Leukaemia

Table 1.2. European Group for the Immunological Characterisation of Leukaemias (EGIL) classification for acute lymphoblastic leukaemia.

1.2.4. Genetics

It is commonly believed that the initial chromosomal/genetic changes in childhood ALL begin *in utero*. Molecular evidence of prenatal leukemogenesis is observed from a twin and backtracking study (Ford, Colman and Greaves, 2023), and over 40% of children with ALL have certain abnormalities (*ETV6::RUNX1* and hyperdiploidy) (Marcotte et al., 2021), which has supported the leukemogenesis initiation *in utero.*

B-cell ALL typically involves two major abnormal genetic events (primary abnormalities), which help initiate leukemic clone: aneuploidy, and chromosomal rearrangements. In the case of aneuploidy, high hyperdiploidy (51-65 chromosomes), near-haploidy (less than 30 chromosomes), low-hypodiploidy (30-39 chromosomes) and hypodiploidy (40-45 chromosomes) occur in this subtype. On the other hand, those with chromosomal rearrangements can have different type of translocations, such as *ETV6::RUNX1*, *TCF3::PBX1*, *BCR::ABL*, *IGH or KMT2A (MLL)* translocations. Deletions and mutations are secondary abnormalities, which present in the subclones. Primary abnormalities are more useful as prognostic indicators and reliable for risk stratification than secondary abnormalities because primary abnormalities reflect the main characteristics of the leukemic clones (Moorman, 2016). Primary and secondary abnormalities can occur at different maturation stages (pro-B cell, common B-cell, pre-B cell and mature B-cell) during cell growth and cell differentiation.

Unlike B-cell ALL, chromosomal abnormalities are rare events in T-cell ALL. Patients with this subgroup can have type A and type B abnormalities. Similar to primary and secondary abnormalities in B-cell ALL, type A and type B abnormalities are responsible for the initiation and maintenance of T-cell leukemic clones. Type A abnormalities are characterised by ectopic expressions of transcription factors, such as *TAL*, *LMO*, *NKX2*, *TLX1*, *TLX3* etc., while type B abnormalities cover mostly insertions, deletions and mutations in certain genes (*NOTCH1*, *FBXW7, CDKN2A/2B* etc.) (Girardi *et al.*, 2017). The changes in expression profiles of T-cell blasts at any differentiation stages contribute to the development of T-cell ALL, as discussed in detail later (section 1.3.3.).

1.2.5. Treatment

Treatment regimens for patients with B-cell ALL and T-cell ALL are the same in most clinical trials. Standard first-line treatment is multiagent chemotherapy with or without cranial radiation. Allogenic stem cell transplantation (SCT) is an option for those with persistent minimal/measurable residual disease (MRD) levels (remaining leukemic cells after treatment). The treatment regimens are given to patients in four treatment phases called

remission induction, consolidation, intensification, and maintenance within 2-3 years. Remission induction phase lasts 4-6 weeks to reduce the disease burden so that patients can have normal haematopoiesis. Steroid (prednisone or dexamethasone), vincristine, asparaginase and chemotherapeutic agents (methotrexate, cytarabine and hydrocortisone) are commonly used in this phase. The consolidation phase is the second step of the treatment lasting 6-8 months. The aim of this phase is to reduce any remaining residual leukemic cells. Patients are classified into risk groups based on treatment response (MRD) at the end of induction and treated with multiple cycles of intensive chemotherapy (the same agents in remission induction plus Ara-C, etoposide, methotrexate and 6-mercaptopurin). The following phase is the intensification step that is the same as the remission induction phase. The last phase of the treatment is the maintenance step aiming to decrease the risk of relapse, and it takes approximately two years. Oral chemotherapy drugs, such as methotrexate and 6-mercaptopurine in the form of pill with or without vincristine and steroid pulses are given to patients at this longest stage (Malard and Mohty, 2020; Kato and Manabe, 2018; Lee and Cho, 2017) (Figure 1.4.).

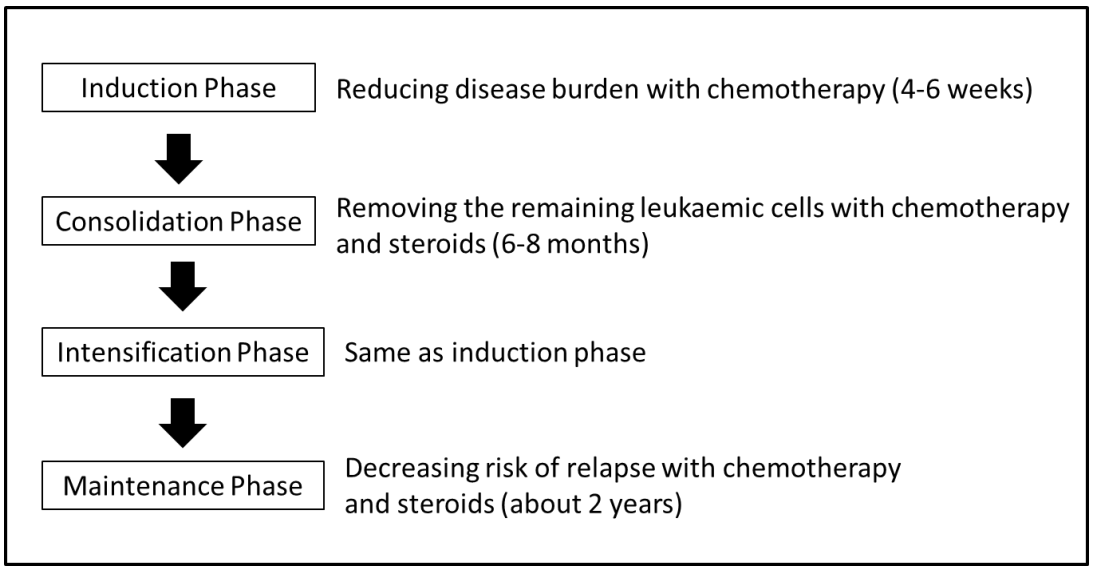

Figure 1.4. Overall view of treatment phases in T-cell ALL.

1.2.6. Clinical Outcome

Prognosis of patients with ALL has dramatically increased with modern therapies and stratifications over time (Hunger *et al.*, 2012; Mitchell *et al.*, 2010), and the gap between B-cell ALL and T-cell ALL has considerably shrunk in event-free survival and overall survival (Campbell *et al.*, 2023; Teachey and Pui, 2019). However, T-cell ALL patients still have inferior overall survival rate after relapse in comparison to B-cell ALL cases (Rheingold *et al.*, 2019).

It is clearly known that NCI-Rome classification (age and white blood cell count at diagnosis) and MRD are the key prognostic factors to categorise patients into risk groups for treatment allocation in B-cell ALL (Cazzaniga *et al.*, 2003; Schultz *et al.*, 2006). In addition to these independent prognostic factors, primary abnormalities are also used to stratify patients into cytogenetic risk groups for personalised medicine. For example, patients with high hyperdiploidy or *ETV6::RUNX1* show favourable prognosis (Paulsson and Johansson, 2009; Wang, Zeng and Zhang, 2018), but those with haploidy, *BCR/ABL1*-like or *KMT2A* rearrangements have worse clinical outcomes (Roberts *et al.*, 2017; Wen *et al.*, 2022; Molina *et al.*, 2022). On the other hand, only MRD is the key prognostic indicator, which is widely used for risk stratification in T-cell ALL, and there is no absolute consensus on prognostication of genetic abnormalities (type A or type B) in this disease (see detail information in section 1.3.3.).

1.3. Overview of T-cell Acute Lymphoblastic Leukaemia

T-cell acute lymphoblastic leukaemia (T-cell ALL) is an aggressive malignant disease and is characterised by the accumulation of immature blasts in the bone marrow. The overproduction of the immature lymphocytes results in the disruption of normal haematopoiesis and failure of production of normal blood cells. In addition to the aggressiveness of the disease, T-cell ALL is a highly heterogenous cancer type in terms of its biology and genetics; hence, it is difficult to treat patients with various abnormalities.

1.3.1. Epidemiology

T-cell ALL is a rare blood cancer type, and it is less common in children compared to adults. In Nordic countries (Sweden, Denmark, Norway, Finland, and Iceland), incidence rate for T-cell ALL cases aged 0-14 was stable from 1985 to 2001, and it was approximately 0.36 cases/100,000 children-years (Hjalgrim *et al.*, 2003). In the United States, the incidence rate of newly diagnosed T-cell ALL patients aged over 20 years was 0.13 per 100,000 persons from 2001 to 2014 (Guru Murthy *et al.*, 2019).

Patients are predominantly male (Teachey and Pui, 2019), and frequency of T-cell ALL gradually rises with increasing age of patients (Toft *et al.*, 2018).

1.3.2. Classification

There are two main classification systems for T-cell ALL: 1) European Group for the Immunological Characterisation of Leukaemias (EGIL) classification and 2) World Health Organisation (WHO) classification.

In the EGIL classification, T-cell ALL can be grouped into different categories based on the level of cell differentiation: pro-T-cell ALL, pre-T-cell ALL, cortical T-cell ALL and mature T-cell ALL (Bene *et al.*, 1995). CD3 (cluster of differentiation 3) is an essential marker for the identification of the lineage, and T-cell ALL cases have CD3 expression. CD1a is a specific indicator for cortical T-cell ALL, while positivity of CD4 or CD8 is a unique expression signature for mature T-cell ALL as single-positive cells (Table 1.3.). Recently, high-risk subtype of T-cell ALL characterised with reduced expression of CD1a, CD5 and CD8 and distinct gene expression profiles has been defined, known as early T-cell precursor (ETP)-ALL (Coustan-Smith *et al.*, 2009). Another subtype, near-ETP, is still controversial in terms of genomic profile/clinical aspects, compared to ETP-ALL (Sin and Man, 2021).

	cCD3	sCD3	CD7	CD1a	TdT	CD2	CD5	CD4/CD8	Stem Cell/Myeloid
ETP-ALL	+	–	+	–	±	–	±	–/–	+/– or –/+ or +/+
Near ETP-ALL	+	–	+	–	±	–	+	–/–	+/– or –/+ or +/+
Pro-T-cell ALL	+	–	+	–	±	–	–	–/–	–
Pre-T-cell ALL	+	±	+	–	±	+	+	–/– or +/+	–
Cortical T-cell ALL	+	±	+	+	±	+	+	–/– or +/+	–
Mature T-cell ALL	+	+	+	–	±	+	+	+/– or –/+	–

Table 1.3. Immunophenotypic classification of T-cell ALL at different maturation stages, taken from (Genescà and la Starza, 2022) (-, negative; +, positive; CD, cluster of differentiation; c, cytoplasmic; s, surface; TdT, terminal deoxynucleotidyl transferase).

The latest WHO classification (5th edition) groups T-cell ALL into two main subgroups as typical T-cell ALL and ETP-ALL based on the immunophenotypic criteria (Table 1.1.) (Alaggio *et al.*, 2022).

In addition to the main classification systems, International Consensus Classification (ICC) subclassifies the disease to T-cell ALL, ETP-ALL, and ETP-ALL with *BCL11B* rearrangement. There is also a provisional entity in the ICC report called natural killer cell ALL, which has been removed in the new version of WHO classification (Arber *et al.*, 2022).

1.3.3. Cytogenetics/Genomics of T-cell ALL

The genetics of T-cell ALL is heterogeneous, which means that multiple steps are required to accumulate genomic abnormalities for the disruption of normal cellular functions, such as cell proliferation, growth, and differentiation. These genomic alterations can be grouped into two categories: 1) type-A abnormalities blocking T-cell differentiation at any stage and facilitating the initiation of disease-specific events with the ectopic expression of transcription factors (*TAL1*, *TAL2*, *LMO1*, *LMO2*, *LYL1*, *TLX1*, *TLX3*, *HOXA,* etc.) or with translocations of *CALM-AF10* or *SET-NUP214*; 2) type-B abnormalities synergising with type-A abnormalities and providing some advantages to the cells in terms of disease progression and relapse by affecting specific genes (*CDKN2A/B*, *NOTCH1*, *FBXW7*, *RAS*, *PTEN,* etc.) that play different roles in cellular processes of cell cycle, differentiation and self-renewal (Van Vlierberghe *et al.*, 2008). Type-A abnormalities are associated with the activation of oncogenes/oncogene fusions, whereas type-B abnormalities contain mutations, insertions and deletions (INDELs).

1.3.3.1. Type-A Abnormalities

Type-A abnormalities can be evaluated in several subtypes based on the expression of specific transcription factors, including type-B alterations. The main mechanisms that cause over expression of transcription factors are gene fusions with T-cell receptor (*TCR*) genes, translocations with other regulatory DNA sequences and molecular alterations to create a new enhancer region, such as duplications, mutations, or insertions. These molecular mechanisms

initiate the genomic changes at the transcriptional and protein levels for leukemogenesis (Girardi *et al.*, 2017). Although T-cell ALL subtypes can be broadly classified as TAL subtype, LMO subtype, TLX1 subtype, TLX3 subtype, HOXA subtype, NKX2.1 subtype and BCL11B subtype, depending on specific transcription factors mostly driven by T-cell receptor (*TCR)* genes (*TCRα/δ or TCRβ*), an unsupervised hierarchical clustering analysis has revealed that these seven subtypes can be narrowed to four broader subgroups using immunophenotype and oncogenic aberrations: 1) immature subtype, 2) TLX3 subtype, 3) proliferative (TLX1/NKX2.1) subtype, and 4) TAL/LMO subtype (Homminga *et al.*, 2012) (Figure 1.5.). Despite the fact that each subtype specifically has more frequent type-B abnormalities, *NOTCH1*, *FBXW7* and *CDKN2A* alterations can be frequently mutated in any of these subgroups (Liu *et al.*, 2017a).

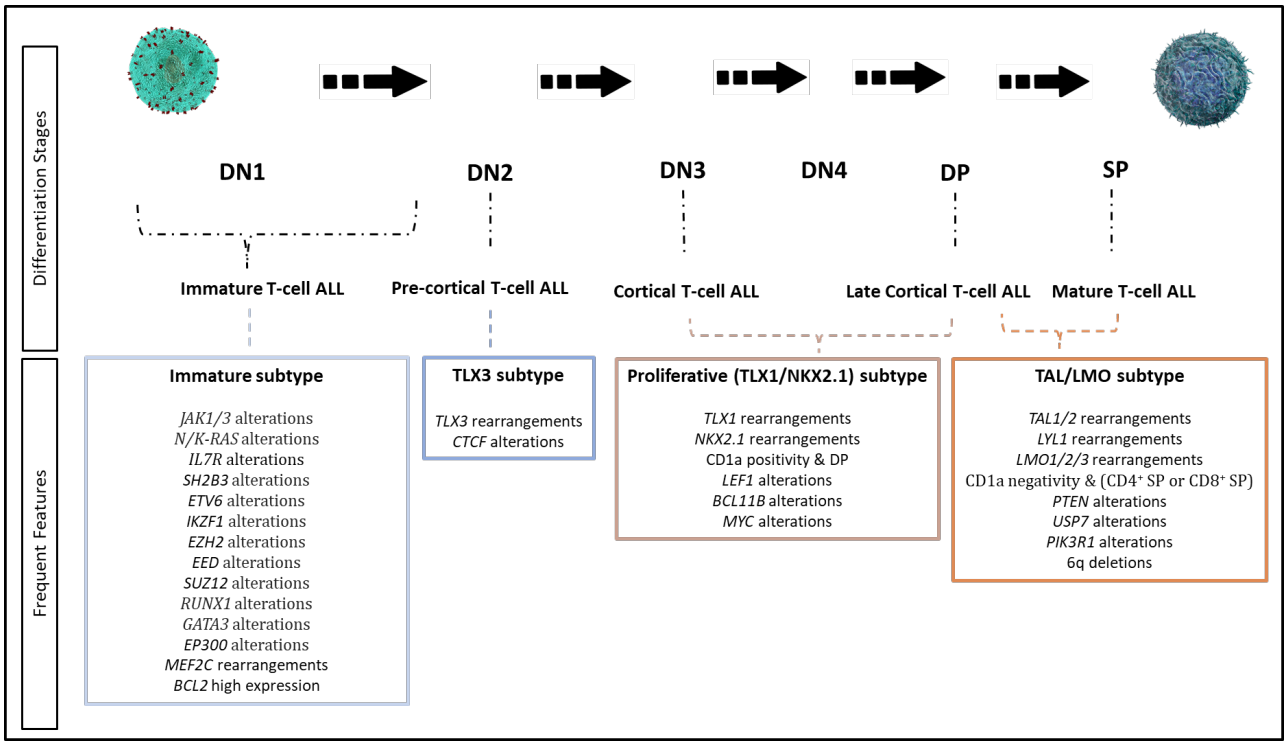

Figure 1.5. Four broader subgroups of T-cell ALL based on differentiation stages and genetic/cytogenetic abnormalities (DN, double negative; DP, double positive; SP, single positive; CD, cluster of differentiation).

1.3.3.1.1. Immature T-cell ALL Subtype

The immature T-cell ALL subgroup consists of early T-cell ALL precursor (ETP) cells, near-ETP cells and pro T-cell ALL cells blocked at early stages of the development process. Most immature cells belong to ETP-ALL, which is defined in the WHO classification system based on CD (cluster of differentiation) markers (Arber *et al.*, 2016). ETP-ALL is highly heterogenous in terms of genetic basis compared to typical T-cell ALL, and it accounts for

10-16% of T-cell ALL cases (Coustan-Smith *et al.*, 2009; Patrick *et al.*, 2014a; Noronha *et al.*, 2021).

A comprehensive study has demonstrated that gene mutations of *NOTCH1* and *FBXW7* are not common in this subtype, but activating mutations are frequent in the following pathways/genes: cytokine receptor and RAS signalling pathway (*N/K RAS*, *FLT3*, *IL7R*, *JAK1/3*, *SH2B3*), alterations in epigenetic regulator genes (*EED*, *EZH2* and *SUZ12*) and loss-of-function alterations in haematopoietic and lymphoid development genes (*RUNX1, IKZF1*, *ETV6*, *GATA3* and *EP300*) (Zhang *et al.*, 2012). Additionally, overexpression of *BCL2* and rearrangements of *MEF2C* have been identified in this subtype (Chonghaile *et al.*, 2014; Homminga *et al.*, 2011).

Although ETP-ALL is a poor indicator, and patients with this subtype have poorer chemotherapy response/clinical outcomes than those with typical T-cell ALL (Inukai *et al.*, 2012; Coustan-Smith *et al.*, 2009; Allen *et al.*, 2013), other researchers have shown that there is no statistically significant difference between ETP-ALL and non ETP-ALL patients in terms of event-free survival and overall survival (Wood *et al.*, 2014; Patrick *et al.*, 2014a).

1.3.3.1.2. TLX3 Subtype

TLX3 subtype is associated with either *TLX3* (*HOX11L2*) rearrangements or immunophenotype of immature lymphocytes blocked at DN2 stage (Cordo *et al.*, 2021). *TLX3* (T-Cell Leukaemia Homeobox 3) is located on chromosome 5 (5q35.1), encoding a transcription factor, which has a role in cell differentiation signalling pathway (GeneCards, 2023c). This gene rearranges the enhancer/promotor region of *BCL11B* (t(5;14)(q35;q32)) and *TCR* (t(5;14)(q35;q11)) in T-cell ALL (Girardi *et al.*, 2017). *CTCF* alterations are predominantly enriched in the subtype (Liu *et al.*, 2017a). There is no consensus on good or poor prognostic impact of *TLX3* rearrangements in children and young adults with T-cell ALL (Ballerini *et al.*, 2002; Gottardo *et al.*, 2005; Paola *et al.*, 2008; Grotel *et al.*, 2006).

1.3.3.1.3. Proliferative (TLX1/NKX2.1) Subtype

The main features of the proliferative subtype are the presence of rearrangements of *TLX1* (*HOX11*) or *NKX2.1*, and T cells are arrested at the differentiation stage of DN3/DP (Cordo *et al.*, 2021). *TLX1* (T-Cell Leukaemia Homeobox 1) and *NKX2.1* (NK2 Homeobox 1), which produce transcription factor proteins related to cell differentiation signalling pathways, are on chromosome 10 (10q24.31) and chromosome 14 (14q13.3), respectively (GeneCards, 2023b; GeneCards, 2023a). Overexpression of *TLX1* is driven by *TCR* genes with the translocations of t(10;14)(q24;q11) and t(7;10)(q34;q24), while a chromosomal abnormality of inv(14)(q11;q13) results in the increased expression of *NKX2.1* (Girardi *et al.*, 2017). *LEF1*,

BCL11B and *MYC* alterations are highly frequent in this proliferative subtype (Liu *et al.*, 2017a). Kees and colleagues reported that paediatric T-cell ALL patients with *TLX1* rearrangements showed a prognostic advantage compared to negative-*TLX1* cases (Kees *et al.*, 2003). In addition to this research, another study demonstrated that paediatric T-cell ALL cases with t(10;14) could have a trend of longer survival (Schneider *et al.*, 2000).

1.3.3.1.4. TAL/LMO Subtype

TAL/LMO subtype is distinguished by the overexpression of *TAL1/2*, *LYL1*, *LMO1/2/3*, and this subtype predominantly shows the immunophenotypic features of single $CD4^+$ or $CD8^+$ cells, indicating a late cortical/mature differentiation stage (Cordo *et al.*, 2021). *TAL1/2* and *LYL1* belong to the gene family called basic helix-loop-helix (bHLH), and *LMO1/2/3* are a member of the cysteine-rich (LIM-domain) family. These gene families are involved in important cellular processes, such as cell differentiation, proliferation, cell cycle regulation and metastasis (Ledent and Vervoort, 2001; Sang *et al.*, 2014). Genomic abnormalities, such as *de novo* formation of enhancer region via insertion, deletion (del(1)(p32p32)), translocations (t(1;14)(p32;q11) with *TCR* genes, and t(7;9)(q34;q32)) cause ectopic expressions of *TAL1* and *TAL2*. t(11;14)(p15;q11) and t(11;14)(p13;q11), which interact with *TCR* genes, lead to overexpression of *LMO1* and *LMO2*, respectively (Girardi *et al.*, 2017; Xia *et al.*, 1991). O'Connor and colleagues reported a higher proportion of *TAL1* cases with *LMO1* abnormalities in induction failure group in comparison with treatment response group (O'Connor *et al.*, 2023). The alterations of *PTEN*, *USP7* and *PIK3R1* are mostly associated with *TAL1*-positive cases, and 6q deletions are commonly seen in this subtype (Liu *et al.*, 2017a). A research group found that *PTEN* abnormalities were associated with poor prognosis in paediatric T-cell ALL cohort (Paganin *et al.*, 2018). However, another study reported that there was no statistically significant difference between the paediatric patients with/without *PTEN* abnormalities in terms of clinical outcomes (Jenkinson *et al.*, 2016).

1.3.3.2. Type-B Abnormalities

Type-B abnormalities refer to mutations or deletions in genes that are responsible for different biological processes. The most frequent type-B abnormalities in T-cell ALL occur in *NOTCH1* signalling pathway, containing the mutations in *NOTCH1* (60%) and *FBXW7* (15%) (Karrman and Johansson, 2017). Although this pathway regulates cell proliferation and cell death, it has a significant role for T-cell differentiation. Some studies reported that activating mutations in *NOTCH1* and/or inactivating mutations in *FBXW7* were an indicator of favourable prognosis (Park *et al.*, 2009; Yuan *et al.*, 2022; Natarajan *et al.*, 2015), whereas others demonstrated that there was no better clinical outcomes in those with the mutations (Clappier *et al.*, 2010;

Zuurbier *et al.*, 2010). In UKALL 2003, patients with *NOTCH1* mutations were associated with a better overall survival rate, compared to those without mutations; however, this patient group did not show a statistically significant difference in event-free survival and relapse rate (Jenkinson *et al.*, 2013).

Mutations/deletions in genes involving in PI3K-AKT signalling pathway are also observed in T-cell ALL, and this pathway controls cell proliferation and apoptosis. Mutations and deletions of *PTEN* (20%) are the most common alterations in this signalling pathway (Karrman and Johansson, 2017). Other alterations in the signalling pathway occur in *AKT* (2%) and *PI3K* (1%) genes (Girardi *et al.*, 2017). Despite the fact that there was a non-significant trend of having worse clinical outcomes in those with the mutations/deletions (Bandapalli *et al.*, 2013; Mendes *et al.*, 2014), the abnormalities in PTEN-PI3K-AKT signalling pathway did not have a prediction ability for event-free survival outcomes (Gutierrez *et al.*, 2009).

The JAK-STAT signalling pathway involves *IL7*, *JAK*, and *STAT*, and this pathway is responsible for controlling cell proliferation. Paediatric T-cell ALL patients can have the activating mutations in *IL7R* (10%), *JAK1* (5%), *JAK3* (5%), and *STAT5B* (Karrman and Johansson, 2017). There was no evidence that mutations in *IL7R*, *JAK1*, and *JAK3* were related to adverse clinical impacts in UKALL 2003 (Vicente *et al.*, 2015). Other research groups found that *JAK* mutations would be associated with a trend of increased risk of clinical events and relapse (Mullighan *et al.*, 2009) and patients with *IL7R* mutations were inferior to those without *IL7R* mutations in event-free survival analysis (Richter-Pechańska *et al.*, 2017). On the other hand, T-cell ALL cases with *STAT5B* mutations had an increased risk of relapse; however, they did not have worse event-free survival outcome (Bandapalli *et al.*, 2014).

Another signalling pathway is the RAS pathway, consisting of *NRAS*, *KRAS* and *HRAS*, which have essential roles in transmitting proliferative signals. *NRAS* mutations (14%) and *KRAS* mutations (6%) can be observed in paediatric T-cell ALL (Girardi *et al.*, 2017). Similar to other type-B abnormalities, the effects of *NRAS* and *KRAS* mutations on clinical outcomes differ across studies with different treatment protocols (Richter-Pechańska *et al.*, 2017; Jenkinson *et al.*, 2016).

Cell cycle regulators control cell cycle transitions during cell division. Monoallelic or biallelic deletions of *CDKN2A* (70%) and *CDKN2B* (60%) are the most common alterations in T-cell ALL (Karrman and Johansson, 2017). Although these deletions are very frequent in T-cell ALL, there is still no consensus on whether or not they lead to good and poor prognosis (Krieger *et al.*, 2010; Mirji *et al.*, 2016). *TP53* alterations can be observed in T-cell ALL; however, the

small number of cases with these alterations is the main limitation to determine their absolute prognostic impact in paediatric and young adult patients (Wada *et al.*, 1993).

1.3.4. Treatment

Backbone treatment for front-line therapy in T-cell ALL is the same as in B-cell ALL (Figure 1.4.). Additionally, targeted immunotherapy drugs have begun use in clinical trials for T-cell ALL patients in recent years as well as traditional therapy. For instance, nelarabine is a purine nucleoside analogue, which inhibits DNA synthesis in T cells. This drug has been added to COG AALL0434 treatment protocol for the relapsed and refractory T-cell ALL patients. The report concluded that addition of nelarabine improved disease-free survival rate and reduced central-nervous system relapses in children and adult patients (Dunsmore *et al.*, 2020). Another potential targeted drug is daratumumab, a monoclonal antibody that targets a surface protein called CD38 to help the immune system attack the tumour cells. This drug has been included in therapy in the DELPHINUS study for children and young adult patients with T-cell ALL. This study reported that it could be safely used with combination of chemotherapy in relapsed and refractory T-cell ALL and help the process of post-stem cell transplantation (Bhatla *et al.*, 2024). Also, a drug called Bortezomib, proteasome inhibitor, has been tested in T-cell ALL cohort treated with modified BFM protocol in AALL1231 trial. It was found that bortezomib offered a survival advantage to patients and decrease the use of cranial radiotherapy (Teachey *et al.*, 2022). The other targets of the drugs containing inhibitors or enzymes can be specific for signalling pathways, such as NOTCH signalling, PI3K/Akt/mTOR signalling, JAK/STAT signalling, MAPK signalling, and cell cycle. Studies investigating these certain signalling targets are going on for potential personalised therapies (Raetz and Teachey, 2016; Pocock *et al.*, 2021).

1.3.5. Clinical Outcome

Historically, the clinical outcomes of T-cell ALL patients are inferior to those with B-cell ALL cases. However, recent outcome for those with T-cell ALL has dramatically increased reaching survival rates of 80%-90% at five years, which results in similar clinical outcomes as those seen in B-cell ALL (Raetz and Teachey, 2016; Hunger *et al.*, 2012). Treatment intensification and stratification based on MRD are the main reasons for this dramatic improvement in T-cell ALL. Despite these major advancements, T-cell ALL patients that relapse still have very poor prognosis. The probability of 10-year event-free survival in relapsed childhood patients is worse than relapsed B-cell ALL patients (21% vs 37%) (Reismüller *et al.*, 2009), and the overall survival rate is 33% at five years (Rheingold *et al.*, 2019). Another report has indicated that children and adolescents with T-cell ALL have an 8-

year overall survival rate of 23% after disease recurrence. Furthermore, in this research, patients with bone marrow relapse have had an 8-year overall survival of 15% compared to those with extramedullary relapse (36%) and combined relapses (24%) (Vinti *et al.*, 2019). These results demonstrate that paediatric and young adult T-cell ALL patients with relapse still have dismal clinical outcomes and need new approaches for treatment strategy.

1.4. Evolution of Risk Stratifications in Acute Lymphoblastic Leukaemia

Prognostic factors are measurements, biomarkers, or clinical characteristics, which are associated with the clinical outcomes of patients. Algorithms/models are based on prognostic factors, and they are used to predict risk probability of future clinical outcomes, helping stratify patients into risk groups for treatment allocation. The primary aim of identifying risk groups in ALL is to treat patients with less or more intensive therapy. More importantly, directed therapies or stem cell transplantation can be treatment options for those assigning to higher risk of relapse (Moorman, 2020).

In B-cell ALL, there are many independent prognostic factors, such as age, white blood cell count at diagnosis, genetic abnormalities, and MRD; however, these factors are not as powerful as in T-cell ALL, except MRD. In this section, the evolution of risk stratifications will be assessing in the context of B-cell ALL and T-cell ALL prognosis over time.

1.4.1. Immunophenotype

T-cell ALL was first identified as a subtype of ALL by Borella and Sen in 1973 (Pui and Evans, 2013). Since then, T-cell ALL subtype has been considered a poor indicator of clinical events in comparison to B-cell ALL. Intensive chemotherapy and MRD-based risk classification have dramatically improved the prognosis of patients with T-cell ALL. Similarly, ETP-ALL can have excellent clinical results with modern treatments although this subtype is associated with a poor prognosis (Pocock *et al.*, 2021).

1.4.2. National Cancer Institute-Rome (NCI-Rome) Criteria

In 1993, the Paediatric Oncology Group (POG) and Children's Cancer Group (CCG) defined a stratification method for childhood ALL. This risk stratification is called National Cancer Institute-Rome (NCI-Rome) classification, which includes initial white blood cell count (WBC) and age (Smith *et al.*, 1996). Patients aged 1-9 years with WBC level of $<50{,}000/mm^3$ ($<50x10^9/L$) are considered as standard-risk group, while others aged less than one year or equal/more than 10 years or WBC $\geq 50{,}000/mm^3$ ($\geq 50x10^9/L$) are defined as high-risk group (Table 1.4.). This classification is commonly used in paediatric ALL trials, and it clearly demonstrates a prognostic difference between the standard-risk group and high-risk group in B-cell ALL cohort. In the UK, NCI-Rome stratification was first used in UKALL99 trial for childhood ALL (Mitchell *et al.*, 2009), and it has been used in the latest UKALL trial (UKALL 2011) for only B-cell ALL, but not for T-cell ALL (Goulden, 2013). The main reason why it has not been included in T-cell ALL stratification is that clinical significance of this classification (age and WBC) is debatable in the context of modern therapy (Burns *et al.*, 2021; Patrick *et al.*, 2014b; Petit *et al.*, 2018; Melchior *et al.*, 2012). Therefore, all T-cell ALL

patients have been treated with a 4-drug induction regimen instead of a 3-drug induction regimen in UKALL 2011 regardless of NCI-Rome stratification (Goulden, 2013).

NCI-Rome	Age (years)	WBC at diagnosis
Standard-risk group	1-9	$<50x10^9$/L
High-risk group	<1 or ≥10	$\geq 50x10^9$/L

Table 1.4. NCI-Rome criteria based on age and WBC stratifying ALL patients into standard-risk and high-risk groups (WBC, white blood cell count, L, litre).

1.4.3. Early Treatment Response

Early treatment response is a measurement of response to chemotherapy by identifying leukemic cells in the bone marrow or peripheral blood. The definition of early treatment response can be evaluated either by percentage of blasts using cytomorphologic features or minimal/measurable residual disease (MRD) using polymerase chain reaction or flow cytometry methods.

1.4.3.1. Marrow/Blood Response

Bone marrow aspiration is a technique to identify the disruption of normal haemopoiesis, which is a marker of leukaemia. ALL patients are internationally classified into three groups based on the count of lymphoblast in the bone marrow: M1 (lymphoblast <5%), M2 (lymphoblast of 5-25%) and M3 (lymphoblast ≥25%) (Buchmann *et al.*, 2022). This procedure also helps assess early treatment response to chemotherapy during the induction phase. Patients can have rapid early response (M1 or M2)/slow early response (M3) according to bone marrow assessment or good/poor response (<1000/μL / ≥1,000/μL) by peripheral blood assessment. Many research groups have reported that paediatric T-cell ALL patients with slow early marrow response or poor early blood response have an increased risk of induction failure, lower event-free and overall survival rates compared to rapid or good early responders (Gaynon *et al.*, 1997; Griffin *et al.*, 2000; Melchior *et al.*, 2012; Möricke *et al.*, 2013). In addition to these findings in the literature, T-cell ALL patients with slow early marrow response at day 8 or day 15 have been transferred to regimen C treatment group (more intensive treatment) during/after induction phase in the latest UKALL trials (UKALL 99, UKALL 2003 and UKALL 2011), indicating the importance of early treatment response in T-cell ALL (Figure 1.6.) (Goulden, 2013; Mitchell *et al.*, 2009; Patrick *et al.*, 2014b).

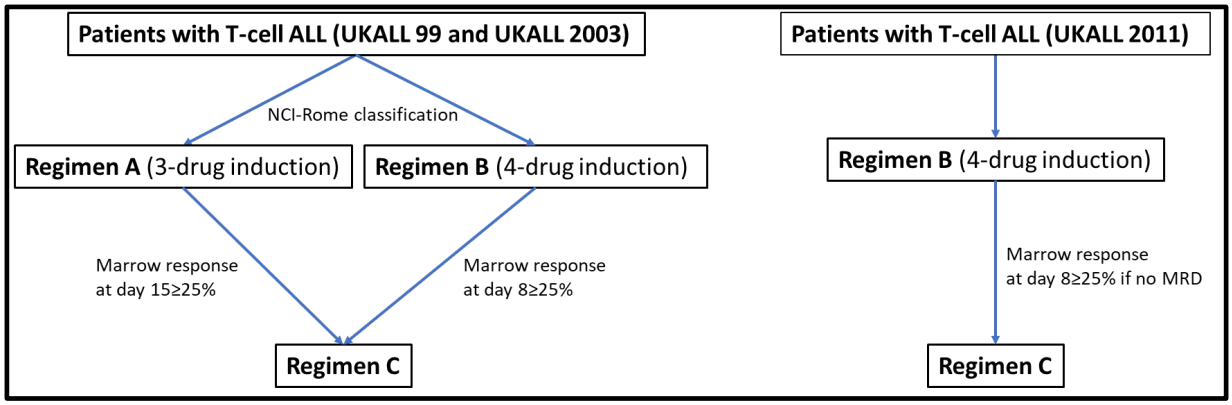

Figure 1.6. Simplified overview of treatment stratification based on NCI-Rome classification and early marrow response in UKALL 99, UKALL 2003 and UKALL 2011 (NCI-Rome, National Cancer Institute-Rome; MRD, minimal/measurable residual disease).

1.4.3.2. Minimal/Measurable Residual Disease

Minimal/measurable residual disease (MRD) is the presence of the leukemic cells detected by Polymerase Chain Reaction (PCR) or Flow Cytometry (FC). PCR and FC are the most common and sensitive (up to 10^{-5} cells) techniques for MRD detection. Gene fusions, such as *TAL1*, *TLX1*, *TLX3* and T-cell receptor rearrangements are detected by PCR, and immunophenotypic abnormalities (CD34, TdT, CD7, cytoplasmic CD3, CD1a) are detected by FC (Saygin *et al.*, 2022).

MRD is a widely accepted independent prognostic factor in T-cell ALL. Patients with higher MRD levels tend to have more relapse and worse clinical outcomes in comparison to those with low MRD levels (Willemse *et al.*, 2002; Berry *et al.*, 2017). Although there is no consensus on absolute MRD threshold and time point, cut-off level of 10^{-3} (0.1%) or 10^{-4} (0.01%) are usually selected for risk stratification, and the threshold of 10^{-4} (0.01%) is traditionally and historically accepted in most studies (Cazzaniga and Biondi, 2005). For instance, the AIEOP-BFM-ALL 2000 study treated T-cell ALL patients (n=464) using a MRD-derived risk stratification protocol between 2000 and 2006. Patients with negative MRD ($\leq 10^{-4}$) at two time points (TP1, end of induction at day 33; TP2, end of consolidation at day 78) were stratified to standard-risk group. Those with MRD $>10^{-4}$ at TP1 or MRD $<10^{-3}$ at TP2 were considered as intermediate-risk group, while MRD $\geq 10^{-3}$ at TP2 was accepted as poor clinical feature for high-risk group (Figure 1.7.) (Schrappe *et al.*, 2011). This study demonstrated that patients with negative MRD at the end of induction experienced excellent clinical outcome. Interestingly, those with negative MRD at only end of consolidation had also favourable outcome, regardless of MRD levels at the end of induction. Additionally, MRD levels at the end of consolidation were strongly associated with relapse incidence in the study. Another report investigating different MRD thresholds at the end of

induction in paediatric T-cell ALL patients treated with DFCI 05-001 and 11-001protocols (n=123) reported that MRD threshold of 10^{-4} (0.01%) could have more prognostic value than the cut-off level of 10^{-3} (0.1%) (5-year disease-free survival; 98% vs 84% for MRD threshold of 10^{-4} vs 94% vs 79% for MRD threshold of 10^{-3}) (Burns *et al.*, 2021). On the other hand, researchers have started focusing on MRD <0.005% in childhood ALL to reduce adverse side effects of the treatment (Goulden *et al.*, 2017).

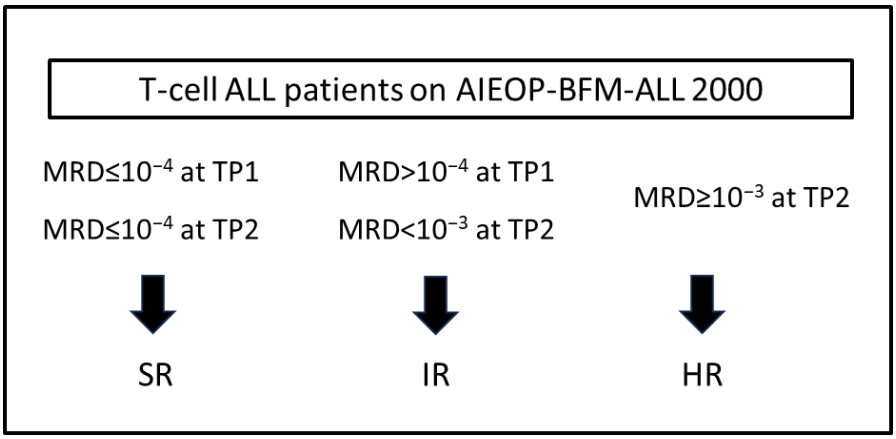

Figure 1.7. MRD-based risk stratification on AIEOP-BFM-ALL 2000 study (MRD, minimal/measurable residual disease; TP1, timepoint 1 at day 33; TP2, time point 2 at day 78; SR, standard-risk; IR, intermediate-risk; HR, high-risk).

1.4.4. Genetics

The biology of T-cell ALL is well established by advanced sequencing technologies, and one of the comprehensive studies investigating genomics of childhood T-cell ALL was published in 2017 (Liu *et al.*, 2017a). In this published article, 264 of paediatric T-cell ALL cases were categorised into the following genetic subgroups based on genetic alterations and gene expressions: *TAL1*, *TAL2*, *TLX1*, *TLX3*, *HOXA*, *LMO1/LMO2*, *LMO2/LYL1* and *NKX2-1*. Gene alterations in certain signalling pathways were found to be associated with the subgroups. For example, *PTEN* and *PIK3R1* mutations (PI3K signalling pathway) were the most common alterations in the *TAL1* subgroup. Another comprehensive study with >1000 participants identified 15 different subtypes based on gene expression data from Whole Transcriptome Sequencing: TAL1 DP-like, TAL1 αβ-like, NKX2-1, NKX2-5, LMO2 γδ-like, HOXA9 TCR, TLX1, TLX3, TME-enriched, STAG2/LMO2, SPI1, MLLT10, ETP-like, BLC11B, KMT2A (Pölönen *et al.*, 2024). Interestingly, deregulation of oncogenes in the subtypes was due to hijacking mechanisms in 50% of cases, and around 60% of driver alterations occurred in non-coding genome regions. In addition, certain signalling pathways were associated with specific subtypes (e.g. alterations in PI3K signalling pathway were mostly observed in NKX2-5 subtype). Although clinical outcomes of the subtypes and individual alterations had a correlation with poorer or better prognosis in this research, the

results should be validated by other research groups because translation of the genetic abnormalities into clinical use is still debatable due to the nature of this heterogenous disease.

1.4.4.1. Genetic Risk Classifications in T-cell ALL

In 2013, the French Group for Research in Adult Acute Lymphoblastic Leukemia (GRAALL) published a paper about oncogenetic risk classification of adult T-cell ALL patients (n=212), which was the first risk classifier based on genetic abnormalities in T-cell ALL. These patients were treated in GRALL-2003 (n=57) and GRALL-2005 (n=155) trials. This group reported that patients with *NOTCH1/FBXW7* (*N/F*) mutations had favourable outcomes, whereas those with *RAS/PTEN* alterations experienced poor prognosis. They also observed an interaction between *N/F* and *RAS/PTEN* mutations, meaning that good clinical impacts of *N/F* mutations were limited to wild type of *RAS/PTEN* mutations. Therefore, they developed an oncogenic classification based on *N/F* mutations and *RAS/PTEN* mutations, and adult T-cell ALL patients were classified into low-risk group (patients with *NOTCH1/FBXW7* mutations and without *PTEN/RAS* mutations) and high-risk group (the rest of patients). The high-risk group had worse clinical outcomes than the low-risk group, in terms of relapse-free survival rate (42% vs 85%, p <0.001) and overall survival rate (44% vs 82%, p <0.001) (Trinquand *et al.*, 2013).

Later, in 2016, Jenkinson and colleagues reported that this oncogenetic classification did not make a significant difference between the risk groups in paediatric UKALL 2003 trial (n=145), in relapse-free survival (87% vs 87%, p-value=0.95) and overall survival (84% vs 94%, p-value=0.06) (Jenkinson *et al.*, 2016). Another article comparing prognostic value of the oncogenic classification in paediatric trials, UKALL 2003 (n=156) and FRALLE 2000T (n=213), was published in 2022. The difference of cumulative relapse incidence was not statistically significant between the low-risk and high-risk groups in UKALL 2003 (12% vs 13%, p=0.77); however, patients in the risk groups showed a significant difference in FRALL 2000T (13% vs 41%, p-value <0.001) (Taj *et al.*, 2022). These findings highlight that differences in treatment protocols can affect the prognostic value of the oncogenetic classification.

In 2024, the same French group (GRAALL) has published a paper reporting a revised genetic classification using Next Generation Sequencing (NGS) method for T-cell ALL patients (Simonin *et al.*, 2024). Those who had favourable gene alterations (*NOTCH1/FBXW7* mutations, *EP300* mutations, or *PHF6* alterations) without adverse alterations (*N/K RAS* mutations, *PI3K* pathway alterations, *IKZF1* alterations, *TP53* alterations, *IDH1/2* mutations, and *DNMT3A* mutations) were classified into NGS-low-risk group, while cases with the

presence of the adverse gene alteration(s) or without mutated favourable genes were categorised into NGS-high-risk group. The NGS-low-risk patients were statistically superior to the NGS-high-risk patients in relapse rate and overall survival. Furthermore, they integrated white blood cell count and MRD into the NGS classification to identify favourable, intermediate, and adverse risk groups and showed a clear separation of these three risk groups in terms of relapse rate and overall survival. Although the new classifier is able to stratify patients into the risk groups, there is a need for validation of this classification in other T-cell ALL cohorts, such as UKALL cohorts.

1.4.4.2. Novel Prognostic Models Developed and Validated for ALL

Enshaei and colleagues have developed and validated a continuous prognostic index (PI_{UKALL}) using paediatric cohorts including B-cell ALL and T-cell ALL patients from UKALL 2003 (n=2405), NOPHO-ALL 2008 (n=1462), DCOG-ALL10 (n=592) and CoALL-07-03 (n=259) (Enshaei *et al.*, 2020). This index contains white blood cell count at diagnosis, cytogenetic abnormalities and MRD at the end of induction, and it predicts the risk of relapse regardless of sex, age and central nervous system involvement. All patients were assigned to the following clinical risk groups: low-risk group, standard-risk group, intermediate-risk group and high-risk group. The risk groups had 5-year relapse rates of 3%, 8%, 17%, 48% in discovery cohort and 4%, 9%, 17%, 35% in validation cohort, respectively. Although T-cell status was not included in the index, researchers mentioned that T-cell status had a strong correlation with the key predicters in the model and subgroup analysis confirmed its applicability in T-cell ALL and B-cell ALL, explaining the exclusion of T-cell status from the index.

Recently, Children's Oncology Group built and validated a prognostic model consisting of clinically known prognostic factors (age, central nervous system involvement, white blood cell count at diagnosis, favourable/unfavourable cytogenetics, MRD at day 8 and day 28) for children and young adults with B-cell acute lymphoblastic leukaemia from the trials of AALL0331 (n=5099), AALL0232 (n=2900), AALL0932 (n=8776) and AALL1131 (n=4424). The COG model stratified patients into low-risk group, standard-risk group, intermediate-risk group and high-risk group, and their relapse-free survival rates at five years were 97%, 93%, 85% and 67% in training cohort, respectively, and these rates were similar in testing cohort (DelRocco *et al.*, 2024).

1.5. Prognostic Models

1.5.1. Overview of Methods Commonly Used in Medicine

Prognostic models containing key variables can predict future clinical outcomes, which are important for treatment decision during therapy. Identifying key variables is a multi-step process, consisting of identification of candidate variables and selection of key predictors through appropriate criteria or stopping rule strategy. Logistic regression for binary outcomes, the proportional hazards regression (Cox regression) for time-to-event outcomes, and classification trees for binary and time-to-event outcomes are commonly used to identify important prognostic factors and to develop a prognostic model (Halabi and Owzar, 2010).

1.5.1.1. Logistic Regression

Logistic regression is a machine learning technique to identify a relationship between dependent variable (binary outcome) and independent (predictor) variable. The following assumptions should be met when this method is used: 1) the dependent variable should be binary, such as yes/no or death/alive; 2) independent variables should not be repeated measures, matched or clustering data; 3) independent variables should not be highly correlated with each other (no multicollinearity); 4) independent variables should have a linear relationship with log-odds of predicted probabilities for dependent variable of interest.

The formula/function of logistic regression is illustrated in Figure 1.8. The formula generates the probability of the dependent variable (p(y)), using the independent variable (x) and the coefficient of the independent variable (β). The probability function of logistic regression has a S-shaped curve with binary dependent and continuous independent variables. In the literature, the results of logistic regression functions are shown with odds ratio (OR) and confidence interval (CI) instead of probability score. OR is the change in the odds of having the dependent variable (outcome) per one-unit change in the independent variable (ratio of two odds). If OR is equal 1, there is no impact of the independent variable on the dependent variable (Harris, 2021).

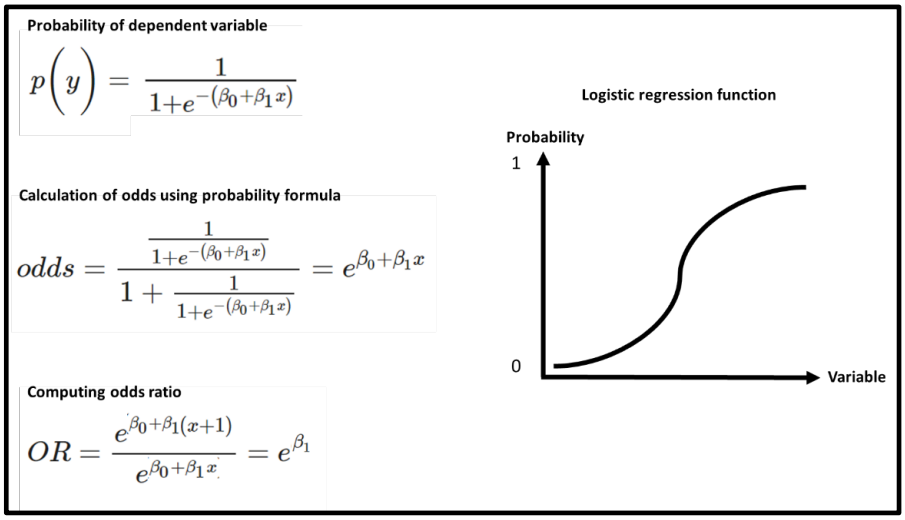

Figure 1.8. The formula and function of logistic regression (p(y), probability of y; x, independent variable; β, the coefficient of the independent variable x).

1.5.1.2. Proportional Hazards Regression (Cox Regression)

Proportional hazards regression (Cox regression) is a method to use in time-to-event analysis in epidemiology or clinical studies. The purpose of using this method is to investigate a relationship between exposure and outcome at a given follow-up time. This statistical technique is a semiparametric method because it does not require an assumption for time distribution. However, the impacts of the variables of interest on survival should be constant over time (the assumption of proportionality).

The result of proportional hazards regression is shown with the score of hazard ratio and confidence interval. Hazard ratio is the ratio of hazard rates corresponding to the events of interest in two groups or conditions. The expected hazard rate at time t (H(t)) is calculated with the formula in Figure 1.9., using baseline hazard ($H_0(t)$, the hazard when all independent variables are equal 0) and coefficient (β) of the independent variable (x). Hazard ratio=1 means there is no difference in survival between two groups or conditions. While hazard ratio of <1 or hazard ratio >1 indicates survival advantage or disadvantage of one group/condition, respectively (Abd ElHafeez *et al.*, 2021).

$$H(t) = H_0(t) \times \exp[\beta_1 x_1 + \beta_2 x_2 + \cdots .. \beta_k x_k]$$

Figure 1.9. Cox regression function (H(t), hazard rate; $H_0(t)$, baseline hazard; β, coefficient, x, independent variable).

1.5.1.3. Classification Tree

Classification tree is a nonparametric technique and defines homogeneous subsets of patients regarding dependent variables (outcome), using independent variables (predictors).

Independent variables can be categorical or continuous data. This method is also able to handle missing values by selecting optimal split points based on available data (no need to exclude missing data). Binary recursive partitioning is commonly used to classify patients until there is no improvement in user specified criteria (Song and Lu, 2015). Leaves represent subsets of patients, whereas branches illustrate independent variables with specific criteria in Figure 1.10.

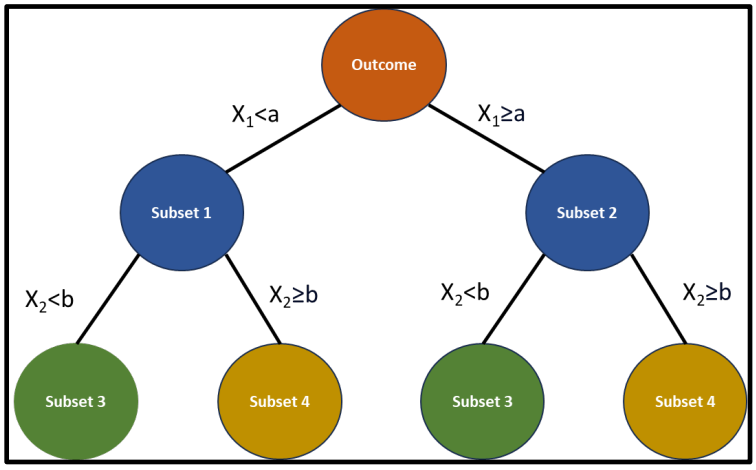

Figure 1.10. Simplified steps of classification tree algorithm (x, independent variable; a or b, thresholds).

1.5.2. Examples of Prognostic Models in Cancer

Prognostic modelling (or index) is a mathematical tool that predicts specific clinical outcomes and improves treatment decision. Regression models are widely used in medical research, and coefficients from the regression methods are useful metrics to build a prognostic model. For example, Liverpool Lung Project (LLP) model was developed using logistic regression method to estimate a risk of developing lung cancer and to identify high-risk patients over the next five years. The first version of this model (LLPv1) has contained sex, age, smoking history, previous diagnosis of pneumonia, occupational exposure to asbestos, previous diagnosis of malignant tumour, and family history of lung cancer (Cassidy *et al.*, 2008). The model has been validated in three international studies from Europe and North America, and researchers have concluded that LLPv1 can be useful to select patients with high-risk features in lung cancer screening using computed tomography (Raji *et al.*, 2012). Later, this model was updated to a second version (LLPv2) with observations from the UK Lung Cancer Screening Trial (UKLS). Other respiratory factors (bronchitis, emphysema, tuberculosis and chronic obstructive pulmonary disease) and other forms of smoking (cigar and pipe) were included in this second version. LLPv2 model has been currently using in the National Health Service (NHS) Targeted Lung Health check programme as a screening tool (NHS, 2022). Recently, the third version of the model (LLPv3) has been developed to improve

overestimation of the second version of the model (Field *et al.*, 2021). Another example is the PREDICT model (version 1) for the survival prediction of breast cancer. In 2010, this model was developed using Cox regression technique with data from patients in the UK (Wishart *et al.*, 2010), and it was validated in different national cohorts (de Glas *et al.*, 2016; Wong *et al.*, 2015). The first version of the model used the following predictors: tumour size (categorical), node status (categorical), tumour grade, oestrogen receptor status and mode of detection. In later versions, human epidermal growth factor 2 status and Ki-67 status were added to the model (Wishart *et al.*, 2014; Down *et al.*, 2014). In the latest version (version 2.2.), age at diagnosis has been included, and categorical predictors, such as tumour size and node status have been converted to numerical predictors (PREDICT, 2024).

As discussed earlier, continuous prognostic index (PI_{UKALL}) was built to stratify paediatric patients with ALL according to the risk of relapse for treatment allocation. Cox regression method was used to develop this prognostic model, and it contained white blood cell count at diagnosis, cytogenetic abnormalities and MRD at the end of induction (Enshaei *et al.*, 2020). Similarly, another prognostic model was created to stratify childhood B-cell ALL patients into risk groups, using a Cox regression method. Age, central nervous system involvement, white blood cell count at diagnosis, favourable/unfavourable cytogenetics, MRD at day 8 and day 28 have been included in the model (DelRocco *et al.*, 2024). In addition to these studies, other research groups have applied state-of-the-art machine learning techniques for the identification of paediatric ALL patients with high-risk features and prediction of the use of cranial radiotherapy treatment (Kashef, Khatibi and Mehrvar, 2020; Pan *et al.*, 2017). The oncogenetic model developed by GRAALL using Cox regression method is the only risk classifier that has been applied across multiple cohorts for T-cell ALL patients (Trinquand *et al.*, 2013). However, this classifier has failed to validate in the paediatric and young adult UKALL cohort (Jenkinson *et al.*, 2016; Taj *et al.*, 2022). Also, NGS-risk model defined by favourable/adverse gene alterations has been built using a least absolute shrinkage and selection operator (LASSO) penalization technique, and it needs external evaluations in different T-cell ALL populations for clinical applicability (Simonin *et al.*, 2024).

It is clearly seen that most research groups working on the development of prognostic models have focused on ALL or B-cell ALL, not only T-cell ALL; therefore, there is no robust and validated prognostic model for T-cell ALL so far.

1.6. Aims and Objectives

Recently, clinical outcomes of patients with T-cell ALL have dramatically improved because of contemporary treatment and minimal/measurable residual disease (MRD) stratification. However, there are still some issues to be solved in T-cell ALL: 1) there is no commonly used key predictors other than MRD and no consensus on absolute threshold and time point for MRD; 2) genetic abnormalities are still controversial in the context of clinical use; 3) prognostic models focus ALL or B-cell ALL, not specifically on T-cell ALL; 4) clinical trials working on targeted drugs are still in progress. With these observations taken together, it is vital to identify additional risk and prognostic factors that improve upon existing MRD prediction in T-cell ALL and make treatment stratification more effective.

Consecutive and comprehensive paediatric and young adult T-cell ALL trials (from UKALL VIII to UKALL 2011) were available to retrospectively investigate potential risk and prognostic factors with an unbiased view. Hence, the purposes of this thesis are to better understand, evaluate and validate risk and prognostic factors to identify key predictors for better clinical estimation and to make treatment decision more effective and useful for T-cell ALL patients.

Objectives of this thesis are to:

1) Investigate prognostic impacts of traditional demographics and clinical risk factors (age, sex, white blood cell count at diagnosis, NCI-Rome classification, central nervous system involvement, organomegalies, marrow response and MRD response, and genetic subtypes) in serial clinical trials with different treatment regimens and randomisations.

2) Develop and validate baseline prognostic models with key predictors identified by selection methods among available single-time point variables and dynamic prognostic models using treatment response variables (marrow and/or MRD responses) measured at multiple timepoints in successive UKALL trials.

3) Examine T-cell ALL genetics using supervised and unsupervised bioinformatic tools.

Chapter 2. Data and Methods

2.1. Data Management

2.1.1. Data Collection and Manipulation

Data was obtained from multiple paediatric UK trials for this thesis (Table 2.1.). Individual patient data for cases treated on UKALL VIII, UKALL X, UKALL XI, UKALL 97/99, UKALL 2003 and UKALL 2011 were collected by the Leukaemia Research Cytogenetics Group (LRCG) from the clinical trial unit that conducted the trial and the National Health Service (NHS) laboratory that performed the diagnostic genetic testing. Additional genetic data was generated by the LRCG as part of previous research projects. The original clinical trial data files were stored securely by the LRCG in either in a bespoke online database or as standalone excel files. The LRCG stores all routine and research genetic and genomic data in a bespoke database called Cytogenetics Information Management System (CIMS). All data used within this thesis was extracted directly from these sources and integrated into a single data-frame. Both UKALL XI and UKALL 97/99 underwent significant modifications during recruitment which affected outcome. Therefore, the following datasets were considered throughout this thesis: UKALL VIII, UKALL X, UKALL XI (UKALLXI+UKALLXI92), UKALL 97/99 (UKALL97+UKALL99), UKALL 2003 and UKALL 2011 (Table 2.1.).

UKALL Trials	Time	Total Patient Number (B-cell ALL / T-cell ALL / unknown)	Age (years)	Median follow-up time for T-cell ALL (years)
UKALL VIII	1980-1984	829 (553 / 58 / 218)	<15	17
UKALL X	1985-1990	1612 (1397 / 139 / 76)	<15	17
UKALL XI	1990-1997	2090 (1425 / 196 / 469)	<15	14
UKALL 97/99	1997-2002	1934 (1676 / 209 / 49)	<18	10
UKALL 2003	2003-2011	3112 (2726 / 386 / 0)	<25	11
UKALL 2011	2012-2018	2554 (2233 / 321 / 0)	<25	5

Table 2.1. Summary information about available paediatric/young adult patients on the UKALL trials (ALL, acute lymphoblastic leukaemia).

The variables of sex, age, white blood cell count (WBC) at diagnosis and central nervous system (CNS) involvement, and clinical events (remission, relapse, second tumour, death) are available in all paediatric/young adult trials. UKALL 97/99, UKALL 2003 and UKALL 2011 have early marrow response variables. MRD values are unknown in early trials (UKALL VIII to UKALL 97/99), while late trials (UKALL 2003 and UKALL 2011) do not have organomegaly information (Table 2.2.) Genetic data are accessible for UKALL 97/99 and UKALL 2003.

	UKALL VIII	UKALL X	UKALL XI	UKALL 97/99	UKALL 2003	UKALL 2011
Sex	A	A	A	A	A	A
Age (years)	A	A	A	A	A	A
WBC (10^9/L)	A	A	A	A	A	A
CNS involvement	A	A	A	A	A	A
Mediastinal mass	A	A	A	A	N/A	N/A
Splenomegaly	A	A	A	A	N/A	N/A
Hepatomegaly	A	A	A	A	N/A	N/A
BM blast % at diagnosis	A	A	A	A	A	N/A
BM blast % at day 8	N/A	N/A	N/A	A	A	A
BM blast % at day 15	N/A	N/A	N/A	A	A	N/A
BM blast % at day 28	N/A	N/A	N/A	A	A	A
MRD % at day 28/29	N/A	N/A	N/A	N/A	A	A
Date of remission	A	A	A	A	A	A
Date of relapse	A	A	A	A	A	A
Date of second tumour	A	A	A	A	A	A
Date of death	A	A	A	A	A	A

Table 2.2. Availability of the variables in paediatric/young adult UKALL trials for T-cell ALL. BM blasts were morphologically measured using light microscope, whereas MRD is measured using PCR technique in all trials (WBC, white blood cell count; CNS, central nervous system; BM, bone marrow; MRD, minimal/measurable residual disease; A, available; N/A, not available).

After obtaining the data from the UKALL trials, file format was converted from excel format (.csv) to STATA format (.dta). STATA files (.do) were created to save STATA commands for data manipulation (format files.do) and analysis (analysis files.do). Data dictionary was used to label available values and variables. The values of available variables were also manipulated (re-coded) for the analysis if necessary. For example, organomegalies (mediastinal mass, splenomegaly, hepatomegaly) were categorised into two subgroups (yes/no) based on the presence/absence of enlargement in relevant organs. Available data from UKALL VIII to UKALL 2011 were checked before the analysis, so available variables and dates of clinical events (dates of remission/relapse/second tumour/death) were controlled if they had incorrect/illogical information (Figure 2.1.). Infant T-cell ALL patients were excluded from all analyses as there were no enough cases for further investigation (n=1). Each trial had one dataset containing all available variables. The data from trials with modified treatment protocols during their active time were merged for the analyses, such as UKALL XI (UKALLXI+UKALLXI92) and UKALL 97/99 (UKALL97+UKALL99).

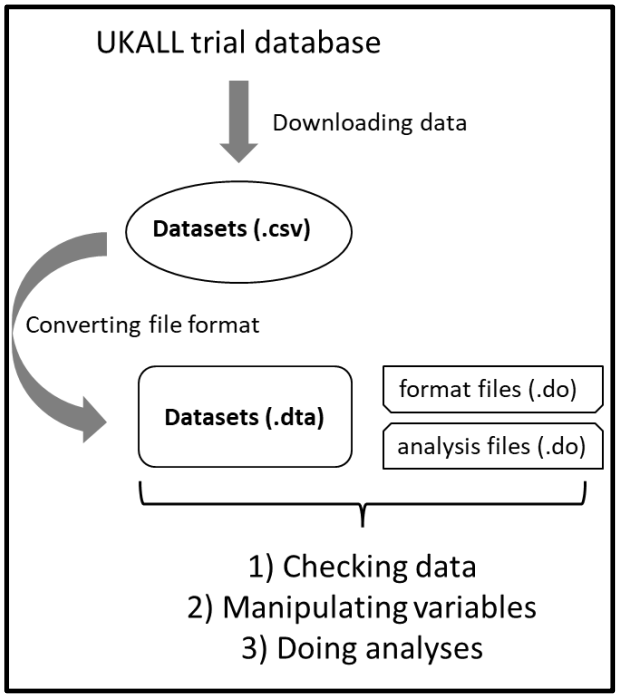

Figure 2.1. Overall steps for data check and data manipulation in the UKALL datasets.

2.1.2. Information about the UKALL Trials

Newly diagnosed paediatric B-cell and T-cell ALL patients were accepted onto UKALL VIII trial between 1980-1984. The main objectives of this trial were to adapt treatment regime used in North America during that time and to also determine the effect of additional daunorubicin in induction and maintenance phases. The therapy mainly consisted of induction phase containing CNS-directed therapy, and maintenance phase (Figure 2.2). The study concluded that daunorubicin could improve disease-free survival despite the increased toxicity, and 3-year maintenance did not show any survival advantage in comparison to 2-year maintenance (Eden *et al.*, 1991).

Children with ALL were registered to UKALL X trial from 1985 to 1990. The aim of this study was to evaluate benefit of additional intensification blocks given at early stage (week 5), late stage (week 20), or early + late stages during the therapy. There were induction phase containing daunorubicin, intensification phase, CNS-directed therapy, and maintenance phase in UKALL X treatment protocol (Figure 2.2.). Researchers reported that the addition of two blocks of intensification therapy (early + late intensifications) could be beneficial in terms of disease-free survival outcome (Chessells, Bailey and Richards, 1995).

Paediatric ALL patients were entered into UKALL XI trial between 1990 and 1997. This trial had three main aims: to test CNS-directed chemotherapy without radiotherapy, to understand on whether or not high dose intravenous methotrexate had to be part of CNS-directed therapy and to assess the third intensification block at week 35 (Figure 2.2.). Two years after UKALL

XI started, in 1992, treatment protocol was revised according to initial UKALL X results. The changes were that daunorubicin omitted from the induction phase because of toxicity concerns and randomisation of the third intensification was added to the protocol. This study indicated that patients with WBC $\leq 50x10^9$/L would be cured without radiotherapy, and those with WBC $\geq 50x10^9$/L could be treated with continuing intrathecal methotrexate alone (ITMTX) or high-dose intravenous methotrexate (HDMTX) instead of radiotherapy (Hill *et al.*, 2004). In addition to these findings, the third intensification block would be beneficial to all patients enrolled in the trial (Hann *et al.*, 2000).

ALL patients were recruited onto UKALL 97/99 trial from 1997 to 2002. The important changes in the trial were about two randomisations for dexamethasone vs prednisolone, and thioguanine vs mercaptopurine (Figure 2.2.). Patients in high-risk group treated with HR1 protocol were identified by chromosomal abnormalities (near haploidy, Philadelphia positive or *KMT2A* (*MLL)* rearrangements) or Oxford regression score, which was based on sex, age, and WBC at diagnosis and applied to both B-cell ALL and T-cell ALL. However, original UKALL 97 treatment protocol was modified for the use of National Cancer Institute-Rome (NCI-Rome) Criteria in 1999, considering age, WBC, and early marrow responses at day 8 and day 15 for treatment allocation. With the revised protocol, UKALL 99 protocol, ALL patients were categorised to treatment groups for regimen A (3-drug induction), regimen B (4-drug induction) and regimen C (augmented treatment). Those in high-risk group, determined by certain chromosomal abnormalities and slow early marrow responses, were treated with regimen C after induction phase. The study indicated that the use of dexamethasone over prednisone made relapse risk decrease in all risk groups, and randomisation group with thioguanine experienced long-term side effect of hepatotoxicity (Vora *et al.*, 2006; Mitchell *et al.*, 2005). Furthermore, the risk stratifications used in UKALL 99 dramatically improved the clinical outcomes of ALL patients in terms of event-free survival and overall survival compared to UKALL X and UKALL XI trials (Mitchell *et al.*, 2009).

Paediatric and young adult patients were enrolled on UKALL 2003 trial between 2003 and 2011. The rationale of this trial was to use MRD risk stratification for the reduction and intensification of the treatment. Patients in MRD low-risk group (MRD <0.01%) were randomised to receive one versus two intensification blocks, whereas MRD high-risk group (MRD ≥0.01%) was randomly treated with standard therapy versus augmented therapy (Figure 2.2). Similar to the UKALL 99 protocol, ALL patients were allocated into three different treatment groups, which were regimen A (3-drug induction), regimen B (4-drug induction) and regimen C (augmented treatment), according to NCI-Rome classification,

chromosomal abnormalities and early marrow responses. The research demonstrated that MRD response was the best available method for risk-adapted therapy, and the reduction of treatment could be possible for patients in the MRD low-risk group. Furthermore, those in the MRD high-risk group can have benefit of augmented therapy after remission (Vora *et al.*, 2014; Vora *et al.*, 2013; Samarasinghe *et al.*, 2021).

Children and young adults with ALL were entered into UKALL 2011 trial from 2012 to 2018. This clinical trial aimed to reduce treatment toxicity in induction phase through short versus standard courses of dexamethasone, to decrease relapse risk of central nervous system through high dose methotrexate versus standard interim maintenance, and to minimise treatment burden in maintenance phase through pulses versus no pulses of vincristine/dexamethasone (Figure 2.2.). Although NCI-Rome risk classification, cytogenetic abnormalities and MRD-based risk stratification were used for treatment allocation of patients into three different regimens (Regimen A, Regimen B and Regimen C), all T-cell ALL patients were considered as NCI-Rome high-risk group and treated with regimen B during induction phase. Therefore, this was the first T-cell ALL specific intervention in the UK trials. Early marrow response was evaluated in those without MRD results, and slow early responders received regimen C. Also, T-cell ALL patients with induction failure went off trial and were recommended for a treatment with nelarabine. Preliminary analysis revealed that there was no statistically significant difference between short and standard dexamethasone arms in terms of treatment toxicity, MRD levels and relapse-free survival rates, and high dose methotrexate was not beneficial to reduce relapse risk of CNS (Kirkwood *et al.*, 2022).

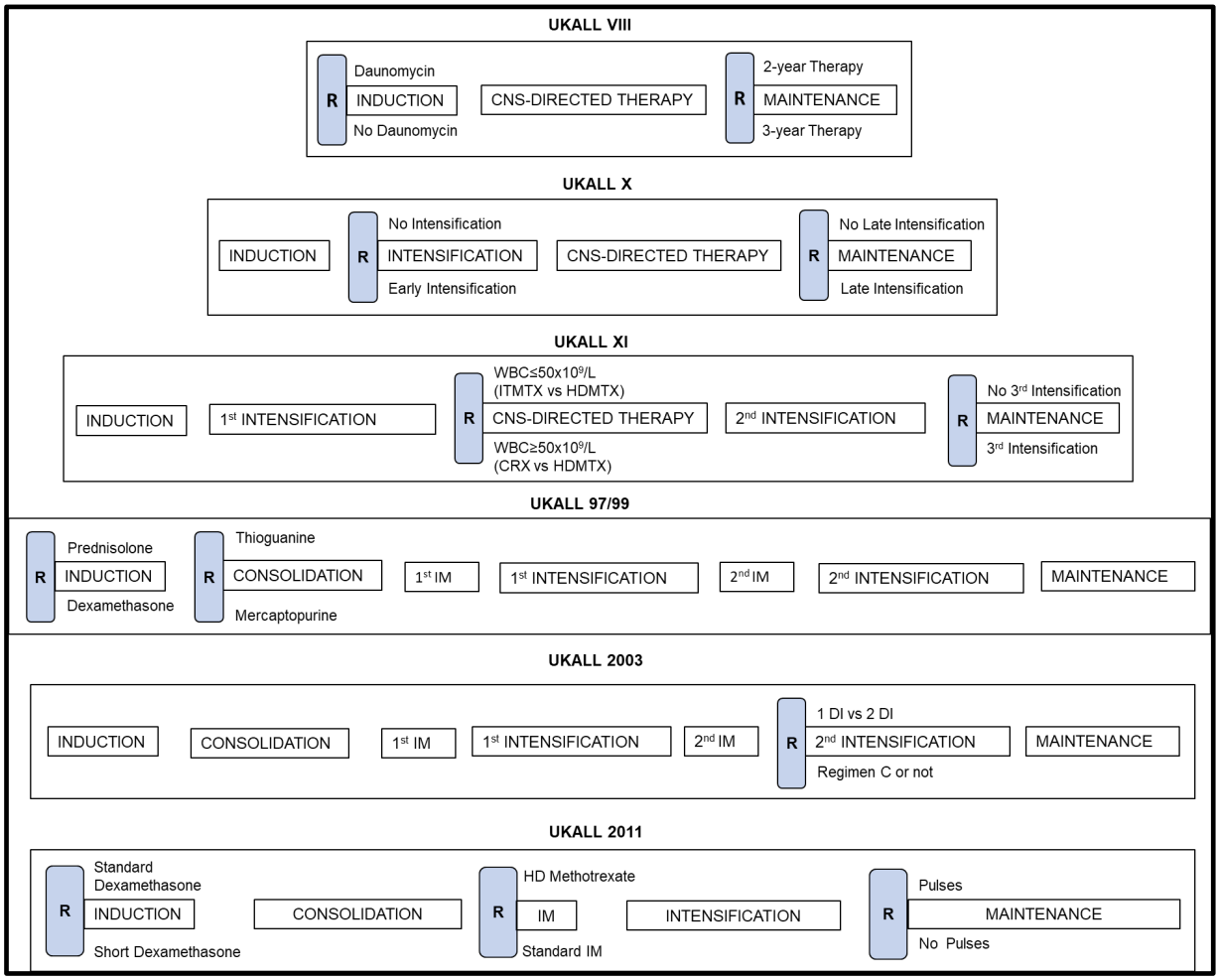

Figure 2.2. Treatment phases and randomisations of UKALL VIII, UKALL X, UKALL XI, UKALL 97/99, UKALL 2003 and UKALL 2011 from 1980 to 2018. (R, randomisation; CNS, central nervous system; WBC, white blood cell count; ITMTX, intrathecal methotrexate; HDMTX, high-dose intravenous methotrexate; CRX, cranial radiotherapy; IM, interim maintenance; DI, delayed intensification).

2.1.3. Manipulation of Genetic Data

As previously mentioned, genetic data was extracted from the Cytogenetics Information Management System (CIMS). This data was generated with different laboratory methods, such as Fluorescence in Situ Hybridization (FISH), Multiplex Ligation-dependent Probe Amplification (MLPA), molecular and cytogenetics techniques. First, known abnormalities (chromosomal translocations/gene fusions, mutations, deletions) were screened separately using FISH, MLPA, molecular, and cytogenetic data, and patients were classified as positive cases if any of these abnormalities detected by the laboratory methods. Next, patients without specific abnormalities observed by any detection method were called negative cases. After labelling positive and negative cases, all patients with or without abnormalities were merged into one dataset (Figure 2.3.). Microsoft Access software (version 2402 build 16.0.17328.20346) was used for the purpose of screening and categorising the genetic

abnormalities in UKALL 97/99 and UKALL 2003. There was no enough genetic data in UKALL 2011 dataset.

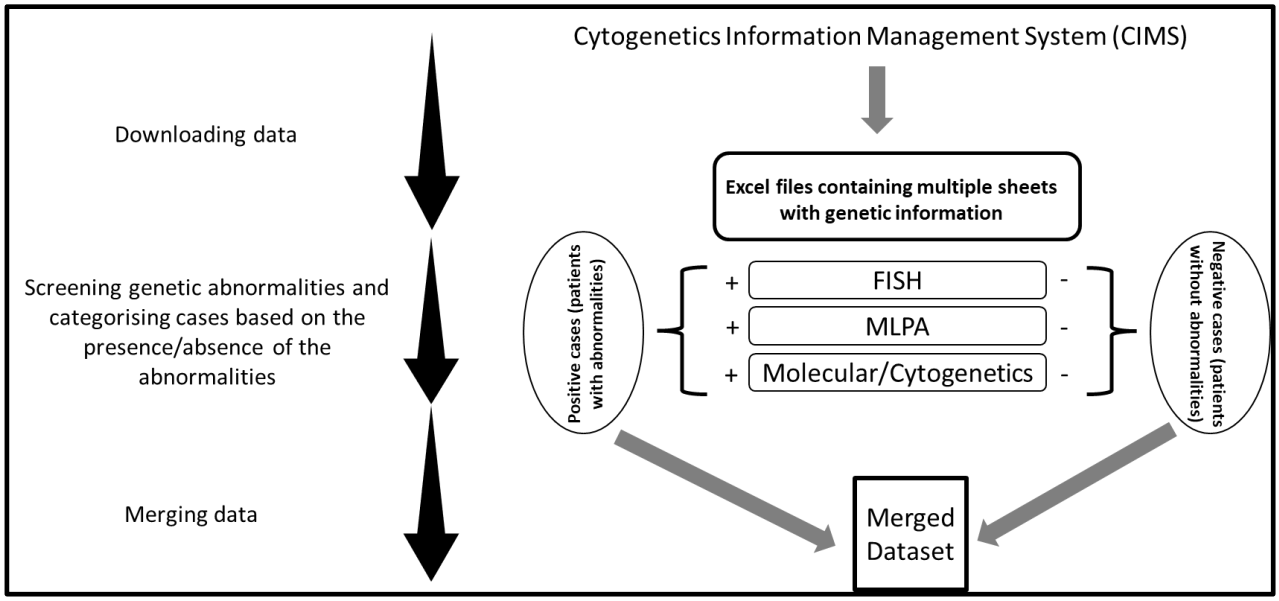

Figure 2.3. Classification of patients according to the presence/absence of known genetic abnormalities detected by different laboratory methods (FISH, Fluorescence in Situ Hybridization; MLPA, Multiplex Ligation-dependent Probe Amplification).

To increase sample size of each genetic subtype, genetic abnormalities in UKALL 97/99 and UKALL 2003 were narrowed to four main subtypes according to a research using an unsupervised hierarchical clustering analysis (Homminga *et al.*, 2012): These subtypes were: 1) immature subtype, 2) TLX3 subtype, 3) proliferative subtype and 4) TAL/LMO subtype. Patients with early T-cell precursor (ETP) status detected by Patrick and colleagues using data from UKALL 2003 (Patrick *et al.*, 2014a), *JAK1/3* mutations, *N/K RAS* mutations, *IL7R* alterations, *ETV6* mutations or *IKZF1* alterations were stratified into immature subtype. Cases with *TLX3* rearrangements was considered as TLX3 subtype. Proliferative subtype was defined as the presence of specific CD markers ($CD1a^+$ & $CD4^+$ & $CD8^+$) or any of the following abnormalities: *TLX1* rearrangements, *NKX2.1* rearrangements, *LEF1* alterations, *BCL11B* alterations. *TAL1/2* rearrangements, *LMO1/2* rearrangements, *LYL1* rearrangements, *PTEN* alterations, or $CD1a^-$ & single positive cells ($CD4^{+/-}$ and $CD8^{-/+}$) were the indicators of TAL/LMO subtype. Figure 1.5. summarises the gene rearrangements, CD markers and gene alterations within the subtypes (see more detail information in section 3.3.4.).

2.2. Data Analysis

2.2.1. Methodology of Model Development and Validation

2.2.1.1. Transformation of Continuous Variables

Transformation is a replacement of continuous values using mathematical functions. The aim of transformation in statistics is to help handle and interpret data easier. Transformation can be used for rescaling values, making equal variability between units, and changing distribution of skewed data to normal distribution via reflection and mathematical formulas (square root, cube root, logarithm etc.).

Logarithmic function was used for the variables of WBC and bone marrow (BM) blast percentage at diagnosis. BM blast percentage at diagnosis was highly skewed; therefore, the values were first reflected and then transformed using logarithmic function. On the other hand, square root function was performed for marrow response variables at day 8, day 15 and day 28. Transformation techniques were utilised before developing and validating prognostic models.

2.2.1.2. Variable Selection Strategy

In the literature, clinical knowledge and statistical methods are used to identify candidate factors for model selection process. After deciding candidate variables, appropriate selection methods with criteria/stopping rule can be utilised to develop final prognostic model. The main purposes of using selection methods to find most relevant variables are to easily interpret a model and to practically use it in the clinics. Although there is no consensus about the number of variables in a prognostic model, the commonly accepted rule is “one in ten rule”, stating that each variable in a model should be investigated for at least 10 clinical events in regression analysis (Chowdhury and Turin, 2020).

Backward and forward selection methods based on Cox regression are traditional techniques for the model development. Backward selection starts with all variables (full model), and they are eliminated step-by-step approach using significance level of a test comparing other models in the absence of previous variables. In forward selection, each candidate variable is added into null model (no variables) step-by-step by starting with the most significant variable until obtaining key variables that depend on the significance level of a statistical test (Royston *et al.*, 2009) (Figure 2.4.).

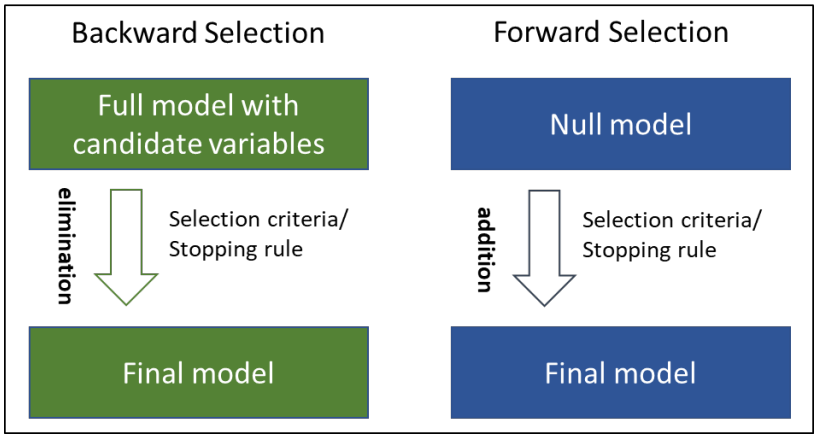

Figure 2.4. Backward and forward selections with criteria/stopping rule. Full model indicates a model with all candidate variables, whereas null model means there is no variable in a model.

In this thesis, univariate Cox regression was performed to identify candidate prognostic variables, and *p*-value of less than 0.05 was considered statistically significant. Sex, age and WBC at diagnosis were included in the model selection process due to literature search and expert knowledge. However, it should be noted that sex, age and WBC did not have statistical significance in the UKALL cohorts. Backward and forward selection techniques based on Cox regression and likelihood-ratio test were used to select key variables for final prognostic model, depending on significance level of 0.05. If the *p*-value was statistically significant after addition or elimination of variables of interest, the variables remained in the model for the next step. Otherwise, they were removed from model selection process in the next step. Interactions between the variables during the modelling process were checked using Cox regression with likelihood-ratio test.

2.2.1.3. Developing Prognostic Indexes and Identifying Optimal Thresholds

Prognostic indexes of baseline models were developed with coefficients from Cox regression, while dynamic model indexes were generated using coefficients of treatment response trend from Nonlinear Mixed-Effects (NLME) regression and Area Under the Curve (AUC) score for repeated measurements. After developing prognostic indexes, clinically relevant thresholds to categorise patients into standard-risk and high-risk groups were defined according to the proportion of the risk groups and relapse rates. Approximately 80% of patients were assigned to the standard-risk group, whereas the rest (~20%) was considered as the high-risk group. Doubled relapse rates between the risk groups was also a criterion to define optimal thresholds for baseline and dynamic prognostic models.

2.2.1.4. Model Validation, Calibration and Evaluation Metrics

There are two types of validation methods that are internal validation and external validation. In the internal validation, a dataset is used for both model development and validation processes with resampling techniques (cross-validation, bootstrapping etc.). In contrast, the external validation indicates that model development and validation steps should be independently performed in different datasets (for example, discovery cohort for model development and validation cohort for model validation). In this thesis, consecutive paediatric/young adult UKALL trials, which had different treatment schedules and randomisations, were mainly used for validation purposes. Also, a publicly available dataset, TARGET, was employed to validate the prognostic model. Cox regression and Kaplan-Meier graphs were performed to check model validation.

Calibration refers a correlation between observed and predicted outcomes of the prognostic models. A graphical assessment and calibration slope score are useful to check calibration accuracy. Calibration of the prognostic models was evaluated with a graph showing observed versus predicted outcomes with a distribution pattern of patients and slope score at five years in validation cohorts. Slope test performs against null hypothesis (p-value >0.05 for slope score=1 means adequate calibration).

Area Under the Receiver Operating Characteristic (AUROC) and C-index were used to performed to evaluate the discrimination ability of relapse/nonrelapse cases and to compare the models, respectively.

2.2.1.5. Missing Values and Imputation Methods

Missing data can exist in clinical studies. There are three types of missing data: 1) Missing Completely at Random (MCAR), indicating no systematic differences between missing data and observed data; 2) Missing at Random (MAR), pointing that systematic differences can be associated with other existing variables, but not with unobserved variables; and 3) Missing not at Random (MNAR), denoting that systematic differences cannot be explained by other existing variables; however, they are related to unobserved values. MCAR and MAR are suitable to use imputation methods to avoid the risk of bias (Ibrahim, Chu and Chen, 2012).

Missing values were checked in the trials of UKALL VIII, UKALL X, UKALL XI, UKALL 97/99, UKALL 2003 and UKALL 2011 (Table 2.3.). There was no high missing data (>5%) in available variables in UKALL VIII and UKALL X. Bone marrow (BM) blast percentage at diagnosis contained high missing values (11%) in UKALL XI. BM blast percentage at diagnosis (17%), BM blast percentage at day 8 (8%), and BM blast percentage at day 28

(37%) had high missingness levels because of HR1 treatment group in UKALL 97/99. UKALL 2003 had high missing data in the variables of BM blast percentage at diagnosis (24%), BM blast percentage at day 8 (16%), BM blast percentage at day 15 (77%), BM blast percentage at day 28 (13%) and minimal/measurable residual (MRD) disease (27%). Similar missing patterns of treatment response variables in UKALL 2003 can be seen in UKALL 2011.

	UKALL VIII (%)	UKALL X (%)	UKALL XI (%)	UKALL 97/99 (%)	UKALL 2003 (%)	UKALL 2011 (%)
Sex	0	0	0	0	0	0
Age (years)	0	0	0	0	0	0
WBC (10^9/L)	0	0	0	0	0	1
CNS involvement	5	0	0	0	0	11
Mediastinal Mass	2	4	3	1	N/A	N/A
Splenomegaly	0	1	1	0	N/A	N/A
Hepatomegaly	0	1	1	1	N/A	N/A
BM blast % at dx	5	5	11	17	24	N/A
BM blast % at d8	N/A	N/A	N/A	8	16	13
BM blast % at d15	N/A	N/A	N/A	78	77	N/A
BM blast % at d28	N/A	N/A	N/A	37	13	12
MRD % at d28/29	N/A	N/A	N/A	N/A	27	21

Table 2.3. Percentage of missing values in the UKALL trials (WBC, white blood cell count, CNS, central nervous system; BM, bone marrow; MRD, minimal/measurable residual disease; dx, diagnosis; d8, day 8; d15, day 15; d28, day 28; d28/29, day 28 or day 29).

K^{th}-Nearest Neighbours (KNN) imputation is one of the imputation techniques to handle missing values. It is based on distance between data points with missing values and other available data. KNN imputation was performed for marrow variables with missing values in which prognostic models were developed. Bone marrow blast percentage at diagnosis was not collected in UKALL 2011. This variable was generated with random values ranging 90-99 as most of patients on UKALL 2003 had values between 90 and 99, so this range was selected to impute missing values in UKALL 2011.

2.3. Bulk RNA-seq Analysis

Researchers at Illumina centre in Cambridge measured RNA quality, prepared libraries, and sequenced T-cell ALL samples (as mentioned below) in 2018 and 2024.

RNA samples from UKALL 2003 were already sequenced in 2018 (n=29). RNA quality was estimated on the Agilent 2100 bioanalyzer (Agilent Technologies). Paired end, poly(A) enriched, stranded RNAseq was performed with the True Seq Stranded mRNA Sample Preparation Kit (Illumina). Libraries were sequenced on a NextSeq500 sequencer (Illumina) with a read length of 75 base-pairs. RNA-seq reads were aligned to the GRCh38 and quantified on Gencode v.24 using STAR v.2.4.2a with parameter quantMode GeneCounts.

Available RNA samples from additional UKALL 2003 cases and UKALL 2011 cases were sent to Illumina Centre in Cambridge in January 2024 (n=97). The quality and concentration of RNA were measured by Agilent 4200 Tapestation System and GeminiTM XPS Microplate Spectroflurometer, respectively. Libraries were prepared using Illumina Stranded Total RNA Prep, Ligation with Ribo-Zero Plus, with additional custom depletion probes (100-400 ng RNA) and were sequenced on a NovaSeq 6000 system using 101 base paired-end chemistry in the centre.

All RNA samples with good quality (n=98) and poor quality (n=28) were included in the analysis. Poor quality was defined as high percentage of unmapped reads because of low RNA concentration, degraded RNA, or DNA contamination.

FASTQ files of the samples sequenced in 2024 were downloaded from Illumina servers using Command Line Interface (Ubuntu, version 20.04.6). In total, 126 raw FASTQ files were used to perform bulk RNA-seq analyses using the Ubuntu (version 20.04.6) and R software (version 4.3.2).

2.4. Comparison Tests and Clinical Endpoints

The Chi-square test (for categorical variables), Fisher exact test (for categorical variables with expected frequency <5), Mann-Whitney U test (for continuous variables with two groups), and Kruskal-Wallis test (for continuous variables with >2 groups) were performed to compare categorical/continuous data between groups of interest.

The Cochran-Armitage test for binary data and Jonckheere-Terpstra test for continuous data were utilised to evaluate trend of the variables across the trials.

Heterogeneity test was performed to check the presence/absence of variations in variables of interest between different trials/groups in survival analysis.

Clinical endpoints were event-free survival (EFS), relapse-free survival (RFS), and overall survival (OS) in time-to-event analysis. EFS was defined as time (from date of diagnosis/treatment start to event date) to event (relapse, second tumour or death), whichever came first, with censoring at date of the last contact. RFS was defined as time (from date of diagnosis/treatment start to relapse date) to relapse for patients having complete remission, with censoring at date of the last contact. OS was defined as time (from date of diagnosis/treatment start to date of death) to death, with censoring at date of the last contact. Cox proportional hazards regression (Cox regression) method, Kaplan-Meier method, and log-rank test were performed to check statistically significant difference between the defined risk groups in terms of clinical outcomes.

The STATA (version 16) and R (version 4.3.2) were used to perform all statistical analyses with the significance level of 0.05, and *p*-values were two-sided values in all analyses.

Chapter 3. Evaluation of Clinical Risk and Prognostic Factors in Paediatric/Young Adult UKALL Trials

3.1. Introduction

Although B-cell acute lymphoblastic leukaemia (B-cell ALL) and T-cell acute lymphoblastic leukaemia (T-cell ALL) are the subtypes of acute lymphoblastic leukaemia (ALL), certain clinical variables in T-cell ALL are not as prognostically powerful as in B-cell ALL. For example, age and white blood cell count (WBC) at diagnosis are considered as independent prognostic factors in B-cell ALL, and they are used to stratify patients according to National Cancer Institute-Rome (NCI-Rome) classification (Smith *et al.*, 1996; Tasian, Loh and Hunger, 2015). However, the prognostic effect of the same NCI-Rome classification (indirectly age & WBC) is debatable in T-cell ALL (Melchior *et al.*, 2012; Tasian, Loh and Hunger, 2015). Specifically, the optimal threshold of WBC might be higher than $50x10^9$/L, such as $200x10^9$/L, for T-cell ALL risk stratification (Petit *et al.*, 2018; Vaitkevičienė *et al.*, 2011). Minimal/measurable residual disease (MRD) is the only strong factor used for risk stratification in T-cell ALL, but absolute optimal threshold (0.01% or 0.1%) and timepoint (end of induction or end of consolidation) are still debatable although MRD at the end of induction with the threshold of 0.01% is traditionally and historically accepted in studies (Cazzaniga *et al.*, 2003; Burns *et al.*, 2021). In addition, early treatment responses (marrow response or blood response) to chemotherapy are indicators showing the prognosis of patients with T-cell ALL (Gaynon *et al.*, 1997; Griffin *et al.*, 2000; Möricke *et al.*, 2013). Early T-cell precursor (ETP)-ALL, a subgroup of T-cell ALL, was initially considered as a subset of patients with high-risk features (Coustan-Smith *et al.*, 2009). However, Patrick and coworkers reported that this subgroup did not experience worse clinical events for experimental therapy or marrow transplant, compared to typical T-cell ALL in UKALL 2003 (Patrick *et al.*, 2014a). Genetic abnormalities are still controversial in relation to their clinical impacts. For instance, mutational status of *NOTCH1* can have superior or neutral prognostic impacts depending on the studies (Yuan *et al.*, 2022; Clappier *et al.*, 2010). Another example is the French oncogenetic classifier (*NOTCH1*/*FBXW7*/*RAS*/*PTEN* classifier), which has externally validated in FRALLE2000T trial, but not in UKALL 2003 trial (Petit *et al.*, 2018; Taj *et al.*, 2022). In summary, the strongest and commonly used prognostic factor is MRD for risk stratification in T-cell ALL.

Chapter 3 aims to investigate traditional risk and prognostic factors (sex, age, WBC, organomegalies, NCI-Rome classification, treatment response variables (marrow and MRD responses)) and genetic abnormalities in paediatric/young adult T-cell ALL.

The objectives of this chapter are to:

1) Compare clinical characteristics and outcomes of T-cell ALL patients in paediatric/young adult UKALL trials, using data from UKALL VIII to UKALL 2011.

2) Examine clinical impacts of potential risk and prognostic factors in the trials.

3.2. Data and Methods

3.2.1. Definition of Subgroups

Available variables were obtained from the clinical trials of UKALL VIII (n=58), UKALL X (n=139), UKALL XI (n=196), UKALL 97/99 (n=209), UKALL 2003 (n=386), and UKALL 2011 (n=321). The NCI-Rome classification is defined according to age with threshold of 10 years and/or WBC with threshold of $50x10^6$/L. Positive central nervous system (CNS) involvement refers the level of blast cells $\geq 5/mm^3$ in cerebrospinal fluid, indicating CNS3 status.

Patients with early marrow responses were grouped to slow early response (SER) and rapid early response (RER), depending on regimens and bone marrow (BM) blast percentage measured at day 8/15. Those treated with regimen A and had BM blast percentage at day 15 <25% or received regimen B and had BM blast percentage measured at day 8 <25% were assigned to RER group, whereas patients had regimen A and BM blast percentage at day 15 ≥25% or regimen B and BM blast percentage at day 8 ≥25% were considered as SER group (Table 3.1.).

		BM response		BM response
Regimen A	Yes	RER	Yes	SER
BM blast % at d15	<25		≥25	
Regimen B	Yes	RER	Yes	SER
BM blast % at d8	<25		≥25	

Table 3.1. Definition of early marrow response groups based on regimens and marrow measurement days (BM, bone marrow; d8, day 8; d15, day 15; RER, rapid early response; SER, slow early response).

The frequency and prognostic importance of genetic abnormalities were examined in UKALL 97/99 and UKALL 2003. There was no available genetic data in UKALL 2011 so the known abnormalities were not able to be screened in this trial. To increase sample size of each genetic subgroup/abnormality, UKALL 97/99 and UKALL 2003 were combined, and patients were grouped into four main subtypes (overall, there was no statistically significant variation in the comparison of main subtypes between these two trials in terms of event-free survival (I^2=2.3%, *p*-value=0.38), relapse-free survival (I^2=0%, *p*-value=0.64), overall survival (I^2=0%, *p*-value=0.43). The following genetic subtypes were based on ETP status, genetic abnormalities and CD markers: 1) immature subtype containing ETP status and most common type-B abnormalities; 2) TLX3 subtype containing *TLX3* rearrangements; 3) proliferative subtype containing *TLX1* or *NKX2.1* rearrangements, certain CD markers ($CD1a^+$,$CD4^+$,$CD8^+$) and most frequent type-B abnormalities; 4) TAL/LMO subtype containing *TAL1/2*,

LMO1/2 or *LYL1* rearrangements, specific CD status ($CD1a^-$, single positive cells ($CD4^-$ & $CD8^+$ or $CD4^+$ & $CD8^-$)) and most frequent type-B abnormalities (Figure 1.5.; Table 3.4.). First, patients with ETP-ALL were directly stratified into immature subtype. Next, those with certain type-A abnormalities (*TLX1* rearrangements, *TLX3* rearrangements, *NKX2.1* rearrangements, *TAL1/2* rearrangements, *LMO1/2* rearrangements, *LYL1* rearrangements) were categorised to other three subtypes. After that, CD markers and type-B abnormalities were used to include patients in any subtypes. This genetic classification of T-cell ALL patients categorised into four main subtypes was retrospectively done, which means that the classification was not available to the treating clinicians and had no impact on treatment decisions. A few patients were excluded from this genetic classification because they had more than one type-A abnormalities.

3.2.2. Statistical Analysis

The comparison tests and survival analysis, which were previously explained in Chapter 2, were performed in this chapter.

3.3. Results

3.3.1. Overview of Variables in Paediatric and Young Adult UKALL Trials

Although UKALL VIII had the smallest cohort of T-cell ALL (n=58), the number of patients recruited in all trials has dramatically increased over time because of better diagnostics (n=139 in UKALL X; n=196 in UKALL XI; n=209 in UKALL 97/99; n=386 in UKALL 2003; n=321 in UKALL 2011). A total of 1309 patients with T-cell ALL from six consecutive paediatric and young adult UKALL trials were analysed in this chapter. Recent trials (UKALL 2003 and UKALL 2011) started recruiting young adult patients as well as paediatric patients. The proportion of male patients had an increased trend over time (*p*-value=0.03). Although the median level of WBC at diagnosis seemed to slightly decrease over the trials, it was not a significant trend (*p*-value=0.11). On the other hand, the median of BM blast percentage at diagnosis were around 90% in all trials (Table 3.2.).

	UKALL VIII	UKALL X	UKALL XI	UKALL 97/99	UKALL 2003	UKALL 2011
No. patients	58	139	196	209	386	321
Sex ratio (M/F)	1.9	3.5	1.9	2.1	2.7	3.4
Age (years, median-range)	7 (1-14)	8 (0-15)	7 (1-15)	9 (1-16)	10 (1-24)	10 (1-24)
WBC (10^9/L, median-range)	104 (2-940)	114 (1-1375)	106 (1-800)	101 (0-1028)	94 (0-881)	74 (0-1800)
CNS involvement (Y, %)	0	3	5	7	4	7
Mediastinal mass (Y, %)	42	47	53	52	N/A	N/A
Splenomegaly (Y, %)	76	87	74	47	N/A	N/A
Hepatomegaly (Y, %)	71	77	77	40	N/A	N/A
BM blast % at dx (median-range)	90 (25-100)	94 (3-100)	93 (2-100)	92 (1-100)	90 (4-100)	N/A
BM blast % at d8 (median-range)	N/A	N/A	N/A	14 (0-93)	10 (0-98)	11 (0-99)
BM blast % at d15 (median-range)	N/A	N/A	N/A	2.5 (0-82)	4 (0-92)	N/A
BM blast % at d28 (median-range)	N/A	N/A	N/A	2 (0-93)	1 (0-95)	1 (0-92)
MRD % at d28/29 (tested, %)	N/A	N/A	N/A	N/A	73	79

Table 3.2. Trends of clinical characteristics of T-cell ALL patients over time (M/F, the number of males divided by the number of females; WBC, white blood cell count; CNS, central nervous system; BM, bone marrow; dx, diagnosis; d8, day 8; d15; day 15; d28, day 28; d28/d29, day 28 or day 29; MRD, minimal/measurable residual disease; Y, yes; N/A, not available).

3.3.2. Clinical Outcomes of Patients in UKALL Trials

Overall, median follow-up time for survivors was 10 years from UKALL VIII with 17 years to UKALL 2011 with five years (Table 2.1.). UKALL 2003 had the best event-free survival (EFS) rate (80%, 95%CI=77-84), relapse rate (RR) (13%, 95%CI=10-17) and overall survival (OS) rate (86%, 95%CI=82-89) at five years, while patients in UKALL VIII experienced the worst clinical outcomes with EFS rate of 40% (95%CI=27-52), relapse rate of 49% (95%CI=36-64) and OS rate of 48% (95%CI=35-60). There was no statistically significant difference in clinical outcomes between UKALL X and UKALL XI (5-year EFS, 47% vs 50%, *p*-value=0.53; 5-year RR, 46% vs 46%, *p*-value=0.87; 5-year OS, 55% vs 64%, *p*-value=0.11). In contrast, patients treated with UKALL 97/99 protocol had inferior rates compared to those who received UKALL

2003 protocol in 5-year EFS (68% vs 80%, *p*-value=0.0003), 5-year RR (24% vs 13%, *p*-value=0.001) and 5-year OS (73% vs 86%, *p*-value=0.0001) (Figure 3.1.). There was a dramatic decrease in relapse rates between UKALL XI and UKALL 97/99 (46% vs 24%, *p*-value=0.0002). However, OS rates gradually changed between these two trials (64% vs 73%, *p*-value=0.05) (Figure 3.1.).

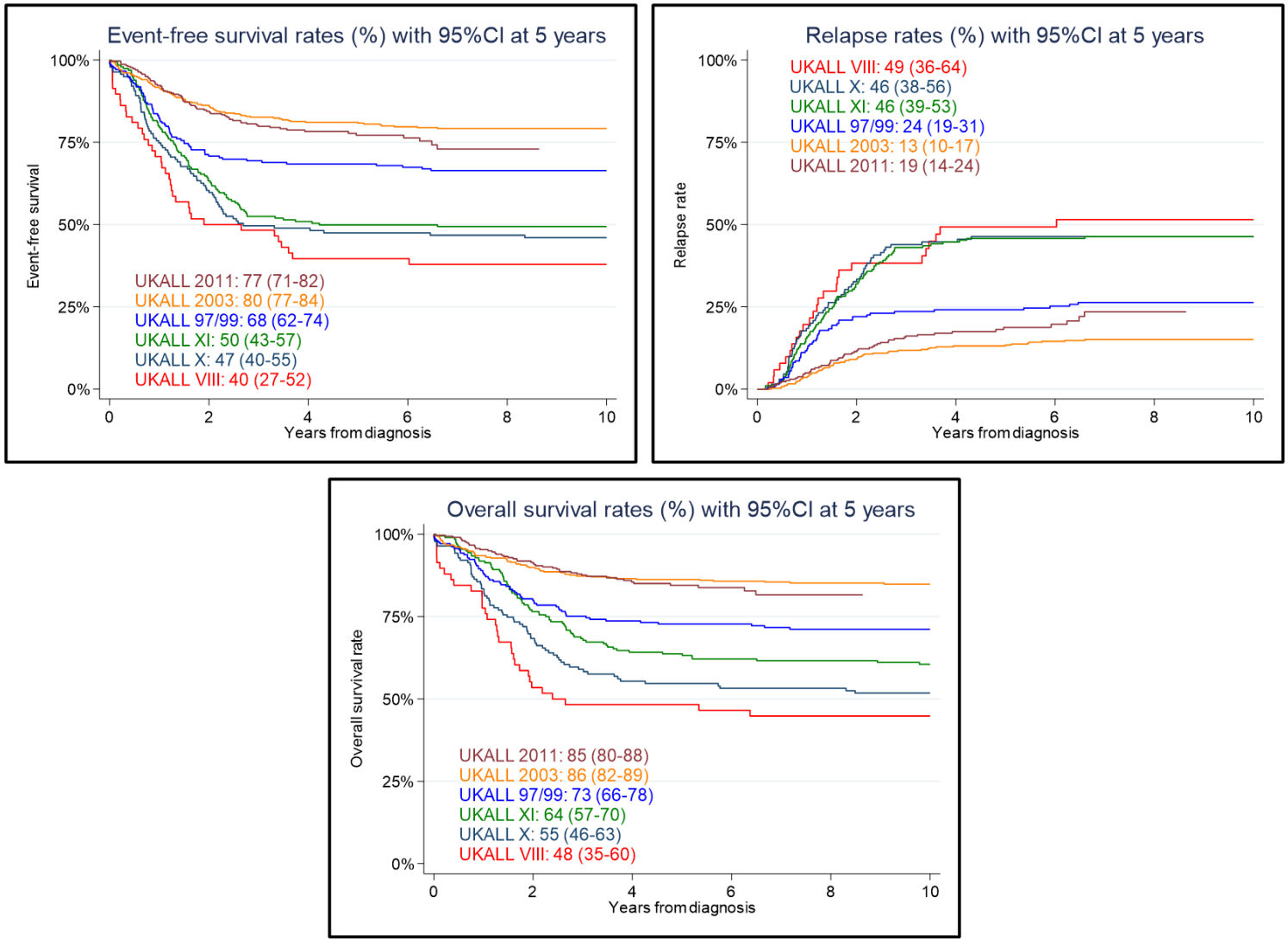

Figure 3.1. Kaplan-Meier graphs for event-free survival, relapse, and overall survival rates at five years in paediatric and young adult T-cell ALL trials. The 95% CIs were included next to the rates in the graphs (CI, confidence interval).

3.3.3. Prognostic Values of Traditional Variables in T-cell ALL

The prognostic effect of traditional variables (sex, age, white blood cell count (WBC), central nervous system (CNS) involvement, organomegalies (anterior mass, splenomegaly, hepatomegaly), marrow and MRD responses) were investigated in all trials. Relapse-free survival (RFS) and overall survival (OS) results from univariable Cox regression analysis indicated that sex, age (<>10 years), WBC (<>$50x10^6$/L or <>$200x10^6$/L), NCI-Rome classification, CNS involvement, and organomegalies did not have consistent prognostic values in the T-cell ALL cohorts (male patients had worse clinical outcomes with significant *p*-value only in UKALL VIII, and high WBC level was associted with poor prognosis only in UKALL X and UKALL XI) (Figure 3.2.). Additional analysis did not reveal statistically significant difference in WBC levels between cases who died with relapse (n=36) and without relapse (n=20) (median, 105.9 vs 169.5, *p*-value=0.21) in UKALL 2003. Also, proportion of WBC levels in different thresholds (50, 100, or 200) did not statistically differ between these two groups. On the other hand, treatment response variables (early marrow response and MRD response with thresholds of 0.01% and 0.1%) were significant in UKALL 97/99, UKALL 2003, and UKALL 2011 (Figure 3.3.).

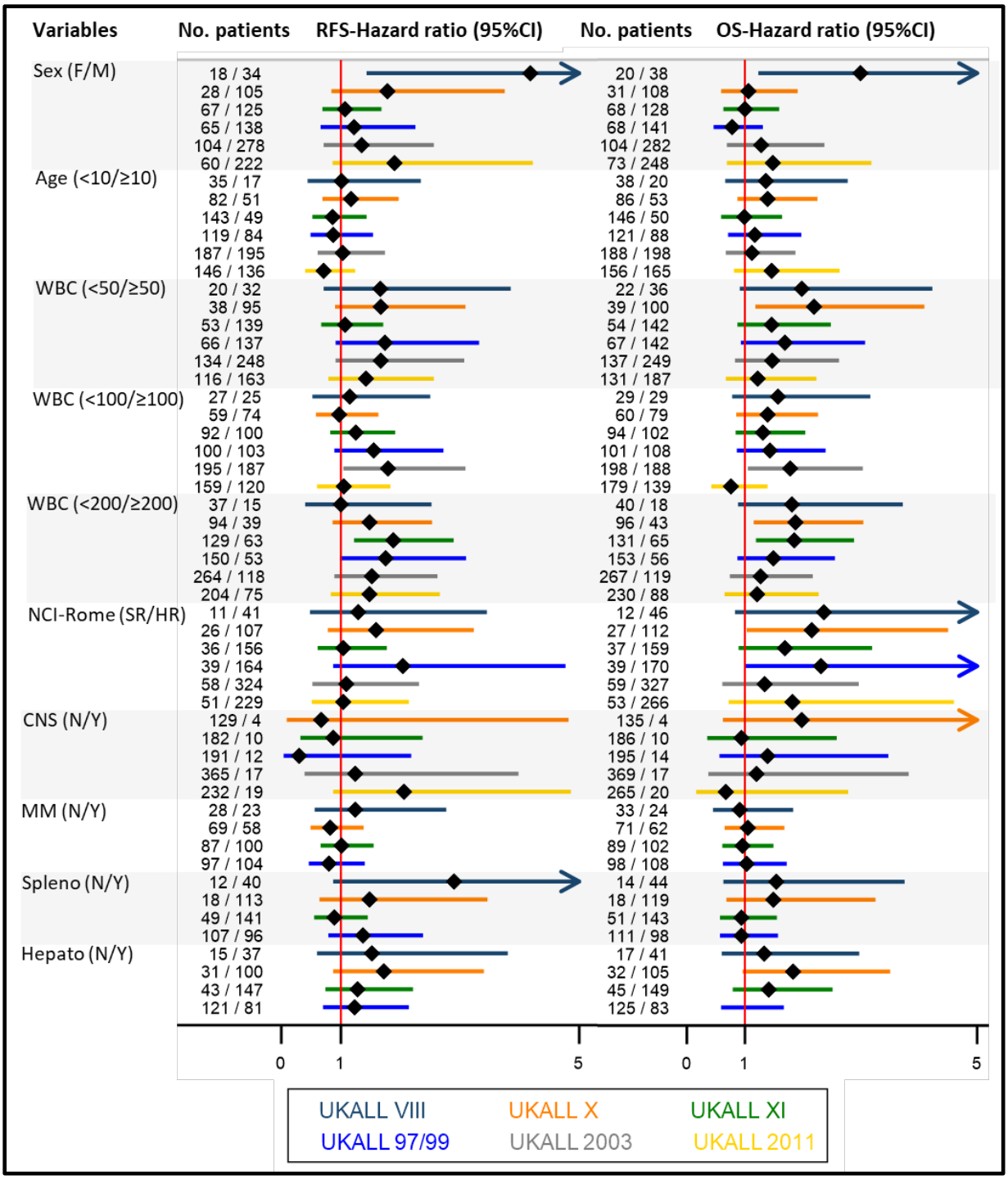

Figure 3.2. Univariable relapse-free survival (RFS) and overall survival (OS) for traditional risk/prognostic variables over the six trials. Different colours represent the trials, and diamond demonstrates hazard ratios (M, male; F, female; WBC, white blood cell count (in unit of 10^6/L at diagnosis); NCI-Rome, National Cancer Institute-Rome; HR, high-risk; SR, standard-risk; CNS, central nervous system; MM, mediastinal mass; Spleno, splenomegaly; Hepato, hepatomegaly).

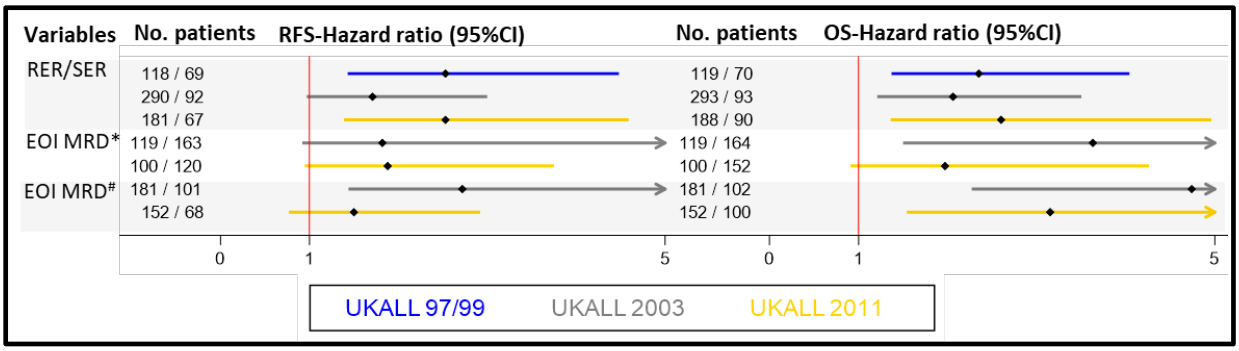

Figure 3.3. Prognostic impact of treatment response variables in univariable relapse-free survival (RFS) and overall survival (OS) (RER, rapid early responder; SER, slow early responder; EOI MRD, minimal/measurable residual disease at the end of induction; *<0.01% vs ≥0.01%; #<0.1% vs ≥0.1%).

3.3.4. Genetic Subtypes and Their Prognostic Impacts on T-cell ALL

T-cell ALL genetic data (type-A and type-B abnormalities) were obtained from UKALL 97/99 and UKALL 2003, but not UKALL 2011 because of unavailable genetic data. The hallmark of type-A abnormalities is overexpression of transcription factors, while type-B abnormalities refer to mutations or deletions in specific genes that have different roles in biological functions (cell cycle and growth regulation, chromatin remodelling, signalling pathways, etc.). Table 3.3. demonstrates the frequency of type-A and type-B abnormalities in UKALL 97/99 and UKALL 2003. Overall, although the proportion of the abnormalities was similar in both cohorts, UKALL 2003 had more positive cases with the abnormalities in comparison to UKALL 97/99. The reasons of this difference could be due to 1) different laboratory techniques, which were used to detect the genetic abnormalities; 2) availability of the samples for retrospective genetic investigation.

	Genes	Abnormalities	Case numbers in UKALL 97/99 (positive/tested (%))	Case numbers in UKALL 2003 (positive/tested (%))
Type-A Abnormalities	*TAL1*	del(1)(p32;p32) [STIL::TAL1] t(1;14)(p32;q11) [TRA/D::TAL1] t(1;7)(p32;q34) [TRB::TAL1]	31/132 (23)	51/281 (18)
	TAL2	t(7;9)(q34;q32) [TRB::TAL2]	3/24 (13)	5/46 (11)
	LMO1	t(11;14)(p15;q11) [LMO1::TRA/D] t(7;11)(q34;p15) [TRB::LMO1]	1/48 (2)	1/88 (1)
	LMO2	t(11;14)(p13;q11) [TRA/D::LMO2] t(7;11)(q34;p13) [TRB::LMO2] del(11)(p13)	12/78 (15)	28/218 (13)
	LYL1	t(7;19)(q34;p13) [TRB::LYL1]	0/54 (0)	1/76 (1)
	TLX1	t(10;14)(q24;q11) [TLX1::TRA/D] t(7;10)(q34;q24) [TLX1::TRB]	1/57 (2)	12/187 (6)
	TLX3	t(5;14)(q35;q11) [TLX3::TRD] t(5;14)(q35;q32) [BCL11B::TLX3]	20/135 (15)	44/263 (17)
	BCL11B	other translocations + deletion + mutation	10/20 (50)	20/48 (42)
	HOXA	inv(7)(p15q34) [TRB::HOXA9/10] t(7;7)(p15;q34) [TRB::HOXA9]	4/26 (15)	8/51 (16)
	KMT2A	t(X;11)(q13;q32) [KMT2A::FOXO4] t(11;19)(q23;p13) [KMT2A::MLLT1] t(6;11)(q27;q23) [KMT2A::MLLT4]	8/184 (4)	10/309 (3)
	AF10 (MLLT10)	t(10;11)(p13;q14) [PICALM::MLLT10]	4/127 (3)	13/131 (10)
	NKX2-1	inv(14)(q11;q13)[NKX2-1::TRA]	1/18 (6)	10/46 (22)
Type-B Abnormalities	*NOTCH1*	mutation	13/16 (81)	138/210 (66)
	FBXW7	mutation	5/18 (28)	37/208 (18)
	PHF6	mutation	4/16 (25)	19/100 (19)
	PTEN	mutation del(10)(q23.31)	5 /19 (26) 5/59 (8)	28/207 (14) 24/252 (10)
	LEF1	mutation del(4)(q25)	2/2 (100) 6/59 (10)	6/6 (100) 12/190 (6)
	CDKN2A/B	del(9)(p21.3)	72/90 (80)	141/200 (71)
	IKZF1	mutation del(7)(p12.2)	1/1 (100) 6/83 (7)	1/1 (100) 3/59 (5)
	WT1	mutation	4/18 (22)	11/99 (11)
	JAK1/3	mutation	3/18 (17)	16/100 (16)
	N/K RAS	mutation	1/1 (100)	16/157 (10)
	IL7R	mutation	4/16 (25)	8/100 (8)
	ETV6	mutation	1/1 (100)	2/2 (100)

Table 3.3. The number of tested type-A and type-B abnormalities in UKALL 97/99 and UKALL 2003. The overexpression of the type-A genes could result from deletions, translocations, inversions, or mutations; therefore, the type-A abnormalities were detected according to these chromosomal/genetic alterations. Different laboratory methods, such as karyotype, FISH, MLPA, PCR, were retrospectively performed to detect the abnormalities, using available samples (pos, positive; neg, negative).

After narrowing abnormalities to four main subtypes, TAL/LMO subtype was the most frequently observed subgroup (n=157) compared to other subtypes (immature subtype (n=53), TLX3 subtype (n=60), and proliferative subtype (n=53)). 272 T-cell ALL patients could not be classified because of lack of information about genetic abnormalities and/or CD markers (Table 3.4.).

Subtypes	Type-A abnormalities	CD markers	Type-B abnormalities	No. patients (UKALL 97/99)	No. patients (UKALL 2003)	Total
Immature	-	ETP-ALL (Patrick et al, 2014a)	*JAK1/3* mut, *N/K RAS* mut, *IL7R* alt, *ETV6* mut, *IKZF1* alt	6	47	53
TLX3	*TLX3* rear	-	-	19	41	60
Proliferative	*TLX1* rear, *NKX2.1* rear	$CD1a^{+}$ & $CD4^{+}$ & $CD8^{+}$	*LEF1* alt, *BCL11B* alt	9	44	53
TAL/LMO	*TAL1/2* rear, *LMO1/2* rear, *LYL1* rear	$CD1a^{-}$ & $CD4^{+/-}$ & $CD8^{-/+}$	*PTEN* alt	49	108	157
Unclassified	-	-	-	126	146	272

Table 3.4. Classification of genetic abnormalities in UKALL 97/99 and UKALL 2003. Four broader subtypes were defined according to ETP status, type-A/B abnormalities, and CD markers (CD, cluster of differentiation; ETP-ALL, early T-cell precursor acute lymphoblastic leukaemia; rear, rearrangement; mut, mutation; alt, alterations (mutation or deletion)).

When comparing clinical characteristics between classified group (immature, TLX3, proliferative, TAL/LMO subtypes) and classified + unclassified group (all cases), there were statistically significant difference in only WBC at diagnosis (median, $122.4x10^9$/L vs $96.6.1x10^9$/L, *p*-value=0.03). Sex, age, diagnostic marrow level, CNS involvement, early marrow responses at day 8 and day 28, MRD response at the end of induction, and rates of event-free survival, relapse and overall survival did not statistically differ between these two groups (Table 3.5.). In general, classified group with genetic information was representative of total T-cell ALL cohort.

Variables	Values	Classified (323, 35%)	Classified + Unclassified (595, 65%)	*p*-value
Sex	male, n (%)	241 (75)	423 (71)	0.26
Age (years)	median (range)	9.4 (1.4-24.0)	9.7 (1.0-24.3)	0.38
WBCx10^9/L	median (range)	122.4 (0.5-1028.0)	96.6 (0.5-1028.0)	**0.03**
BM blast % at diagnosis	median (range)	91 (1-100)	91 (1-100)	0.65
CNS involvement	n (%)	17 (5)	31 (5)	0.97
Regimen C (UKALL99)	n (%)	6 (13)	26 (29)	0.05
Regimen C (UKALL2003)	n (%)	94 (39)	142 (37)	0.55
BM blast % at day 8	median (range)	11 (0-98)	11 (0-98)	0.82
BM blast % at day 28	median (range)	1 (0-50)	1 (0-95)	0.41
EOI MRD (%)*	≥0.01, n (%)	110 (60)	164 (58)	0.64
Event-free survival	% (95%CI)	80 (75-84)	76 (73-80)	0.34
Relapse rate	% (95%CI)	85 (80-89)	83 (79-86)	0.49
Overall survival	% (95%CI)	84 (79-87)	81 (78-84)	0.43

Table 3.5. Clinical characteristics of classified group (immature, TLX3, proliferative, TAL/LMO) and classified + unclassified group in UKALL 97/99 and UKALL 2003
(WBC, white blood cell count; BM, bone marrow; CNS, central nervous system; EOI MRD, minimal/measurable residual disease at the end of induction; *only UKALL 2003).

To compare clinical features among the classified subtypes, WBC at diagnosis, bone marrow (BM) blast percentage at diagnosis and day 8, proportion of cases treated with regimen C, and proportion of positive MRD cases at the end of induction had statistically significant differences (Table 3.6.). However, there was no statistically significant variations in terms of 5-year event-free survival (77% vs 78% vs 81% vs 80%, *p*-value=0.92), 5-year relapse rate (13% vs 19% vs 10% vs 16%, *p*-value=0.47), and 5-year overall survival (79% vs 83% vs 85% vs 85%, *p*-value=0.73) (Figure 3.4.). Notably, although no significant differences observed in the clinical outcomes, patients with proliferative subtype received less intensive treatment (regimen C) (Table 3.6.).

UKALL 97/99 & UKALL 2003	Immature subtype (n=53, 16%)	TLX3 subtype (n=60, 19%)	Proliferative subtype (n=53, 16%)	TAL/LMO subtype (n=157, 49%)	*p*-value
Sex (Male, n (%))	33 (62)	46 (77)	40 (75)	122 (78)	0.16
Age (years, median (range))	10.1 (2.4-24.0)	9.4 (3.2-19.0)	7.5 (1.4-19.2)	9.6 (1.9-20.6)	0.05
WBC (10^9/L, median (range))	22.5 (0.5-615.0)	111.3 (1.0-693.3)	88.7 (3.7-497.0)	167.0 (0.7-1028.0)	**<0.001**
CNS involvement (Y, n (%))	2 (4)	1 (2)	1 (2)	13 (8)	0.12
BM blast % at dx (median (range))	90 (1-100)	95 (9-100)	93 (30-99)	90 (4-100)	**0.001**
Regimen C (Y, n (%))	25 (48)	19 (38)	10 (20)	46 (35)	**0.03**
BM blast % at d8 (median (range))	15 (0-98)	20 (1-95)	7 (0-87)	9 (0-96)	**0.005**
BM blast % at d28 (median (range))	1 (0-50)	1 (0-14)	1 (0-5)	1 (0-18)	0.31
MRD % at EOI (≥0.01, n (%))*	23 (82)	20 (57)	11 (32)	56 (65)	**0.001**

Table 3.6. Clinical characteristics of patients with the main genetic subtypes (WBC, white blood cell counts; CNS, central nervous system; BM, bone marrow; dx, diagnosis; d8, day 8; d28, day 28; EOI MRD, minimal/measurable residual disease at end of induction; *only UKALL 2003).

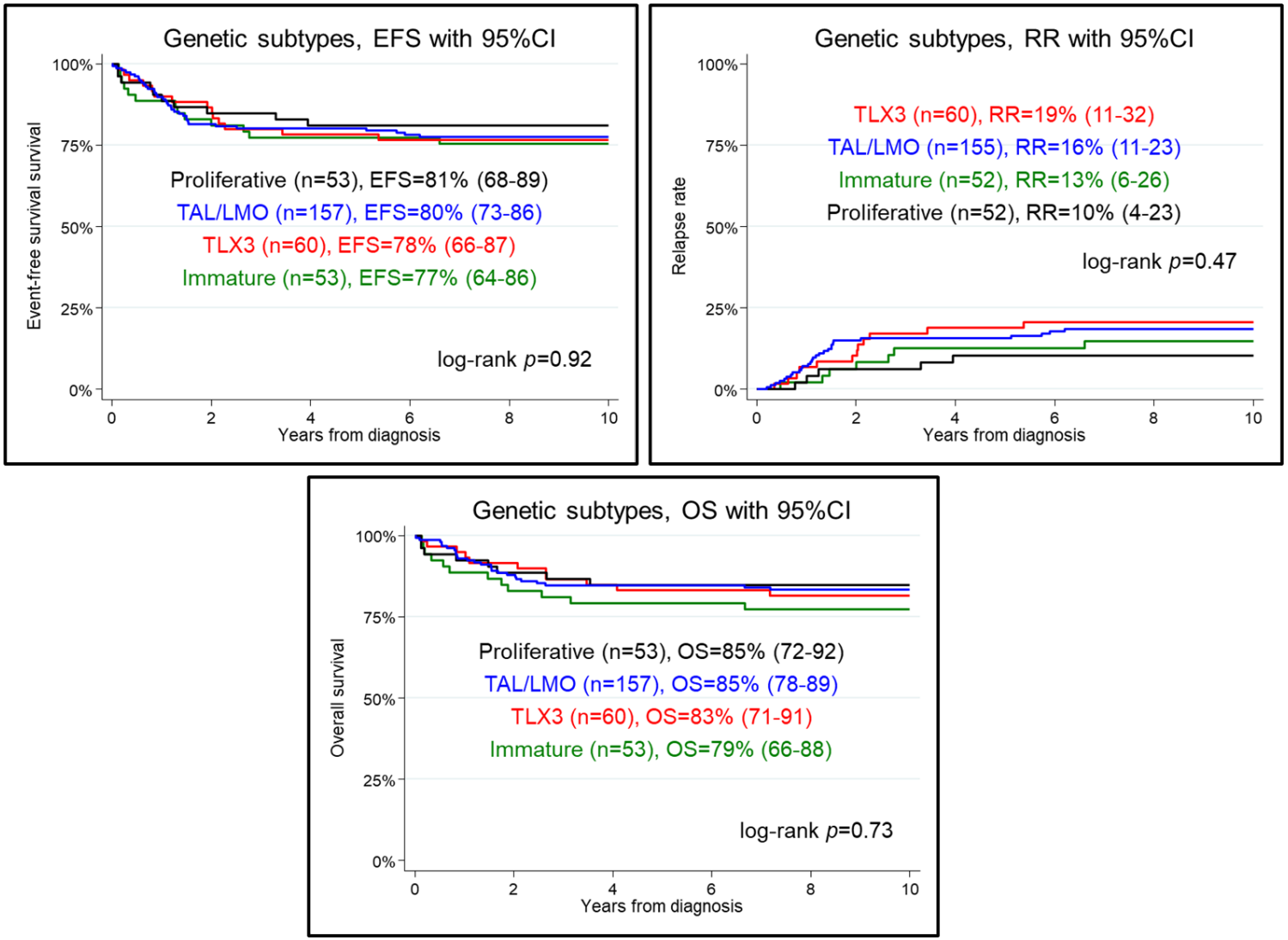

Figure 3.4. Event-free survival (EFS), relapse rate (RR) and overall survival (OS) at five years for the main genetic subtypes. The 95% CIs were included next to the rates in the graphs (CI, confidence interval).

3.4. Discussion

In this chapter, risk and prognostic factors in patients with T-cell ALL (n=1309) were investigated in the following six independent UKALL trials: UKALL VIII (n=58), UKALL X (n=139), UKALL XI (n=196), UKALL 97/99 (n=209), UKALL 2003 (n=386) and UKALL 2011 (n=321). Sex, age, WBC at diagnosis, CNS status, organomegalies (in early trials), and marrow & MRD responses (in late trials) were available in the trials. Young adult patients have been recruited in the recent trials (UKALL 2003 and UKALL 2011), and proportion of male patients has increased over time (Table 3.2.).

T-cell ALL patients treated with UKALL VIII protocol had the worst clinical results in event-free survival (40%), relapse rate (49%) and overall survival (48%) at five years (Figure 3.1.). The bright side of these results was that approximately 50% of these patients cured with less intensive treatment protocol as there was no intensification phase in UKALL VIII (Figure 2.2.). Therefore, it has highlighted that low-risk patients may not need intensification so clinical trials have started focusing on reduction of chemotherapeutic agents to prevent serious side effects of the treatment (Pieters *et al.*, 2016; Vora *et al.*, 2013).

5-year overall survival (OS) rate of T-cell ALL patients have dramatically increased from 48% to 86% over paediatric/young adult UKALL trials (Figure 3.1.). Recent phase-3 clinical trial carried out by Children's Oncology Group (AALL1231) has reported similar OS rate at four years for those with T-cell ALL (n=615, 88.1% ± 1.5%) (Teachey *et al.*, 2022). Another study showed that patients treated with Dana-Farber Cancer Institute (DFCI) ALL Consortium Protocols 05-001 and 11-001 (n=123) had an excellent 5-year OS rate of 90% (95%CI=83-94) (Burns *et al.*, 2021). In addition to this meaningful progression for T-cell ALL patients in this thesis, a big gap between UKALL XI and UKALL 97/99 was observed in relapse rate (RR) graph, but not in OS graph (Figure 3.1.). The main reason of this dramatic change could be due to early marrow classification (slow early response/rapid early response), which was firstly used in UKALL 99 in the UK. The marrow classification allowed to stratify patients with high-risk features for intensive treatment protocol at very early stage in this trial.

UKALL 97 treatment protocol was revised in 1999, and HR1 treatment stratification based on chromosomal abnormalities or Oxford regression score (sex, age, WBC at diagnosis) was abandoned. Importantly, NCI-Rome classification was used to stratify patients into 3 treatment groups (Regimen A, Regimen B and Regimen C) in the revised protocol. Although

therapies in UKALL 99, UKALL 2003 and UKALL 2011 were driven by the NCI-Rome risk classification, but not in UKALL VIII, UKALL X, UKALL XI and UKALL 97. In other words, intensified treatment protocols have been used since UKALL 99 trial. The NCI-Rome classification has had important prognostic value in B-cell ALL, it is not a prognostically strong classification in T-cell ALL. Figure 3.2. demonstrated that standard-risk and high-risk groups defined by NCI-Rome classification did not have significant clinical differences in either early or late trials. Studies have reported that NCI-Rome risk stratification (indirectly age and WBC) is not an effective classification in T-cell ALL (Raetz and Teachey, 2016; Burns *et al.*, 2021; Liu *et al.*, 2020; Melchior *et al.*, 2012). Therefore, it is clearly seen that T-cell ALL does not share similar aspects and is different disease compared to B-cell ALL. Sex, CNS involvement, mediastinal mass, splenomegaly and hepatomegaly did not have obvious clinical impacts on T-cell ALL (Figure 3.2.). In the literature, sex does not have a prognostic value in T-cell ALL patients treated with current therapy (Gupta *et al.*, 2022). Although organomegalies were involved in the calculation of hazard scores in early studies, they lost prognostic significance in contemporary therapy (Vora, 2017). CNS involvement is rare in T-cell ALL, and CNS-targeted therapies have been given to those with CNS disease at diagnosis to eliminate its worse clinical impact (Gossai *et al.*, 2023).

Only early treatment marrow response and MRD response had prognostic importance in T-cell ALL among other variables (Figure 3.2.; Figure 3.3.). Some research groups have reported that T-cell ALL patients with slow marrow response at day 15 (marrow blast ≥25%) have a trend of poor clinical outcomes in event-free survival and overall survival (Wei *et al.*, 2015; Lauten *et al.*, 2012), and many studies have highlighted that T-cell ALL cases with high MRD ($\geq 10^{-4}$ or $\geq 10^{-3}$) experience inferior clinical outcomes compared to those with low MRD ($<10^{-4}$ or $<10^{-3}$) (Basso *et al.*, 2009; Schrappe *et al.*, 2011; Wood *et al.*, 2014; Berry *et al.*, 2017). An interesting observation from my thesis indicated that MRD with cut-off of 0.1% had a superior advantage to MRD with threshold of 0.01% in univariable overall survival analysis, but not univariable relapse-free survival analysis (Figure 3.3.). A report by Burns and colleagues has demonstrated that MRD threshold of 10^{-4} (0.01%) can be better than 10^{-3} (0.1%) for the prognostic ability of MRD in EFS and OS in T-cell ALL (Burns *et al.*, 2021). These inconsistent findings highlight that further investigations are needed to define an absolute/robust MRD threshold for T-cell ALL patients.

In terms of genetic classification, patients were categorised into four subtypes using ETP status, type-A/B abnormalities and CD markers if information was accessible in UKALL

97/99 and UKALL 2003 (Table 3.4.). Classified group (immature subtype, TLX3 subtype, proliferative subtype and TAL/LMO subtype) and classified + unclassified group were compared to each other, and there was no statistically significant difference in clinically relevant variables, such as early marrow response, MRD level at the end of induction, and clinical outcomes (Table 3.5). Therefore, the classified group was representative of T-cell ALL cohort in the combined dataset. Kaplan-Meier graphs clearly showed that these genetic subtypes did not statistically differ from each other in the clinical endpoints (Figure 3.4.). In the literature, there is no consensus on the prognosis of genetic abnormalities in T-cell ALL, which indicates that different treatment protocols may affect clinical outcomes of patients with the known genetic subtypes. For example, positivity of *TLX3* expression in T-cell ALL patients treated on FRALLE-93 was an indicator for poor prognosis (Paola *et al.*, 2008); however, patients with *TLX3*-positive treated with ALL-BFM (Berlin–Frankfurt–Münster) protocol had an excellent clinical outcome (Attarbaschi *et al.*, 2010). Similarly, ETP-ALL (majority of immature subtype) was initially considered a T-cell ALL subgroup with high-risk features (Coustan-Smith *et al.*, 2009; Inukai *et al.*, 2012); however, nonsignificant difference of clinical outcomes between typical T-cell ALL and ETP-ALL was reported by others (Wood *et al.*, 2014; Patrick *et al.*, 2014a). On the other hand, those with proliferative subtype, including *TLX1* and *NKX2* rearrangements, treated with less intensive therapy (Table 3.6.). Some studies reported that *TLX1* positivity indicated good prognosis in paediatric T-cell ALL (Ferrando *et al.*, 2002; Kees *et al.*, 2003).

Unavailable marrow response variables in early trials could be a limitation to check their prognostic impacts in different protocols over the trials. Although the first studies investigating MRD detection in ALL were carried out in 1980s, MRD was firstly used for the purpose of randomisation/stratification of patients in UKALL 2003 in the UK. Therefore, another limitation can be availability of MRD response in early trials. Other limitation would be genetic classification (four broader genetic subtypes) to increase sample size of each subtype because genetic abnormalities were not fully tested for all patients in UKALL 97/99 and UKALL 2003.

In summary, the findings suggest that treatment response variables (early marrow response and MRD response) are important prognostic factors in T-cell ALL, but not sex, age, WBC at diagnosis, NCI-Rome classification, CNS involvement, organomegalies, or the known genetic subtypes.

Chapter 4. Integration of Early Treatment Response Variables into Risk Stratification in T-cell Acute Lymphoblastic Leukaemia

4.1. Introduction

The only prognostic factor commonly used for risk stratification is minimal/measurable residual disease (MRD) in T-cell ALL (Raetz and Teachey, 2016). Overall survival rate of patients with T-cell ALL has dramatically reached around 80-90%; however, the rate of those with relapse/refractory disease has remained poor, only ~33% at five years (Raetz and Teachey, 2016; Hunger *et al.*, 2012; Rheingold *et al.*, 2019). Although MRD-based stratification with intensive treatment protocols has mainly made a dramatic improvement in the first line-treatment in T-cell ALL, the following issues still need to be optimised, or even solved: 1) there is no agreement about absolute optimal MRD threshold (0.01% or 0.1%) and timepoint (end of induction or end of consolidation) (Willemse *et al.*, 2002; Burns *et al.*, 2021; Szczepański *et al.*, 2002; Schrappe *et al.*, 2011); 2) there is no robust and validated genetic predictor or risk stratification model improving MRD prediction specifically for T-cell ALL so far (e.g. the prognostic value of oncogenetic classifier (*NOTCH1/FBXW7/RAS/PTEN*) was checked in different T-cell ALL cohorts, and it could not be validated in UKALL 2003 (Trinquand *et al.*, 2013; Taj *et al.*, 2022)). Hence, a new approach is needed to improve current MRD-based stratification for T-cell ALL patients.

The aims of this chapter are to develop and validate a baseline prognostic model for risk prediction of clinical events, which improves current MRD stratification in T-cell ALL.

The objectives of the fourth chapter are to:

1) Develop a prognostic model for the prediction of clinical outcomes using selection methods with candidate prognostic variables.

2) Validate the prognostic model in the next UKALL trials.

3) Check distribution and clinical impact of the risk groups defined by the prognostic model within different subsets of patients (sex, age, white blood cell count, MRD, and genetics).

4) Investigate interactions between the prognostic model and MRD at the end of induction.

4.2. Data and Methods

Data from UKALL X (n=139), UKALL XI (n=196), UKALL 97/99 (n=209), UKALL 2003 (n=386), and UKALL 2011 (n=321) were used to develop and validate baseline prognostic models. UKALL VIII (n=58) was excluded from model development/validation processes because of insufficient sample size and treatment schedule without an intensification phase. In addition to above criteria, infants with T-cell ALL (age <1 year) were not included in the model development and validation steps in all UKALL trials because the number of these patients were not sufficient for the analysis (only 1 infant patient in total). UKALL X and UKALL XI (UKALL X+XI) were merged to increase sample size for model development as clinical outcomes (event-free survival, relapse rate and overall survival) of these two trials did not statistically differ from each other (Figure 3.1.).

The key change in UKALL treatment protocols was made in UKALL 97 trial, which was revised in 1999. In this revised treatment protocol, HR1 stratification based on chromosomal abnormalities or Oxford regression score (sex, age, WBC at diagnosis) was abandoned, and NCI-Rome classification was firstly used to stratify patients into 3 treatment groups (Regimen A, Regimen B and Regimen C) along with marrow classification. Later, in UKALL 2003, MRD stratification with the threshold of 0.01% was firstly used in the UK as well as the NCI-Rome and marrow classifications. In UKALL 2011, the NCI-Rome classification was abandoned for all T-cell ALL patients who were treated with regimen B, and MRD stratification with the threshold of 0.005% was defined for treatment allocation along with marrow response.

Transformation methods were used for the variables with non-normal distribution. White blood cell count (WBC) at diagnosis and marrow responses at day 8/28 were transformed to logarithmic values (log(WBC+1) and square root values, respectively. Similarly, bone marrow (BM) percentage at diagnosis was transformed using reflection values and square root transformation. Univariable Cox regression method was performed for the identification of potential prognostic variables, and historically relevant factors such as sex, age, and WBC at diagnosis were always added to selection process regardless of statistical significance. Backward and forward selection criteria based on univariable and multivariable Cox regression with likelihood-ratio test were the selection methods to build a baseline prognostic model. Interaction terms in the model development step were checked on whether or not they improved the final model.

Prognostic indexes were created with coefficients of the variables from Cox regression analysis after selection methods. Optimal threshold was defined to categorise patients into standard-risk group and high-risk group, depending on sample size (standard-risk group: ~80% versus high-risk group: ~20%) and relapse rates (at least doubled) of the risk groups. The models were validated in the trials using univariable Cox regression and Kaplan Meier methods (Figure 4.1.). Additionally, publicly available TARGET data was used to check validation results. Calibration curves and statistics were examined for accuracy of risk estimates of the models.

Comparison tests and clinical endpoints were mentioned in the Chapter 2. K^{th}-Nearest Neighbour (KNN) imputation was performed to impute missing values in BM blast percentages at diagnosis, day 8 and day 28 before model development. BM blast percentage at day 15 was not imputed because of very high missingness level.

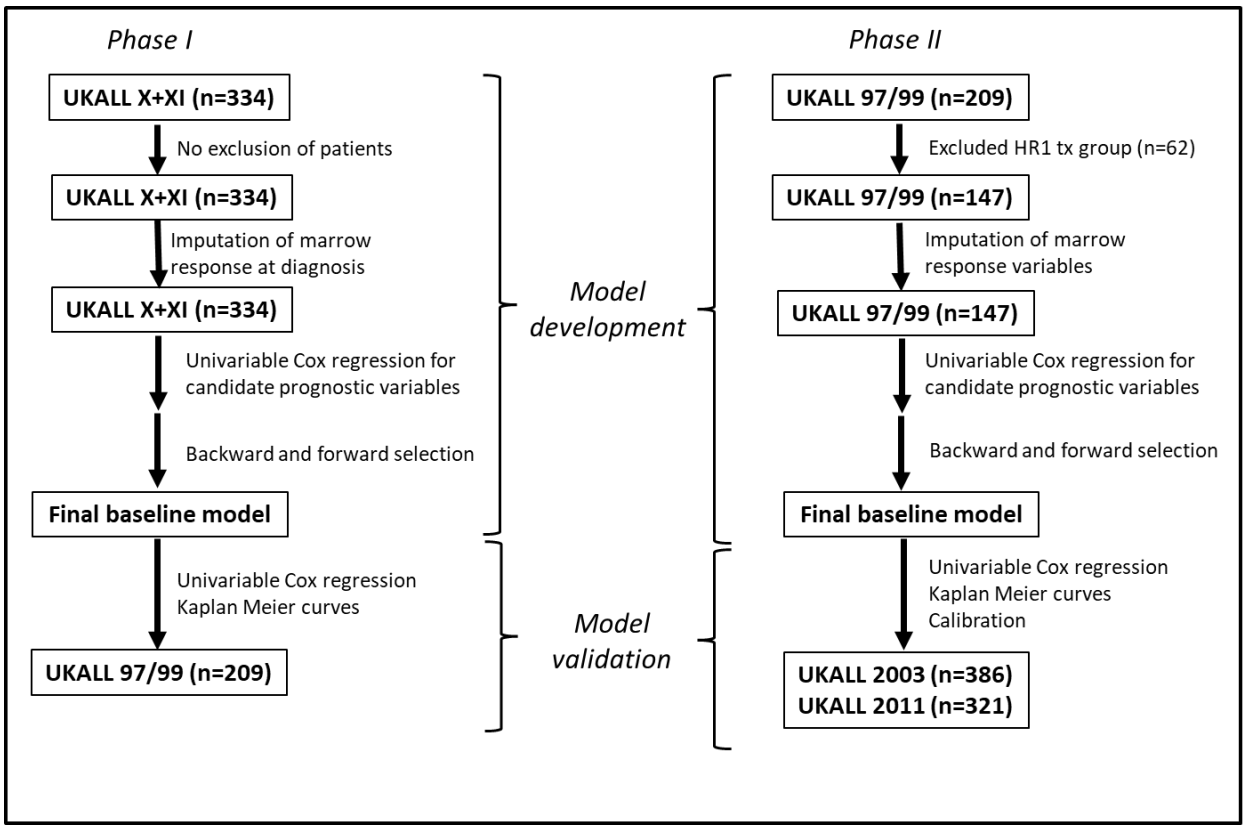

Figure 4.1. Main steps of the model development and validation in phase I (UKALL X+XI and UKALL 97/99) and phase II (UKALL 97/99, UKALL 2003 and UKALL 2011) (tx, treatment).

4.3. Results

4.3.1. Distribution and Handling of Missing Values in UKALL X+XI and UKALL 97/99

Missing data represents incomplete information, which can cause bias and lead to incorrect results in the analysis. In clinical trials, missing data can be observed when participants do not attend scheduled visits, or with incomplete clinical records and measurements, as well as loss of follow-up during/after treatment. There are three main imputation techniques to replace missing values by imputed values: 1) univariate single imputation methods (e.g. mean imputation or median imputation) are used to replace the missing values by single values (mean or median values); 2) multivariate single imputation strategies, for example K^{th}-nearest neighbours (KNN) or random forest, are used to estimate missing data based on other available data; 3) multiple imputation techniques generate and replace missing values many times (Awan *et al.*, 2022). In this thesis, KNN imputation was performed for the missing values because this method preserves original distribution of data and does not need a prediction model (Murti *et al.*, 2019).

Missing values were checked and imputed in combined UKALL X+XI and UKALL 97/99 trials before the steps of prognostic model development. Missing marrow values in HR1 treatment group were not imputed because of very high missingness levels, so patients treated with HR1 protocol were excluded from all analyses in UKALL 97/99 (n=62) (Table 4.1).

Treatment groups	Total no (n=209)	Missing values		
		BM blast % at dx (%)	BM blast % at d8 (%)	BM blast % at d28 (%)
Main (UKALL97)	56	12 (21)	2 (4)	11 (20)
HR1 (UKALL97)	62	7 (11)	13 (21)	59 (95)
Regimen A (UKALL99)	13	2 (15)	1 (8)	0 (0)
Regimen B (UKALL99)	52	10 (19)	0 (0)	4 (8)
Regimen C (UKALL99)	26	4 (15)	1 (4)	3 (12)

Table 4.1. Distribution of missing values in marrow variables by treatment (BM, bone marrow; dx, diagnosis; d8, day 8; d28, day 28).

Missingness pattern of the marrow variables was examined in UKALL X+XI and UKALL 97/99. Complete data (patients without missing values) and missing data (patients with missing values) were compared to each other. There was no statistically significant difference between complete data and missing data in both trials; therefore, incomplete values were randomly distributed in UKALL X+XI and UKALL 97/99 (Table 4.2.). In addition, specific pattern of missingness was not observed in data with information of patients recruited by different hospitals and treatment time in UKALL 97/99 (Table 4.3.).

UKALL X+XI	Complete data (n=296, 89%)	Missing data (n=38, 11%)	*p*-value
Sex (male, n (%))	211 (71)	24 (63)	0.30
Age (years, median (range))	7.3 (1-15)	7.5 (1-15)	0.64
WBC (10^9/L, median (range))	108.5 (1-1375)	131 (5-846)	0.38
CNS involvement (Y, n (%))	11 (4)	3 (8)	0.21
Mediastinal mass (Y, n (%))	149 (50)	15 (56)	0.60
Splenomegaly (Y, n (%))	234 (79)	27 (79)	0.96
Hepatomegaly (Y, n (%))	227 (77)	26 (76)	0.98
BM blast % at dx (median (range))	93 (2-100)	95 (75-100)	0.97
Treatment-DIx2 (Y, n (%))	246 (83)	34 (89)	0.32
Event (Y, n (%))	158 (53)	19 (50)	0.69
Relapse (Y, n (%))	129 (45)	15 (42)	0.72
Death (Y, n (%))	133 (45)	15 (39)	0.52

UKALL 97/99 (no HR1 tx group)	Complete data (n=103, 70%)	Missing data (n=44, 30%)	*p*-value
Sex (male, n (%))	62 (60)	26 (59)	0.90
Age (years, median (range))	9.1 (1-16)	7.3 (1-16)	0.13
WBC (10^9/L, median (range))	60.4 (1-869)	76.5 (2-711)	0.50
CNS involvement (Y, n (%))	9 (9)	2 (5)	0.51
Mediastinal mass (Y, n (%))	50 (49)	28 (64)	0.09
Splenomegaly (Y, n (%))	41 (40)	24 (55)	0.10
Hepatomegaly (Y, n (%))	38 (37)	18 (41)	0.65
BM blast % at dx (median (range))	90 (1-100)	90 (1-100)	0.88
UKALL99-regimen C (Y, n (%))	18 (26)	8 (35)	0.45
BM blast % at d8 (median (range))	13.5 (0-88)	9.5 (2-93)	0.49
BM blast % at d28 (median (range))	2 (0-93)	2 (0-42)	0.69
Event (Y, n (%))	32 (31)	14 (32)	0.93
Relapse (Y, n (%))	25 (24)	10 (23)	0.84
Death (Y, n (%))	29 (28)	12 (27)	0.91

Table 4.2. Distribution of clinical characteristics between complete data (patient with complete values) and missing data (patients with missing values) in UKALL X+XI and UKALL 97/99. HR1 treatment group was excluded because of high level missing pattern in UKALL 97/99 (tx, treatment; WBC, white blood cell count; CNS, central nervous system; BM, bone marrow; dx, diagnosis; d8, day 8; d28, day 28; DIx2, early and late delayed intensification (2 times); Y, yes).

Hospitals in 1997	Total (n=16)	Complete (n=10)	Incomplete (n=6)
AH, Cambridge	1	1	0
HH, Birmingham	2	2	0
JRH, Oxford	3	0	3
RHSC, Bristol	1	1	0
RHSC, Glasgow	1	1	0
RMCH, Manchester	3	2	1
RMH, Sutton	2	0	2
SBH, London	2	2	0
SJH, Leeds	1	1	0

Hospitals in 1998	Total (n=22)	Complete (n=16)	Incomplete (n=6)
ARIH, Aberdeen	1	1	0
AHCH, Liverpool	1	1	0
HH, Birmingham	2	1	1
GOH, London	4	2	2
JRH, Oxford	1	0	1
LH, Penarth	1	1	0
UH, Nottingham	2	2	0
RHSC, Bristol	1	1	0
RHSC, Edinburgh	1	1	0
RHSC, Glasgow	2	2	0
RVI, Newcastle	1	0	1
CH, Sheffield	1	1	0
GH, Southampton	1	1	0
SBH, London	1	0	1
SJH, Leeds	2	2	0

Hospitals in 1999	Total (n=26)	Complete (n=12)	Incomplete (n=14)
AH, Cambridge	1	1	0
AHCH, Liverpool	2	0	2
CH, Birmingham	1	0	1
GOH, London	2	0	2
JRH, Oxford	2	2	0
LH, Penarth	1	1	0
UH, Nottingham	3	1	2
RHSC, Belfast	1	0	1
RHSC, Bristol	1	1	0
RHSC, Edinburgh	2	1	1
RMH, Sutton	2	0	2
GH, Southampton	2	1	1
SBH, London	3	3	0
SGH, London	2	0	2
TLH, London	1	1	0

Hospitals in 2000	Total (n=26)	Complete (n=20)	Incomplete (n=6)
AHCH, Liverpool	3	2	1
GOH, London	2	1	1
JRH, Oxford	1	0	1
LYI, Leicester	2	2	0
LH, Penarth	2	2	0
UH, Nottingham	1	1	0
OLH, Dublin	1	1	0
RHSC, Bristol	2	1	1
RHSC, Edinburgh	1	1	0
RHSC, Glasgow	3	3	0
RMCH, Manchester	2	1	1
RMH, Sutton	3	2	1
CH, Sheffield	1	1	0
GH, Southampton	1	1	0
SBH, London	1	1	0

Hospitals in 2001	Total (n=44)	Complete (n=36)	Incomplete (n=8)
AHCH, Liverpool	5	5	0
CH, Birmingham	6	4	2
JRH, Oxford	3	2	1
LYI, Leicester	1	1	0
LH, Penarth	1	1	0
OLH, Dublin	4	3	1
RHSC, Belfast	1	1	0
RHSC, Bristol	3	2	1
RHSC, Edinburgh	1	1	0
RMCH, Manchester	7	7	0
RMH, Sutton	6	5	1
RVI, Newcastle	1	1	0
CH, Sheffield	2	1	1
GH, Southampton	2	1	1
TMH, London	1	1	0

Hospitals in 2002	Total (n=13)	Complete (n=9)	Incomplete (n=4)
CH, Birmingham	2	1	1
JRH, Oxford	1	1	0
LYI, Leicester	2	2	0
LH, Penarth	1	1	0
OLH, Dublin	2	2	0
RHSC, Bristol	2	0	2
RHSC, Glasgow	1	1	0
RMCH, Manchester	1	1	0
RMH, Sutton	1	0	1

Table 4.3. The number of patients recruited to different hospitals for UKALL97/99 trial between 1997 and 2002. Complete means patients with complete data for marrow values at diagnosis, day 8, and day 28, while incomplete indicates patents with missing marrow values at diagnosis, day 8, or day 28 (ARIH, Aberdeen Royal Infirmary Hospital; AH, Addenbrookes Hospital; AHCH, Alder Hey Children's Hospital; CH, Children's Hospital; HH, Heartlands Hospital; GOH, Great Ormond Hospital; JRH, John Radcliffe Hospital; LYI, Leicester Royal Infirmary; LH, Llandough Hospital; UH, University Hospital; OLH, Our Lady's Hospital, RHSC, Royal Hospital for Sick Children; RMCH, Royal Manchester Children's Hospital; RMH, Royal Marsden Hospital; RVI, Royal Victoria Infirmary; GH, General Hospital; SBH, St. Bartholomews Hospital; SGH, St. Georges Hospital; SJH, St. James Hospital; TLH, The London Hospital; TMH, The Middlesex Hospital).

In UKALL X+XI and UKALL 97/99, BM blast percentages measured at diagnosis, day 8 and day 28 with missing values were imputed using the KNN imputation method with reference to sex, age, WBC at diagnosis, central nervous system (CNS) status, and organomegalies. An optimal k threshold was selected as 5 (k=5) because the coefficients, standard errors and C-indexes of complete values were close to the coefficients, standard errors, and C-indexes of combined values (complete + imputed values). Also, this threshold could be used in all variables in both trials, instead of choosing different thresholds for each variable. These evaluation metrics were generated from univariable event-free survival analysis with bootstrap technique (replication=100) (Table 4.4).

UKALL X+XI (BM blast % at dx)	n	coefficient	SE	*p*-value	C-index
Complete values	305	0.011	0.0066	0.086	0.54
Complete+imputed values (k=1)	334	0.009	0.0053	0.094	0.53
Complete+imputed values (k=3)	334	0.009	0.0051	0.082	0.53
Complete+imputed values (k=5)	334	0.010	0.0055	0.069	0.53
Complete+imputed values (k=7)	334	0.010	0.0055	0.072	0.53
Complete+imputed values (k=9)	334	0.010	0.0056	0.067	0.53

UKALL 97/99 (BM blast % at dx)	n	coefficient	SE	*p*-value	C-index
Complete values	119	-0.002	0.007	0.76	0.55
Complete+imputed values (k=1)	147	-0.002	0.007	0.81	0.57
Complete+imputed values (k=3)	147	0.001	0.007	0.9	0.46
Complete+imputed values (k=5)	147	-0.001	0.007	0.86	0.55
Complete+imputed values (k=7)	147	0.0001	0.007	0.99	0.46
Complete+imputed values (k=9)	147	0.0010	0.007	0.86	0.47

UKALL 97/99 (BM blast % at d8)	n	coefficient	SE	*p*-value	C-index
Complete values	143	0.019	0.006	0.001	0.65
Complete+imputed values (k=1)	147	0.018	0.005	0.001	0.63
Complete+imputed values (k=3)	147	0.018	0.005	0.001	0.63
Complete+imputed values (k=5)	147	0.018	0.005	0.001	0.63
Complete+imputed values (k=7)	147	0.018	0.005	<0.001	0.64
Complete+imputed values (k=9)	147	0.018	0.005	<0.001	0.64

UKALL 97/99 (BM blast % at d28)	n	coefficient	SE	p-value	C-index
Complete values	129	0.041	0.008	<0.001	0.58
Complete+imputed values (k=1)	147	0.041	0.009	<0.001	0.59
Complete+imputed values (k=3)	147	0.039	0.008	<0.001	0.59
Complete+imputed values (k=5)	147	0.039	0.008	<0.001	0.59
Complete+imputed values (k=7)	147	0.039	0.008	<0.001	0.58
Complete+imputed values (k=9)	147	0.039	0.008	<0.001	0.58

Table 4.4. Selection of optimal k threshold in the KNN imputation for marrow variables in UKALL X+XI and UKALL 97/99 (BM, bone marrow; dx, diagnosis; d8, day 8; d28, day 28; SE, standard error).

After imputing missing values in the marrow variables, a significant difference between complete data and combined data (complete data+imputed data) was not observed in clinical characteristics of patients in UKALL X+XI and UKALL 97/99 (Table 4.5.). This finding indicated that imputed values, using KNN imputation with k=5, did not affect the distribution of clinical variables in both clinical trials; hence, the cohorts with imputed values were representative of all T-cell ALL patients from UKALL X+XI and UKALL 97/99.

UKALL X+XI	Complete data (n=296)	Complete + imputed data (n=334)	*p*-value
Sex (male, n (%))	211 (71)	235 (70)	0.80
Age (years, median (range))	7.3 (1-15)	7.31 (1-15)	0.91
WBC (10^9/L, median (range))	108.5 (1-1375)	110.0 (1-1375)	0.83
CNS involvement (Y, n (%))	11 (4)	14 (4)	0.76
Mediastinal mass (Y, n (%))	149 (50)	164 (51)	0.91
Splenomegaly (Y, n (%))	234 (79)	261 (79)	0.99
Hepatomegaly (Y, n (%))	227 (77)	253 (77)	1.00
BM blast % at dx (median (range))	93 (2-100)	94 (2-100)	1.00
Treatment-DIx2 (Y, n (%))	246 (83)	280 (84)	0.81
Event (Y, n (%))	158 (53)	177 (53)	0.92
Relapse (Y, n (%))	129 (45)	144 (44)	0.93
Death (Y, n (%))	133 (45)	148 (44)	0.88

UKALL 97/99 (no HR1 tx group)	Complete data (n=103)	Complete + imputed data (n=147)	*p*-value
Sex (male, n (%))	62 (60)	88 (60)	0.96
Age (years, median (range))	9.1 (1-16)	7.9 (1-16)	0.53
WBC (10^9/L, median (range))	60.4 (1-869)	62.0 (1-869)	0.78
CNS involvement (Y, n (%))	9 (9)	11 (7)	0.72
Mediastinal mass (Y, n (%))	50 (49)	78 (53)	0.48
Splenomegaly (Y, n (%))	41 (40)	65 (44)	0.49
Hepatomegaly (Y, n (%))	38 (37)	56 (38)	0.85
BM blast % at dx (median (range))	90 (1-100)	90 (1-100)	0.88
UKALL99-regimen C (Y, n (%))	18 (26)	26 (29)	0.77
BM blast % at d8 (median (range))	13.5 (0-88)	12.0 (0-92.5)	0.77
BM blast % at d28 (median (range))	2 (0-93)	2 (0-93)	0.91
Event (Y, n (%))	32 (31)	46 (31)	0.97
Relapse (Y, n (%))	25 (24)	35 (24)	0.93
Death (Y, n (%))	29 (28)	41 (28)	0.96

Table 4.5. Comparison of clinical characteristics of patients with complete and complete+imputed data in UKALL X+XI and UKALL 97/99 (tx, treatment; WBC, white blood cell count; CNS, central nervous system; BM, bone marrow; dx, diagnosis; d8, day 8; d28, day 28; DIx2, early and late delayed intensification (2 times); Y, yes).

4.3.2. Model Development and Validation Using UKALL X+XI and UKALL 97/99

4.3.2.1. Identification of Candidate Prognostic Variables in UKALL X+XI

Sex, age, WBC at diagnosis, CNS status, organomegalies, and BM blast percentage at diagnosis were available in UKALL X+XI, and they were evaluated in univariable Cox regression to check their possible prognostic impacts. WBC (log(WBC+1)) had significant

prognostic values in event-free survival, relapse-free survival and overall survival, whereas hepatomegaly and BM blast percentage at diagnosis (reflected and then log(BM percentage at diagnosis+1) had trends of poor and good clinical outcomes, respectively (Figure 4.2.).

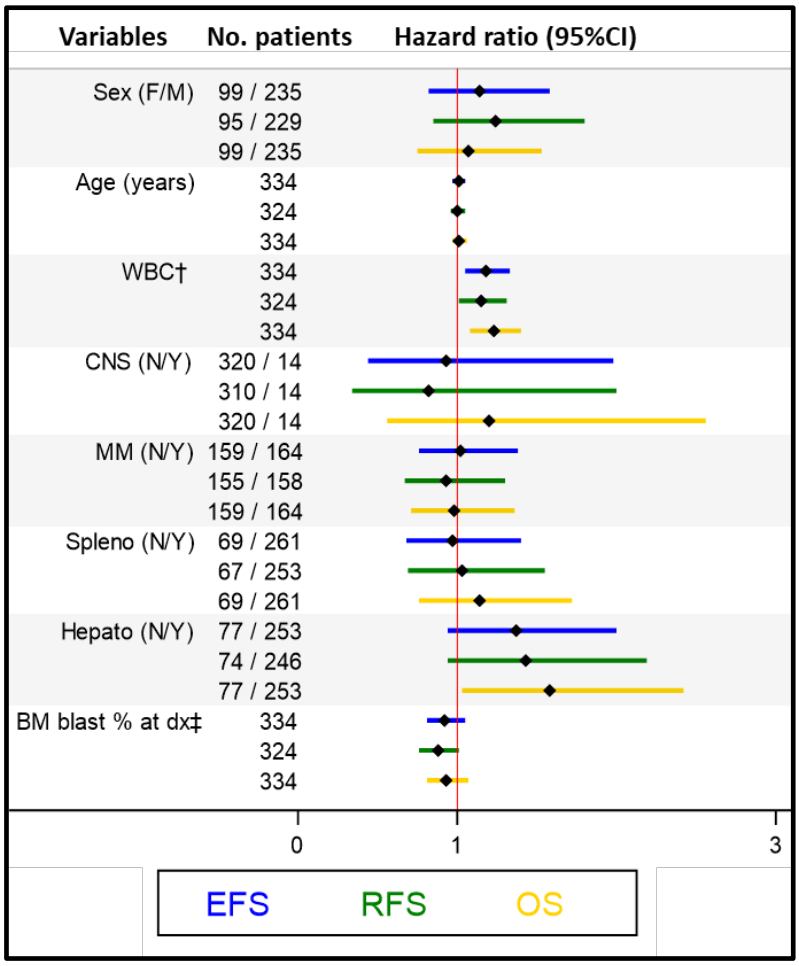

Figure 4.2. Prognostic effect of clinical variables in univariable Cox regression in the combined UKALL X+XI trial. Different colours indicate clinical endpoints, and diamond demonstrates hazard ratios (F/M, female/male; WBC, white blood cell count; CNS, central nervous system, MM, mediastinal mass; BM, bone marrow; dx, diagnosis; N/Y, no/yes; †log transformation; ‡log transformation after reflecting values).

4.3.2.2. Developing and Validating Prognostic Models in UKALL X+XI and UKALL 97/99

Results from univariable Cox regression analysis revealed that WBC at diagnosis, hepatomegaly and BM blast percentage at diagnosis could be included in the model selection processes (backward and forward selections) (Figure 4.2.). Although sex and age did not have prognostic importance in the analysis, these variables were included in the selection processes because of historical relevance and controversy in the literature. Hence, these five clinical variables (sex, age, WBC at diagnosis, hepatomegaly and BM blast percentage at diagnosis) were used to develop prognostic models for the risk estimation of clinical events in UKALL X+XI.

Three models were built using Cox regression coefficients from event-free survival (EFS), relapse-free survival (RFS) and overall survival (OS): EFS-model, RFS-model, and OS-model. The assumption of proportionality for Cox regression was checked by the plot of Schoenfeld residuals and its statistical test. Horizontal line and nonsignificant *p*-value demonstrate that there is no violation of this assumption. In case of violation of this assumption, time varying covariates were examined and used to develop baseline prognostic models (Figure 4.3.). WBC at diagnosis and hepatomegaly remained in all final models after the selection processes, including interaction terms step (Table 4.6.). The models were: 1) EFS-model=log(WBC+1) x 0.155934 + Hepatomegaly x 0.1360978; 2) RFS-model=log(WBC+1) x 0.1051928 + Hepatomegaly x 0.2303952; 3) OS-model=log(WBC+1) x (-0.0005777) + Hepatomegaly x 0.2575416. Normalised scores of the models (0-10) were statistically significant in univariable event-free survival and relapse-free survival analyses, but not in overall survival analysis (Table 4.7.). The validation results of the models in UKALL 97/99 indicated similar hazard ratios that were not statistically significant (Table 4.7.). As the hazard ratios were not high and failed to reach statistical significance in the validation cohort, I did not pursue this model any further investigation.

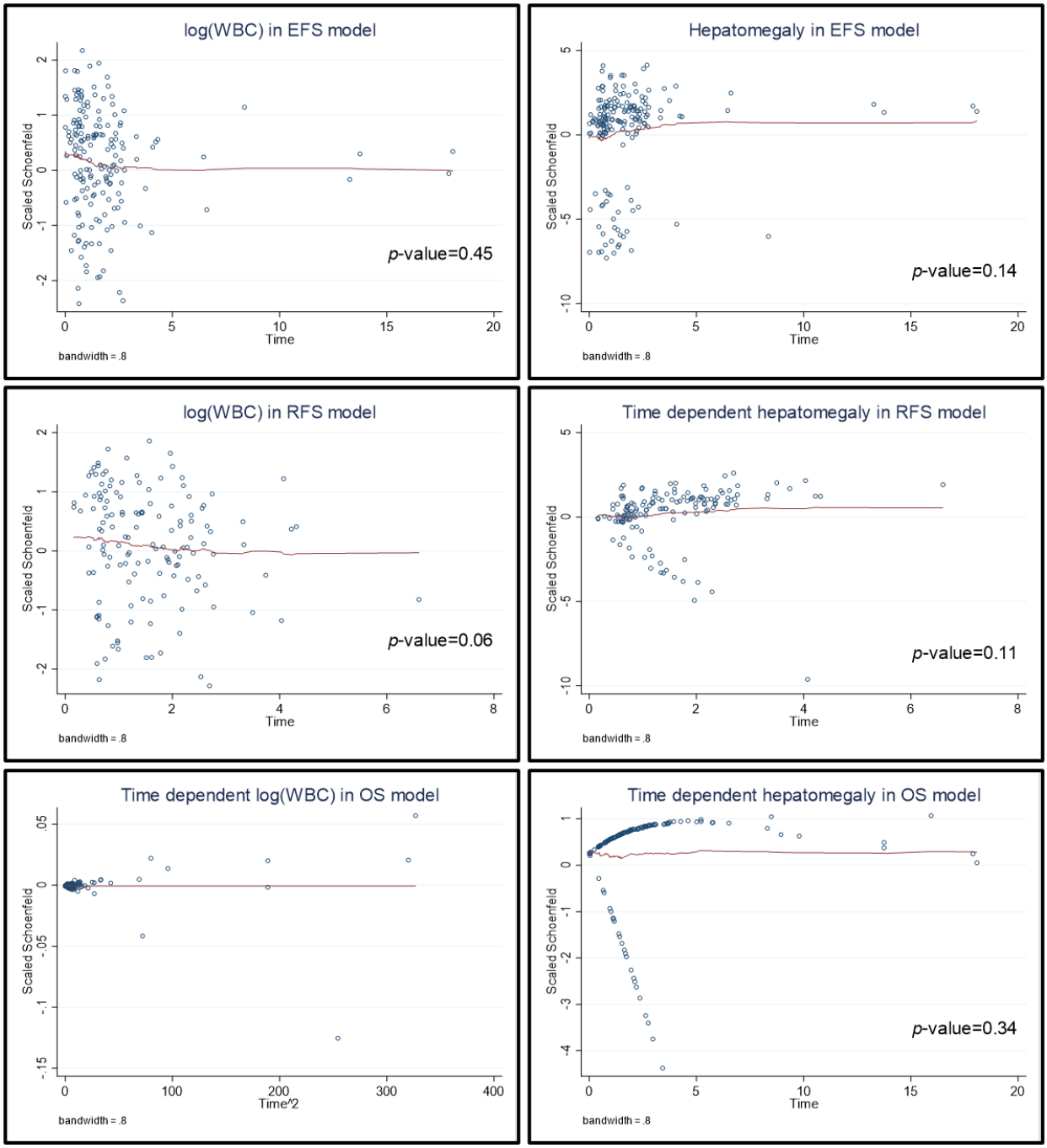

Figure 4.3. Evaluation of Cox regression assumption for the variables in baseline prognostic models developed in UKALL X+XI (WBC, white blood cell count; EFS, event-free survival; RFS, relapse-free survival; OS, overall survival).

EFS (Backward)	Models	Elimination	*p*-value
Step 1	WBC† + Hepatomegaly + Age + BM blast % at dx‡ + Sex	Sex	0.66
Step 2	WBC† + Hepatomegaly + Age + BM blast % at dx‡	BM blast % at dx‡	0.65
Step 3	WBC† + Hepatomegaly + Age	Age	0.61
Step 4	WBC† + Hepatomegaly	Hepatomegaly	**<0.0001**
Step 5	WBC† + Hepatomegaly + (WBC† x Hepatomegaly)	(WBC† x Hepatomegaly)	0.39
Final model	**WBC† + Hepatomegaly**		

EFS (Forward)	Models	Addition	*p*-value
Step 1	WBC†	Hepatomegaly	**<0.0001**
Step 2	WBC† + Hepatomegaly	BM blast % at dx‡	0.63
Step 3	WBC† + Hepatomegaly	Sex	0.66
Step 4	WBC† + Hepatomegaly	Age	0.61
Step 5	WBC† + Hepatomegaly	(WBC† x Hepatomegaly)	0.39
Final model	**WBC† + Hepatomegaly**		

RFS (Backward)	Models	Elimination	*p*-value
Step 1	WBC† + BM blast % at dx‡ + Sex + Hepatomegaly + Age	Age	0.87
Step 2	WBC† + BM blast % at dx‡ + Sex + Hepatomegaly	Hepatomegaly	**<0.0001**
Step 3	WBC† + BM blast % at dx‡ + Sex + Hepatomegaly	Sex	0.43
Step 4	WBC† + BM blast % at dx‡ + Hepatomegaly	BM blast % at dx‡	0.25
Step 5	WBC† + Hepatomegaly + (WBC† x Hepatomegaly)	(WBC† x Hepatomegaly)	0.10
Final model	**WBC† + Hepatomegaly**		

RFS (Forward)	Models	Addition	*p*-value
Step 1	WBC†	BM blast % at dx‡	0.19
Step 2	WBC†	Hepatomegaly	**<0.0001**
Step 3	WBC† + Hepatomegaly	Sex	0.40
Step 4	WBC† + Hepatomegaly	Age	0.92
Step 5	WBC† + Hepatomegaly	(WBC† x Hepatomegaly)	0.10
Final model	**WBC† + Hepatomegaly**		

OS (Backward)	Models	Elimination	*p*-value
Step 1	WBC† + Hepatomegaly + Age + Sex + BM blast % at dx‡	BM blast % at dx‡	0.98
Step 2	WBC† + Hepatomegaly + Age + Sex	Sex	0.96
Step 3	WBC† + Hepatomegaly + Age	Age	0.56
Step 4	WBC† + Hepatomegaly	Hepatomegaly	**0.0001**
Step 5	WBC† + Hepatomegaly + (WBC† x Hepatomegaly)	(WBC† x Hepatomegaly)	0.92
Final model	**WBC† + Hepatomegaly**		

OS (Forward)	Models	Addition	*p*-value
Step 1	WBC†	Hepatomegaly	**0.0001**
Step 2	WBC† + Hepatomegaly	BM blast % at dx‡	0.94
Step 3	WBC† + Hepatomegaly	Age	0.56
Step 4	WBC† + Hepatomegaly	Sex	0.97
Step 5	WBC† + Hepatomegaly	(WBC† x Hepatomegaly)	0.92
Final model	**WBC† + Hepatomegaly**		

Table 4.6. Selection methods for the model development in time-to-event analysis. *p*-values are from likelihood-ratio test comparing two models with/without the variables, which were added or eliminated in the selection process (EFS, event-free survival; RFS, relapse-free survival; OS, overall survival; WBC, white blood cell count; BM, bone marrow; dx, diagnosis; †logarithmic transformation; ‡logarithmic transformation after reflecting values).

UKALL X+XI	No. patients	EFS-model	RFS-model	OS-model
Hazard ratio in EFS (95%CI)	330	1.12 (1.04-1.21)	1.11 (1.03-1.20)	1.03 (0.99-1.07)
Hazard ratio in RFS (95%CI)	320	1.11 (1.02-1.20)	1.10 (1.02-1.20)	1.04 (0.99-1.08)
Hazard ratio in OS (95%CI)	330	1.15 (1.06-1.26)	1.15 (1.05-1.25)	1.05 (1.00-1.09)

UKALL 97/99	No. patients	EFS-model	RFS-model	OS-model
Hazard ratio in EFS (95%CI)	208	1.11 (0.99-1.24)	1.08 (0.98-1.20)	1.01 (0.96-1.06)
Hazard ratio in RFS (95%CI)	202	1.10 (0.97-1.26)	1.09 (0.97-1.22)	1.02 (0.96-1.08)
Hazard ratio in OS (95%CI)	208	1.07 (0.95-1.20)	1.05 (0.94-1.17)	1.00 (0.95-1.05)

Table 4.7. Hazard ratios of normalised scores of the prognostic models developed in UKALL X+XI and validated in UKALL 97/99. The scores from the models were normalised between 0-10. Hazard ratio indicates relative change in risk associated with each one-unit increase in the model scores (EFS, event-free survival; RFS, relapse-free survival; OS, overall survival).

4.3.3. Model Development and Validation in UKALL 97/99, UKALL 2003 and UKALL 2011

Models were developed using data from UKALL 97/99, which additionally contained treatment response variables, such as BM blast percentages at day 8 and day 28. UKALL 2003 and UKALL 2011 were used as validation cohorts to check prognostic importance of the prognostic models developed in UKALL 97/99.

4.3.3.1. Identification of Candidate Predictors in UKALL 97/99

Available variables, such as sex, age, WBC at diagnosis, CNS status, mediastinal mass, splenomegaly, hepatomegaly, BM percentage at diagnosis and marrow responses at day 8/28 were used to check potential prognostic predictors for model development in UKALL 97/99 after excluding patients in HR treatment group (n=62) and imputing missing values in marrow variables (Figure 4.1). Univariable Cox regression revealed that WBC at diagnosis in EFS and BM blast percentages at day 8 and day 28 in EFS/RFS/OS had significant prognostic values in T-cell ALL (Figure 4.4.).

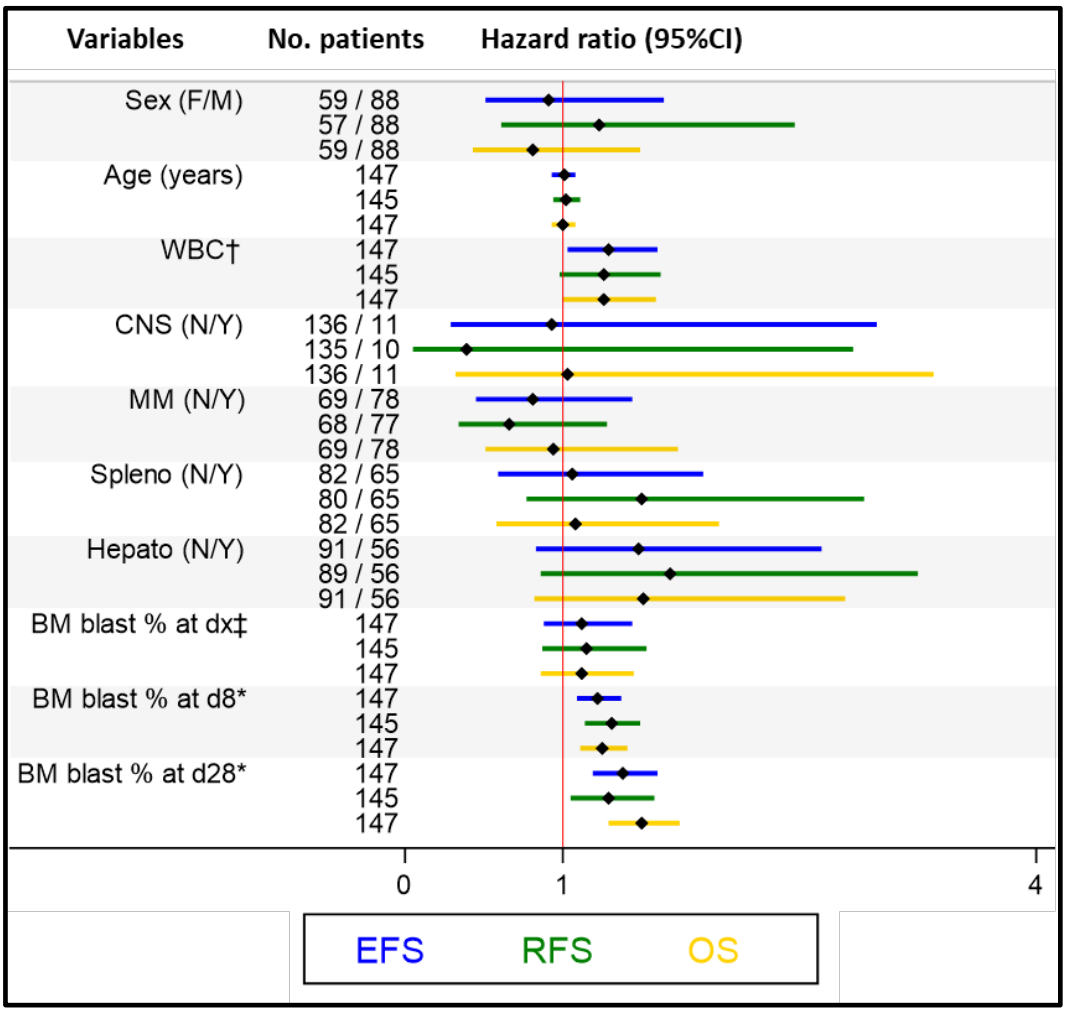

Figure 4.4. Prognostic impact of available variables from univariable Cox regression analyses in UKALL 97/99. Different colours indicate clinical endpoints, and diamond demonstrates hazard ratios (F/M, female/male; WBC, white blood cell count; CNS, central nervous system; BM, bone marrow; dx, diagnosis; d8, day 8; d28, day 28; N/Y, no/yes; EFS, event-free survival; RFS, relapse-free survival; OS, overall survival; †log transformation; ‡log transformation after reflecting values; *square root transformation).

4.3.3.2. Developing Prognostic Models in UKALL 97/99

Historically relevant factors in ALL (sex, age, WBC at diagnosis) and the variables with significant prognostic association in the univariable survival analysis (BM blast percentages at day 8 and day 28) were included in the model selection process. The assumption of Cox regression, which was used in selection process, was checked by the plot of Schoenfeld residuals and its statistical test. Horizontal line and nonsignificant *p*-value demonstrate that there is no violation of this assumption (Figure 4.5.). After backward and forward selection methods, three prognostic models with different predictors were built using EFS, RFS and OS coefficients: 1) EFS-model=square root(BM blast percentage at day 28) x 0.3309185 + log(WBC+1) x 0.2526421); 2) RFS-model=square root(BM blast percentage at day 8) x 0.2664043); 3) OS-model=square root(BM blast percentage at day 28) x 0.04161459 + log(WBC+1) x 0.2457903 (Table 4.8). The three models had significant prognostic values in UKALL 97/99 as showed in Table 4.9. The highest C-index for discrimination of relapse status belonged to the RFS-model (0.69, (0.60-0.79)), compared to the EFS-model (0.60 (0.51-0.69)) and OS-model (0.59 (0.50-0.68)) (Table 4.9.).

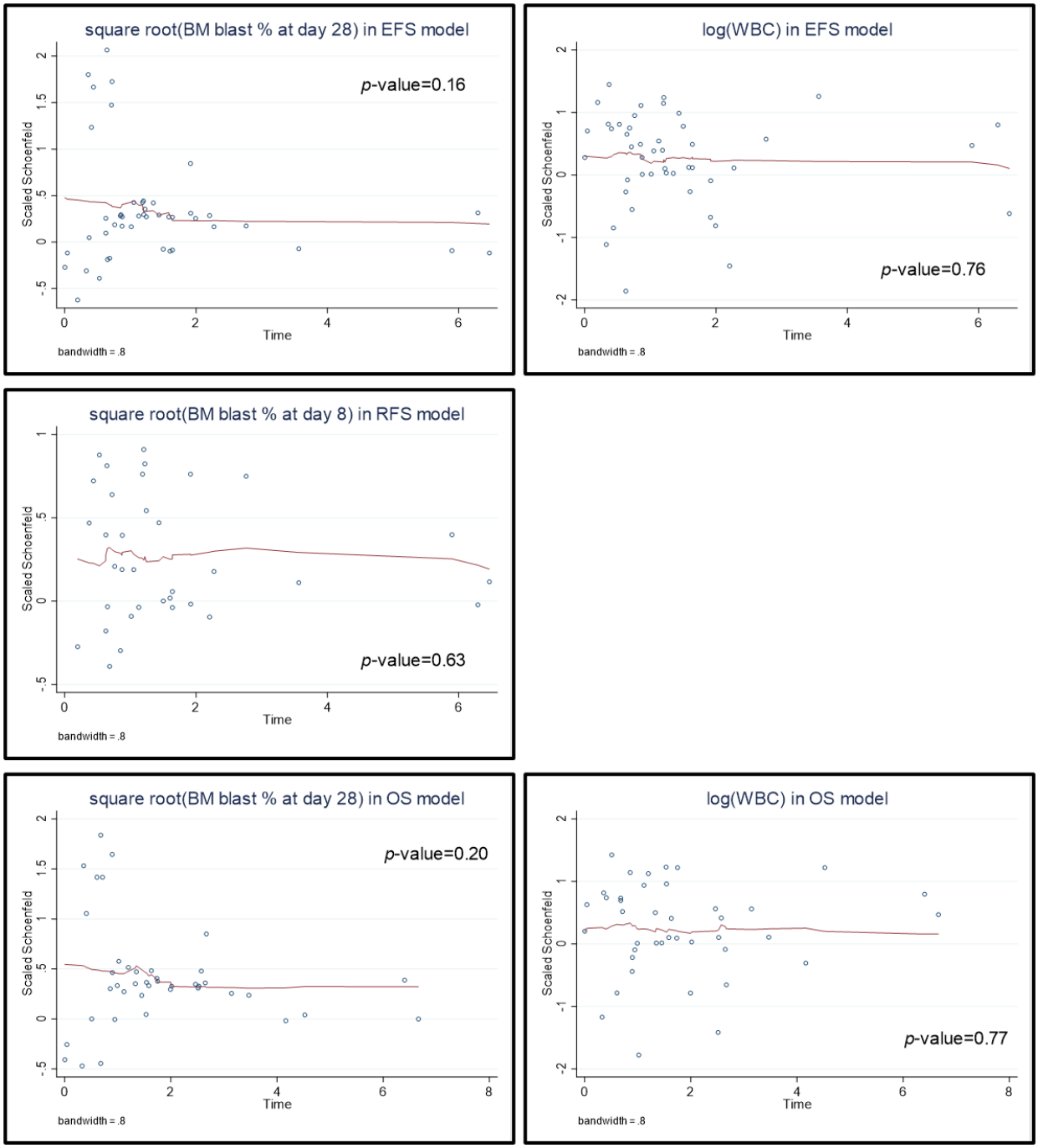

Figure 4.5. Evaluation of Cox regression assumption for the variables in baseline prognostic models developed in UKALL 97/99 (BM, bone marrow; WBC, white blood cell count; EFS, event-free survival; RFS, relapse-free survival; OS, overall survival).

EFS (Backward)	Models	Elimination	*p*-value
Step 1	BM blast % at d28* + WBC† + BM blast % at d8* + Age + Sex	Sex	0.93
Step 2	BM blast % at d28* + WBC† + BM blast % at d8* + Age	Age	0.68
Step 3	BM blast % at d28* + WBC† + BM blast % at d8*	BM blast % at d8*	0.08
Step 4	BM blast % at d28* + WBC†	WBC†	**0.02**
Step 5	BM blast % at d28* + WBC† + (BM blast % at d28* x WBC†)	(BM blast % at d28* x WBC†)	0.37
Final model	**BM blast % at d28* + WBC†**		

EFS (Forward)	Models	Addition	*p*-value
Step 1	BM blast % at d28*	BM blast % at d8*	0.05
Step 2	BM blast % at d28*	WBC†	**0.02**
Step 3	BM blast % at d28* + WBC†	Sex	0.70
Step 4	BM blast % at d28* + WBC†	Age	0.57
Step 5	BM blast % at d28* + WBC†	(BM blast % at d28* + WBC†)	0.37
Final model	**BM blast % at d28* + WBC†**		

RFS (Backward)	Models	Elimination	*p*-value
Step 1	BM blast % at d8* + WBC† + Sex + Age + BM blast % at d28*	BM blast % at d28*	0.52
Step 2	BM blast % at d8* + WBC† + Sex + Age	Age	0.45
Step 3	BM blast % at d8* + WBC† + Sex	WBC†	0.14
Step 4	BM blast % at d8* + Sex	Sex	0.10
Final model	**BM blast % at d8***		

RFS (Forward)	Models	Addition	*p*-value
Step 1	BM blast % at d8*	BM blast % at d28*	0.34
Step 2	BM blast % at d8*	WBC†	0.16
Step 3	BM blast % at d8*	Sex	0.10
Step 4	BM blast % at d8*	Age	0.81
Final model	**BM blast % at d8***		

OS (Backward)	Models	Elimination	*p*-value
Step 1	BM blast % at d28* + WBC† + BM blast % at d8* + Sex + Age	Age	0.81
Step 2	BM blast % at d28* + WBC† + BM blast % at d8* + Sex	Sex	0.62
Step 3	BM blast % at d28* + WBC† + BM blast % at d8*	BM blast % at d8*	0.09
Step 4	BM blast % at d28* + WBC†	WBC†	**0.03**
Step 5	BM blast % at d28* + WBC† + (BM blast % at d28* x WBC†)	(BM blast % at d28* x WBC†)	0.50
Final model	**BM blast % at d28* + WBC†**		

OS (Forward)	Models	Addition	*p*-value
Step 1	BM blast % at d28*	BM blast % at d8*	0.06
Step 2	BM blast % at d28*	WBC†	**0.03**
Step 3	BM blast % at d28* + WBC†	Sex	0.40
Step 4	BM blast % at d28* + WBC†	Age	0.66
Step 5	BM blast % at d28* + WBC†	(BM blast % at d28* + WBC†)	0.50
Final model	**BM blast % at d28* + WBC†**		

Table 4.8. Backward and forward selection strategies to develop baseline prognostic models in UKALL 97/99. *p*-values are from likelihood-ratio test, comparing two models with/without the variables, which were added or eliminated in the selection process (WBC, white blood cell count; BM, bone marrow; d8, day 8; d28, day 28; EFS, event-free survival; RFS, relapse-free survival; OS, overall survival; †logarithmic transformation; *square root transformation).

UKALL 97/99	No. patients	EFS-model	RFS-model	OS-model
Hazard ratio in EFS (95%CI)	147	1.55 (1.30-1.85)	1.22 (1.09-1.37)	1.53 (1.30-1.81)
Hazard ratio in RFS (95%CI)	145	1.44 (1.14-1.82)	1.31 (1.14-1.49)	1.43 (1.14-1.79)
Hazard ratio in OS (95%CI)	147	1.70 (1.41-2.04)	1.25 (1.11-1.41)	1.68 (1.41-2.00)

UKALL 97/99	No. patients	C-index (relapse)	95% CI	C-index (death)	95% CI
EFS-model	147	0.60	0.51-0.69	0.67	0.58-0.76
RFS-model	145	0.69	0.60-0.79	0.65	0.56-0.74
OS-model	147	0.59	0.50-0.68	0.67	0.59-0.76

Table 4.9. Hazard ratios and C-indexes of normalised scores of the prognostic models developed in UKALL 97/99. The prognostic indexes from the models were normalised between 0-10. Hazard ratio indicates relative change in risk associated with each one-unit increase in the model scores (EFS, event-free survival; RFS, relapse-free survival; OS, overall survival).

To further investigate one of the prognostic models, RFS-model (D8BM% model) was chosen to continue downstream analyses because of the highest C-index score in discrimination of relapse cases. Additionally, C-index scores of the models for death status were close to each other, meaning that the discriminative ability of the models was similar in the context of alive/death status (Table 4.9). The prognostic index of the RFS-model (D8BM% model) was evaluated in both UKALL 2003 and UKALL 2011. Prognostic values of the model were stable and validated properly in both trials (Table 4.10).

UKALL 2003	No. patients	RFS-model
Hazard ratio in EFS (95%CI)	326	1.18 (1.09-1.28)
Hazard ratio in RFS (95%CI)	325	1.17 (1.06-1.29)
Hazard ratio in OS (95%CI)	326	1.22 (1.11-1.34)

UKALL 2011	No. patients	RFS-model
Hazard ratio in EFS (95%CI)	278	1.22 (1.11-1.33)
Hazard ratio in RFS (95%CI)	248	1.20 (1.08-1.33)
Hazard ratio in OS (95%CI)	278	1.25 (1.11-1.39)

Table 4.10. Hazard ratios of normalised index of RFS-model (D8BM% model) in UKALL 2003 and UKALL 2011. The prognostic indexes from the models were normalised between 0-10. Hazard ratio indicates relative change in risk associated with each one-unit increase in the model scores (EFS, event-free survival; RFS, relapse-free survival; OS, overall survival).

4.3.3.3. Identifying Risk Groups Using Optimal Threshold

The D8BM% model with threshold of 1.7 stratified T-cell ALL patients into standard-risk group (<1.7) and high-risk group (≥1.7) in UKALL97/99. The threshold of 1.7 was defined according to the following criteria: 1) size of risk groups should be approximately 80% for the standard-risk group and 20% for the high-risk group because the proportion of relapse is

around 20% of all T-cell ALL cases (Chessells *et al.*, 2003; Raetz and Teachey, 2016); 2) relapse rate should be at least doubled between the risk groups. As a result of using the threshold, the standard-risk group had 115 (78%) T-cell ALL patients, while the high-risk group consisted of 32 (22%) patients with T-cell ALL (Figure 4.6.).

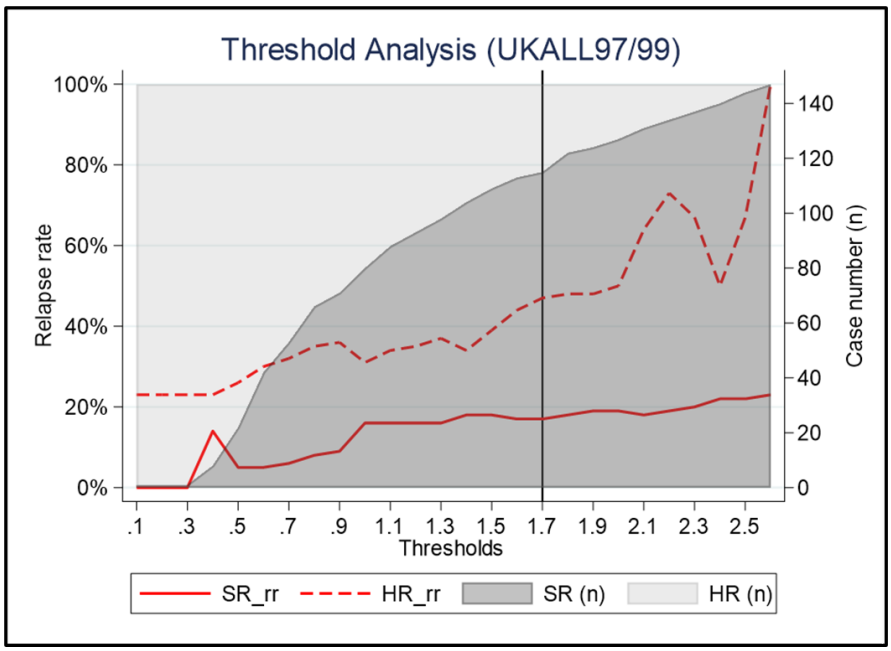

Figure 4.6. Identification of optimal threshold for the D8BM% model. Cut-off points of D8BM% score are shown on the x-axis. The y-axis on the left side indicates relapse rates, and the y-axis on the right side shows the number of patients. Solid and dashed red lines represent relapse rates of the standard-risk group and high-risk group, respectively. The line where grey colours (dark grey and bright grey) cross in the graph demonstrates the number/proportion of patients in the standard-risk group (SR, standard-risk; HR, high-risk; rr, relapse rate; n, patient number).

4.3.3.4. Clinical Characteristics and Outcomes of the Risk Groups Defined by the D8BM% Model in UKALL 97/99, UKALL 2003 and UKALL 2011

The proportion of patients in the risk groups identified by the D8BM% model with a threshold of 1.7 was consistent on UKALL 2003 (SR=256 (79%), HR=70 (21%)) and UKALL 2011 (SR=211 (76%), HR=67 (24%)) in comparison to UKALL 97/99 (SR=115 (78%), HR=32 (22%)) (Figure 4.7, Table 4.11.).

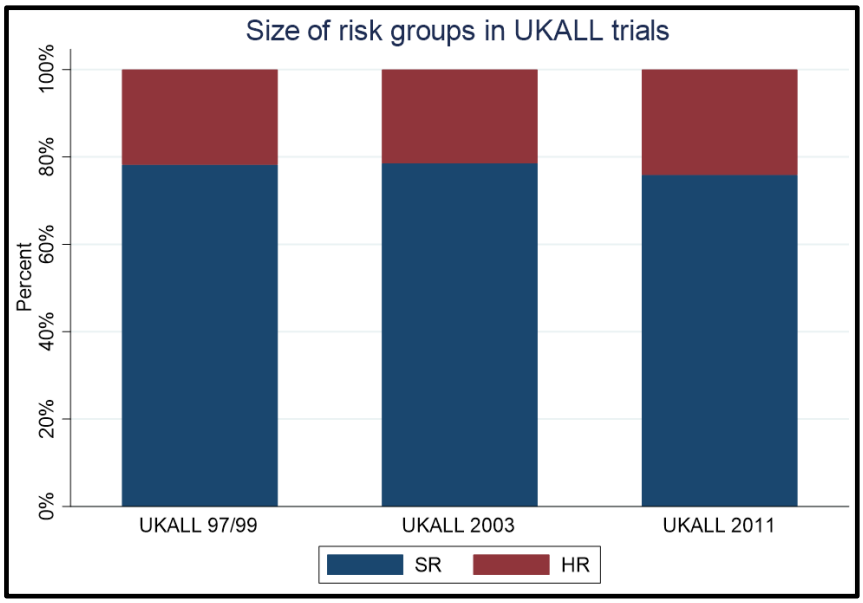

Figure 4.7. The proportion of T-cell ALL patients within the risk groups defined by the D8BM% model on UKALL 97/99, UKALL 2003 and UKALL 2011 (SR, standard-risk; HR, high-risk).

Different MRD thresholds were used to stratify cases in UKALL 2003 and UKALL 2011, so MRD positivity with these thresholds were investigated in the two trials. The proportion of positive MRD cases at the end of induction (MRD ≥0.01% in UKALL 2003 and MRD ≥0.005% in UKALL 2011) was higher in the high-risk group in UKALL 2003 (104 (52%) vs 42 (88%), *p*-value <0.001) and UKALL 2011 (96 (57%) vs 53 (95%), *p*-value <0.001), compared to positive MRD cases in the standard-risk group (Figure 4.8., Table 4.11.).

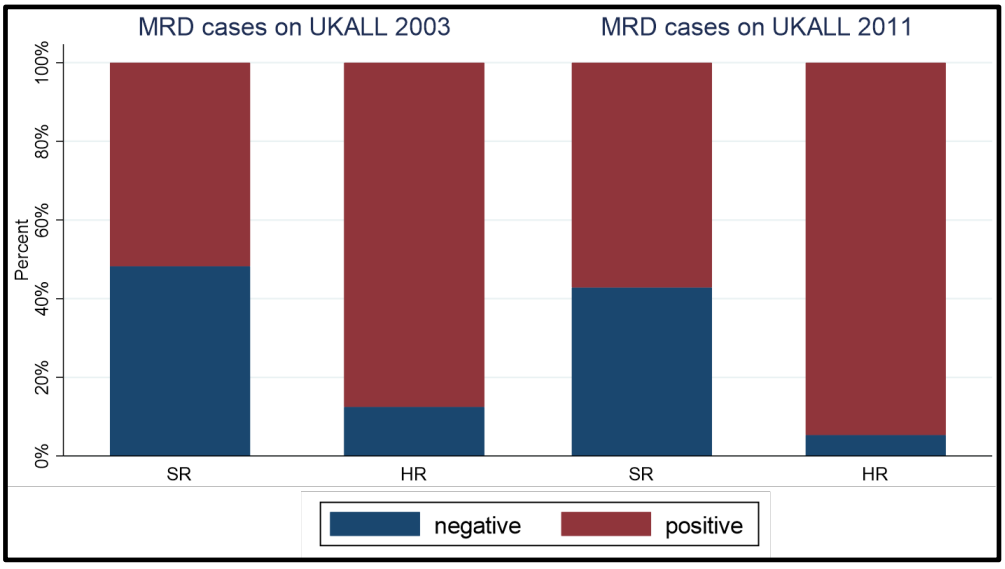

Figure 4.8. The percentage of MRD cases at the end of induction in the risk groups in UKALL 2003 and UKALL 2011 (SR, standard-risk; HR, high-risk; negative, cases with MRD <0.01% on UKALL 2003 or MRD <0.005% on UKALL 2011; positive, cases with MRD ≥0.01% on UKALL 2003 or MRD ≥0.005% on UKALL 2011).

UKALL 97/99 (no HR1 tx group)	SR (n=115, 78%)	HR (n=32, 22%)	*p*-value
Sex (male, n (%))	71 (62)	17 (53)	0.38
Age (years, median (range))	7.9 (1.1-15.2)	9.0 (1.0-16.0)	0.41
WBC (10^9/L, median (range))	57.5 (0.7-869.0)	106 (1.0-711.0)	0.10
CNS involvement (Y, n (%))	11 (10)	0 (0)	0.12
Mediastinal mass (Y, n (%))	63 (55)	15 (47)	0.43
Splenomegaly (Y, n (%))	48 (42)	17 (53)	0.25
Hepatomegaly (Y, n (%))	41 (36)	15 (47)	0.25
BM blast % at dx (median (range))	90 (1-100)	94 (40-100)	**0.03**
UKALL97-main tx (Y, n (%))	46 (100)	10 (100)	-
UKALL99-regimen C (Y, n (%))	7 (10)	19 (86)	**<0.001**
BM blast % at d8 (median (range))	8 (0-39)	62 (41-93)	**<0.0001**
BM blast % at d28 (median (range))	2 (0-42)	3 (0-93)	**0.0001**

UKALL 2003	SR (n=256, 79%)	HR (n=70, 21%)	*p*-value
Sex (male, n (%))	185 (72)	54 (77)	0.41
Age (years, median (range))	9.4 (1.4-24.3)	12.1 (2.6-23.3)	**0.01**
WBC (10^9/L, median (range))	95.0 (1.0-881.0)	145.7 (0.5-777.0)	0.09
CNS involvement (Y, n (%))	11 (4)	3 (4)	1.00
BM blast % at dx (median (range))	90 (4-100)	91 (53-100)	0.22
Regimen C (Y, n (%))	67 (26)	62 (89)	**<0.001**
BM blast % at d8 (median (range))	7 (0-40)	71 (41-98)	**<0.0001**
BM blast % at d28 (median (range))	1 (0-14)	2 (0-95)	**0.0001**
MRD % at EoI (≥0.01, n (%))	104 (52)	42 (88)	**<0.001**

UKALL 2011	SR (n=211, 76%)	HR (n=67, 24%)	*p*-value
Sex (male, n (%))	164 (78)	50 (75)	0.60
Age (years, median (range))	9.0 (1.0-24.0)	12.0 (1.0-24.0)	0.06
WBC (10^9/L, median (range))	72.1 (0.4-883.0)	154.0 (0.8-1000.0)	0.06
CNS involvement (Y, n (%))	17 (9)	2 (3)	0.26
Regimen C (Y, n (%))	104 (51)	50 (91)	**<0.001**
Trial-off patients (n (%))	6 (100)	12 (100)	-
BM blast % at d8 (median (range))	7 (0-40)	70 (42-99)	**<0.0001**
BM blast % at d28 (median (range))	1 (0-30)	2 (0-92)	**0.001**
MRD % at EoI (≥0.005, n (%))	96 (57)	53 (95)	**<0.001**

Table 4.11. Clinical characteristics of the risk groups defined by the D8BM% model. HR1 treatment group was excluded from the analysis, and missing values in marrow variables were imputed in UKALL 97/99 (WBC, white blood cell count; CNS, central nervous system; BM, bone marrow; dx, diagnosis; d8, day 8; d28, day 28; tx, treatment; Y, yes; SR, standard-risk; HR, high-risk).

The size of the risk groups defined by the prognostic model was similar in UKALL97/99, UKALL 2003 and UKALL 2011, and the model also correlated with the strongest prognostic factor (MRD) in UKALL 2003 and UKALL 2011. Prognosis of patients within the standard-risk and high-risk groups was checked in three clinical endpoints: event-free survival (EFS), relapse rate (RR) and overall survival (OS). Clinical outcomes of patients in the standard-risk group were superior to those in the high-risk group in UKALL 97/99 (77% vs 50%, *p*-value=0.003 in EFS ;17% vs 47%, *p*-value=0.0006 in RR; 80% vs 50%, *p*-value=0.0004 in OS), in UKALL 2003 (84% vs 70%, *p*-value=0.014 in EFS; 11% vs 22%, *p*-value=0.06 in RR; 89% vs 76%, *p*-value=0.01 in OS) and in UKALL 2011 (84% vs 62%, *p*-value=0.0001 in EFS; 15% vs 30%, *p*-value=0.007 in RR; 69% vs 90%, *p*-value=0.0005 in OS) (Figure 4.9.).

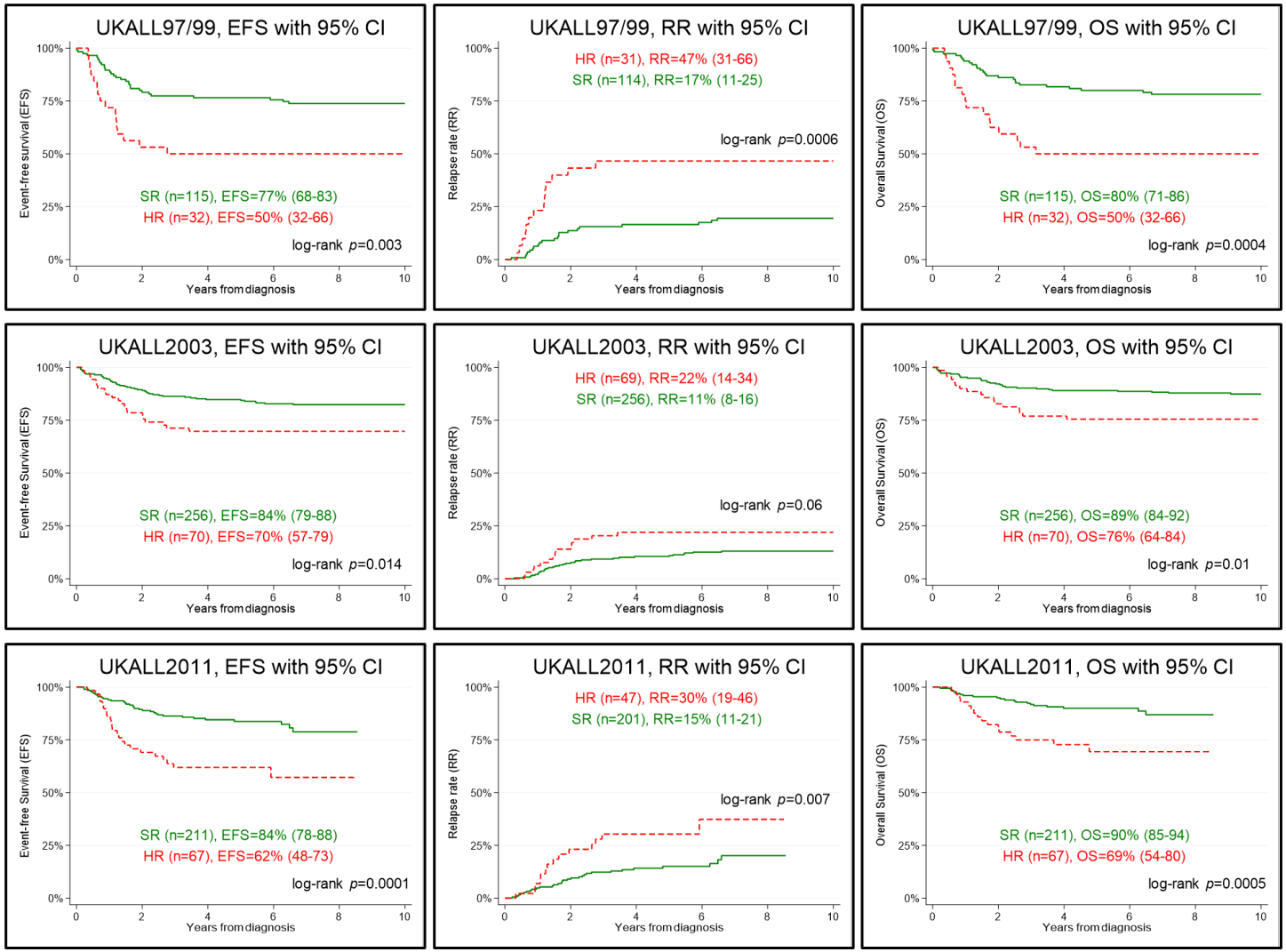

Figure 4.9. Kaplan-Meier graphs for 5-year event-free survival, relapse rate and overall survival of the risk groups. The 95% CIs were included next to the rates in the graphs (EFS, event-free survival; RR, relapse rate; OS, overall survival; SR, standard-risk; HR, high-risk; CI, confidence interval).

4.3.3.5. Model Calibration in UKALL 2003 and UKALL 2011

Calibration methods assess the ability of predicting clinical outcomes or probability of clinical events for a prognostic model, and a calibration graph compares predicted clinical outcomes/probabilities with observed clinical outcomes/probabilities. The main difference between validation and calibration is that validation reflects overall performance of a model for discrimination ability of clinical outcomes (relapse, death) in various datasets, whereas calibration is about the accuracy of predicting clinical outcomes/probabilities. The overall accuracy of estimated relapse risk of the D8BM% model was checked with calibration curve and slope test. Calibration slope in relapse-free survival analysis at five years statistically deviated from 1 in UKALL 2003, indicating a poor calibration (slope score=0.58, *p*-value=0.03); however, it did not statistically differ from 1 in UKALL 2011, demonstrating a good calibration (slope score=0.68, *p*-value=0.11) (Figure 4.10).

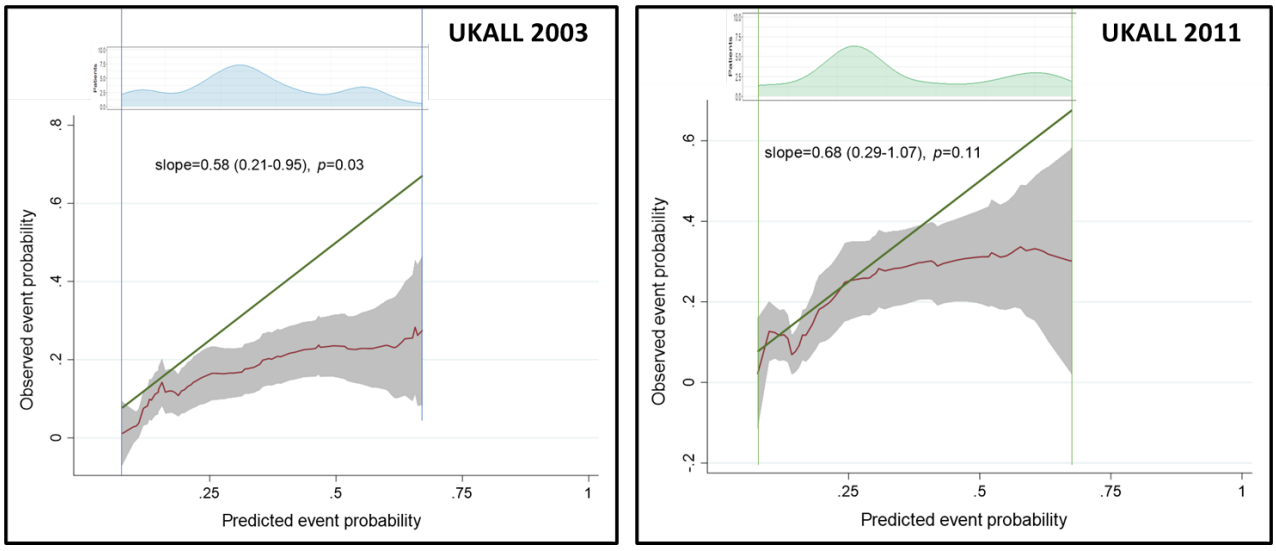

Figure 4.10. Calibration graphs and scores of the D8BM% model in relapse-free survival analysis at five years in UKALL 2003 and UKALL 2011.

4.3.4. Genetic Abnormalities within the Risk Groups in UKALL 97/99 and UKALL 2003

Type-A and type-B alterations are the key genetic abnormalities to initiate and contribute to the development of T-cell ALL. Type-A abnormalities are related to overexpression of oncogenes, and they have been previously grouped to four broader subtypes: immature subtype, TLX3 subtype, proliferative subtype and TAL/LMO subtype (Homminga *et al.*, 2012) (Table 3.4.). Although prognostic values of the genetic abnormalities are still debatable in T-cell ALL, some studies have demonstrated favourable/unfavourable impacts of genetic alterations on T-cell ALL outcomes. For example, *NOTCH1* mutations are the most common genetic alterations in T-cell ALL, and they have been observed in patients with good clinical

outcomes (Park *et al.*, 2009; Yuan *et al.*, 2022; Natarajan *et al.*, 2015; Jenkinson *et al.*, 2013) (see more information about T-cell ALL genetics in section 1.3.3.).

In this thesis, sufficient positive cases with the genetic abnormalities (n ≥10) were examined within the risk groups in a merged dataset (UKALL 97/99 + UKALL 2003). More specifically, cases with/without genetic abnormalities were identified using data containing laboratory results from FISH, MLPA, molecular/cytogenetic methods (Figure 2.3.). The high-risk group had a higher proportion of immature subtype (27% vs 15%) and TLX3 subtype (22% vs 16%), while proliferative subtype was more frequent in the standard-risk group (22% vs 6%) (Table 4.12). Regarding type-B abnormalities, comparison results between the risk groups revealed that certain genetic alterations were enriched in the standard-risk group (*NOTCH1* mutations, 72% vs 42%, *p*-value=0.001; *CDKN2A/B* deletions, 81% vs 59%, *p*-value <0.001) and in the high-risk group (*JAK1/3* mutations, 38% vs 11%, *p*-value=0.008) (Table 4.12.).

D8BM% model	Values	SR (n=371, 78%)	HR (n=102, 22%)	*p*-value
Main genetic subtypes				
Immature		33 (15%)	15 (27%)	**0.01**
TLX3		34 (16%)	12 (22%)	
Proliferative		47 (22%)	3 (6%)	
TAL/LMO		101 (47%)	25 (45%)	
Type-B abnormalities (positive)				
NOTCH1 mutation	Positive	111 (72%)	17 (42%)	**0.001**
	Negative	44 (28%)	23 (58%)	
FBXW7 mutation	Positive	28 (18%)	4 (10%)	0.24
	Negative	127 (82%)	35 (90%)	
PTEN mutation	Positive	24 (15%)	8 (20%)	0.49
	Negative	131 (85%)	32 (80%)	
PHF6 mutation	Positive	15 (18%)	5 (31%)	0.21
	Negative	70 (82%)	11 (69%)	
WT1 mutation	Positive	9 (11%)	3 (19%)	0.36
	Negative	76 (89%)	13 (81%)	
JAK1/3 mutation	Positive	10 (11%)	6 (38%)	**0.008**
	Negative	77 (89%)	10 (62%)	
N/K-RAS mutation	Positive	10 (9%)	6 (19%)	0.15
	Negative	96 (91%)	26 (81%)	
IL7R mutation	Positive	10 (12%)	1 (6%)	0.52
	Negative	75 (88%)	15 (94%)	
CDKN2A/B deletion	Positive	149 (81%)	37 (59%)	**<0.001**
	Negative	34 (19%)	26 (41%)	
PTEN deletion	Positive	20 (10%)	4 (7%)	0.45
	Negative	179 (90%)	55 (93%)	
LEF1 deletion	Positive	10 (6%)	2 (4%)	0.51
	Negative	146 (94%)	49 (96%)	

Table 4.12. Distribution of the genetic subtypes and type-B abnormalities within the risk groups in the combined dataset (UKALL 97/99 + UKALL 2003) (pos, positive cases; SR, standard-risk; HR, high-risk).

4.3.5. The D8BM% Model and MRD Interaction in UKALL 2003 and UKALL 2011

Although MRD is widely used and strongest prognostic indicator in T-cell ALL, there is no consensus on absolute threshold (0.01% or 0.1%) and optimal timepoint (end of induction or end of consolidation) (Willemse *et al.*, 2002; Burns *et al.*, 2021; Szczepański *et al.*, 2002; Schrappe *et al.*, 2011). A threshold of 0.01% was traditionally and historically used to stratify T-cell ALL patients for treatment allocation (Cazzaniga and Biondi, 2005). The MRD cut-off of 0.01% has been investigated in UKALL 2003 (Vora *et al.*, 2013), while the MRD threshold of 0.005% has been used in UKALL 2011 (Goulden, 2013). In this section, these levels (0.01% and 0.005%) were checked for the interactions between the D8BM% model and MRD. Firstly, a positive correlation was observed between the model and MRD levels at the end of induction (EOI). The standard-risk group had more negative MRD cases in UKALL 2003 (48% vs 13%) and UKALL 2011 (43% vs 5%) in comparison to the high-risk group. In parallel, cases with high MRD levels (MRD ≥5%) were less frequent in the standard-risk group in UKALL 2003 (5% vs 29%) and UKALL 2011 (4% vs 32%) (Table 4.13.).

EOI MRD (UKALL 2003)	SR (n, %)	HR (n, %)	EOI MRD (UKALL 2011)	SR (n, %)	HR (n, %)
<0.01%	97 (48)	6 (13)	<0.005%	72 (43)	3 (5)
0.01%-5%	95 (47)	28 (58)	0.005%-5%	89 (53)	35 (63)
≥5%	9 (5)	14 (29)	≥5%	7 (4)	18 (32)

Table 4.13. Correlation between the D8BM% model and MRD levels (SR, standard-risk; HR, high-risk; EOI MRD, end of induction measurable/minimal residual disease).

Secondly, Kaplan-Meier graphs with relapse rates demonstrated that the risk groups defined by the D8BM% model separated patients within MRD-intermediate risk (MRD-IR) group in UKALL 2003 (SR+MRD(0.01%-5%) vs HR+MRD(0.01%-5%), 6% vs 27%, *p*-value=0.007) and in UKALL 2011 (SR+MRD(0.005%-5%) vs HR+MRD(0.005%-5%), 14% vs 32%, *p*-value=0.02) (Figure 4.11.). On the other hand, the model did not show a clear separation for patients with negative MRD levels because of insufficient number of patients in the high-risk group.

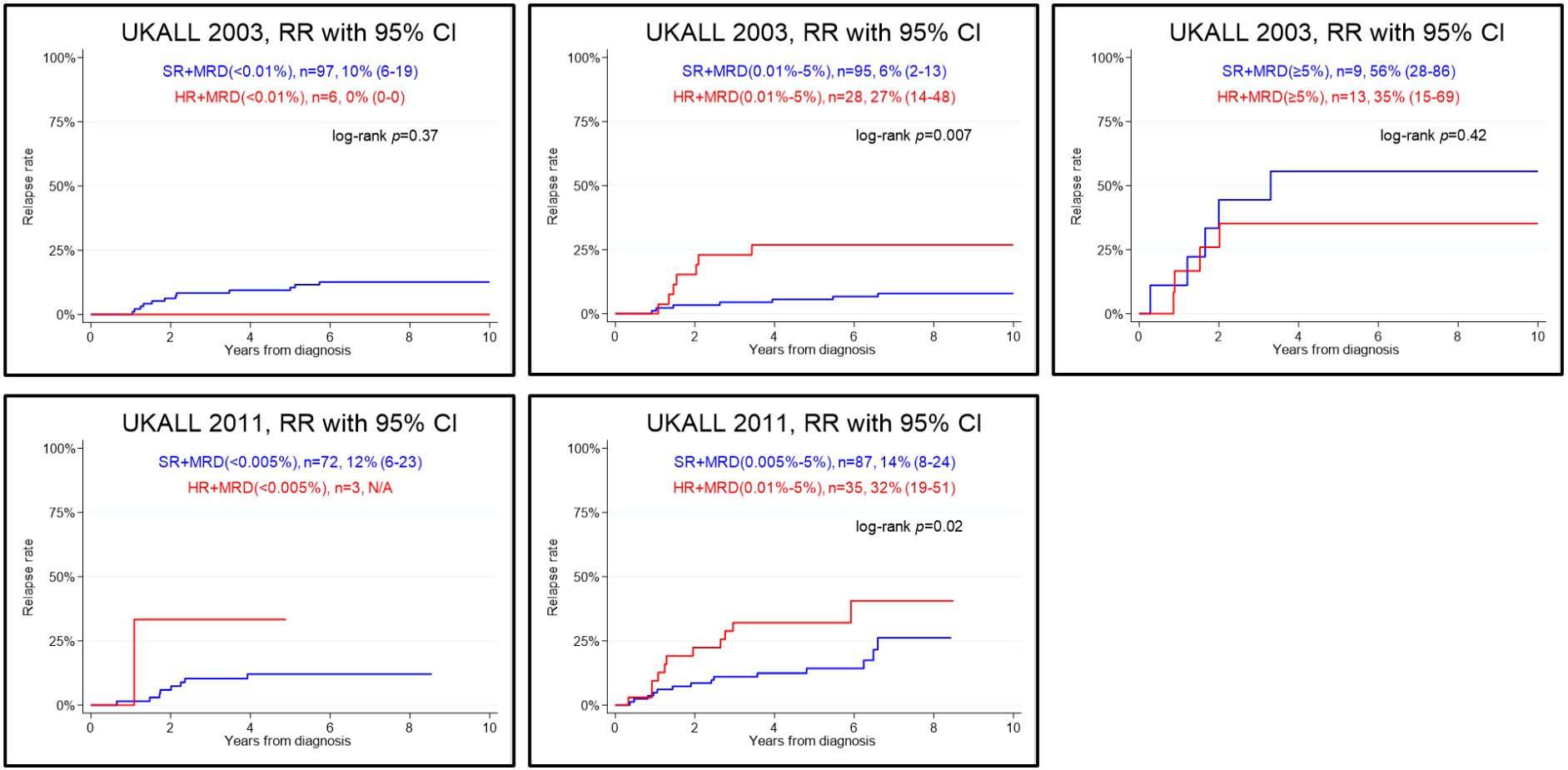

Figure 4.11. Interactions between the risk groups and MRD levels at the end of induction in terms of 5-year relapse rate. All patients with high MRD levels (≥5%) considered as no complete remission on UKALL 2011; therefore, there is no graph for this subset of patients. The 95% CIs were included next to the rates in the graphs (MRD, measurable/minimal residual disease; SR, standard-risk; HR, high-risk; RR, relapse rate; CI, confidence interval).

Lastly, discrimination abilities of the D8BM% model and MRD stratification at the end of induction were compared using the Area Under the Receiver Operating Characteristic (AUROC) method. Comparison results in the AUROC analysis indicated that there was no statistically significant difference between these two stratifications in UKALL 2003 (0.57 vs 0.54, *p*-value=0.50) and UKALL 2011 (0.60 vs 0.57, *p*-value=0.61) in terms of relapse status (Figure 4.12.).

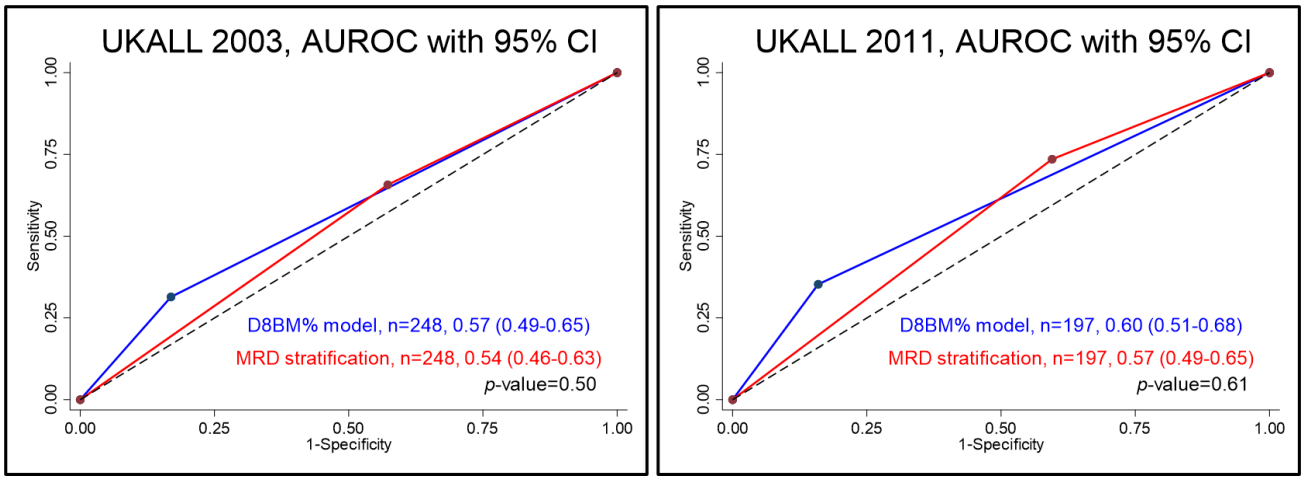

Figure 4.12. Discriminative abilities of the D8BM% model and MRD stratification at the end of induction in terms of relapse status in UKALL 2003 and UKALL 2011. MRD thresholds of 0.01% and 0.005% were used in UKALL 2003 and UKALL 2011, respectively. The 95% CIs were included next to the scores in the graphs (AUROC, Area Under the Receiver Operating Characteristic; MRD, minimal/measurable residual disease; CI, confidence interval).

Additional analysis using another MRD threshold of 0.1% showed that the number of negative MRD cases (MRD <0.1%) was higher in the standard-risk group than in the high-risk group (71% vs 27%) (Table 4.14.). Furthermore, relapse rate graphs demonstrated that the D8BM% model separated patients within negative MRD group (MRD <0.1%) in UKALL 2011 (13% vs 42%, *p*-value=0.006), but not in UKALL 2003 (9% vs 8%, *p*-value=0.76). Also, those in the MRD-IR group (0.1%-5%) were significantly separated by the model in UKALL 2003 (7% vs 30%, *p*-value=0.02), but not in UKALL 2011 (15% vs 29%, *p*-value=0.16) (Figure 4.13). There was no statistically significant difference between the model and MRD stratification in terms of separating patients with/without relapse (0.57 vs 0.60, *p*-value=0.58 in UKALL 2003; 0.60 vs 0.55, *p*-value=0.30 in UKALL 2011) (Figure 4.14.).

EOI-MRD (UKALL 2003)	SR (n, %)	HR (n, %)	EOI-MRD (UKALL 2011)	SR (n, %)	HR (n, %)
<0.1%	142 (71)	13 (27)	<0.1%	126 (75)	11 (20)
0.1%-5%	50 (25)	21 (44)	0.1%-5%	35 (21)	27 (48)
≥5%	9 (4)	14 (29)	≥5%	7 (4)	18 (32)

Table 4.14. Correlation between the D8BM% model and MRD levels in UKALL 2003 and UKALL 2011 (SR, standard-risk; HR, high-risk; EOI MRD, end of induction measurable/minimal residual disease).

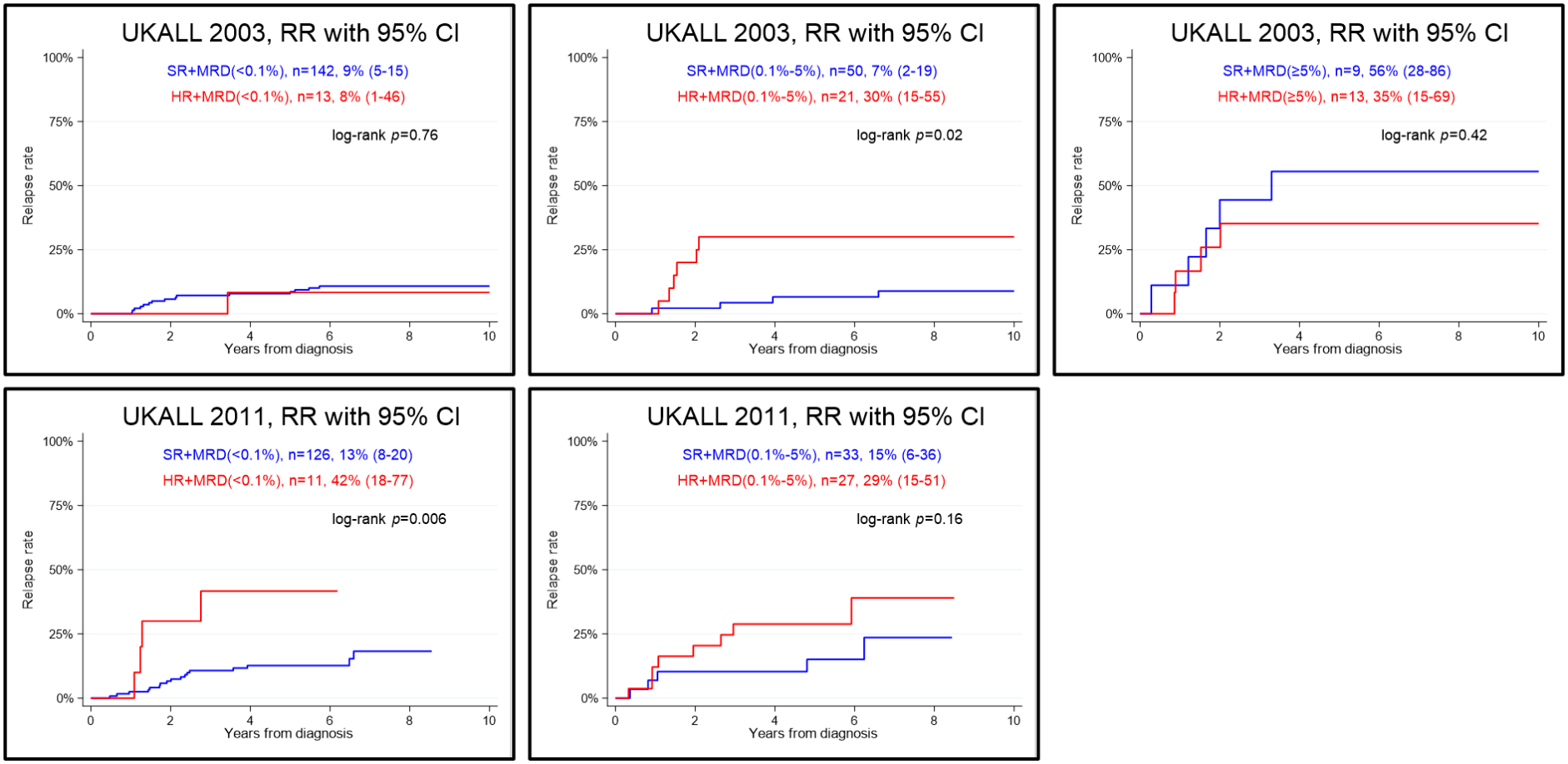

Figure 4.13. Interactions between the risk groups and MRD levels at the end of induction (<0.1%, 0.1%-5%, ≥5%) in terms of 5-year relapse rate. All patients with high MRD levels (≥5%) considered as no complete remission on UKALL 2011; therefore, there is no graph for this subset of patients. The 95% CIs were included next to the rates in the graphs (MRD, measurable/minimal residual disease; SR, standard-risk; HR, high-risk; RR, relapse rate; CI, confidence interval).

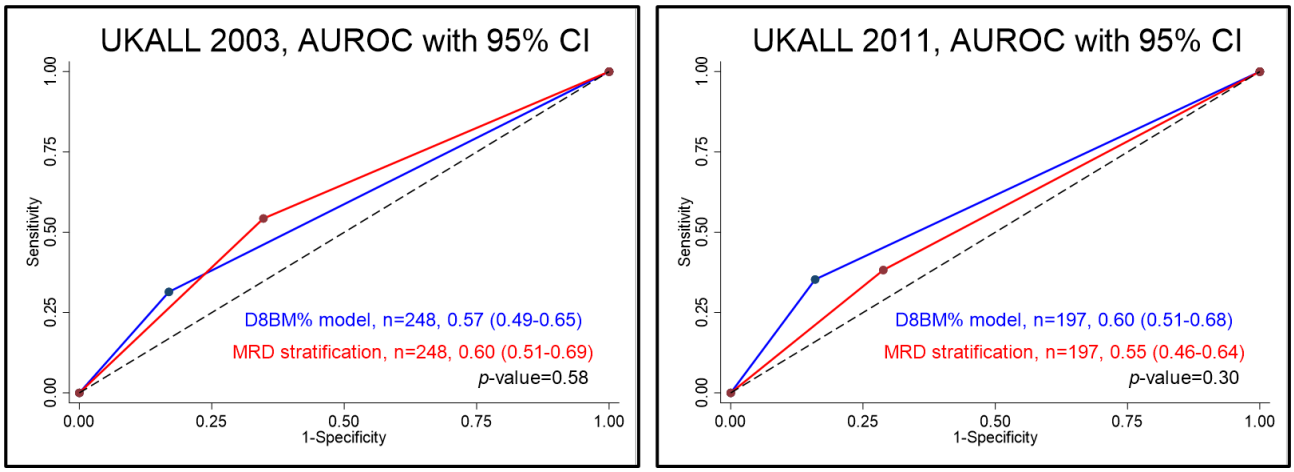

Figure 4.14. Discriminative abilities of the D8BM% model and MRD stratification at the end of induction in terms of relapse status in UKALL 2003 and UKALL 2011. MRD threshold of 0.1% was used in both trials. The 95% CIs were included next to the scores in the graphs (AUROC, Area Under the Receiver Operating Characteristic; MRD, minimal/measurable residual disease; CI, confidence interval).

4.3.6. Robustness of the D8BM% Model in Different Patients Subgroups

The D8BM% model was validated in two independent UKALL cohorts (Figure 4.7., Figure 4.9.) and correlated with the strongest prognostic factor in T-cell ALL (Table 4.13, Table 4.14.). In addition to validation and correlation results, robustness of this model in different patient subgroups (sex, age, WBC at diagnosis, MRD at the end of induction, main genetic subtypes, genetic alterations) were examined in Figure 4.15 Subgroup analysis based on relapse-free survival and overall survival indicated that the D8BM% model was robust across the subgroups within UKALL 97/99, UKALL 2003, and UKALL 2011. Specifically, WBC at diagnosis with different thresholds and MRD at the end of induction with threshold of 0.1% did not show significant variations within the risk groups in the trials.

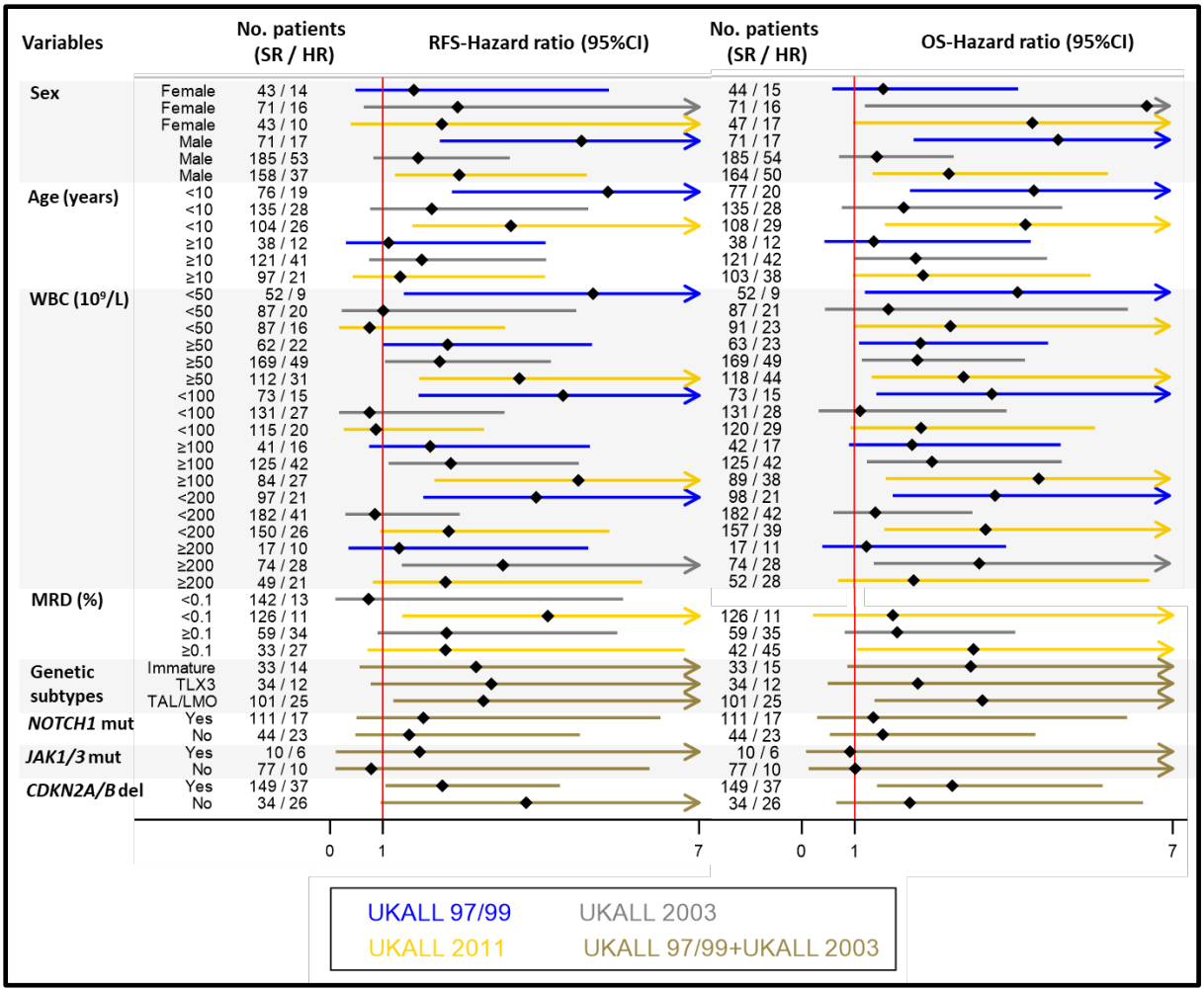

Figure 4.15. Forest plot showing the hazard ratios of the subgroups in UKALL 97/99, UKALL 2003, and UKALL 2011. UKALL 97/99 and UKALL 2003 were merged for analysis of genetic abnormalities, and the results derived from relapse-free survival (RFS) and overall survival (OS) (SR, standard-risk; HR, high-risk; WBC, white blood cell count, MRD, minimal/measurable residual disease; mut, mutations; del, deletions).

4.3.7. The D8BM% Model and Traditional Marrow Classification

Conventional marrow classification was proposed for ALL decades ago. Lymphoblast levels in the bone marrow have been evaluated in this traditional classification, categorising ALL patients into three marrow status: M1 status (lymphoblast level <5%), M2 status (lymphoblast level between 5% and 25%) and M3 status (lymphoblast level ≥25%). The marrow classification has indicated remission status of the bone marrow at the end of induction, and M3 status has been used to identify patients with rapid early response or slow early response during induction phase (Gaynon *et al.*, 1997; Hann *et al.*, 1979; Miller *et al.*, 1983; Miller *et al.*, 1980; Buchmann *et al.*, 2022).

The D8BM% model, containing only BM blast percentage at day 8, predicted clinical events in UKALL 97/99, UKALL 2003 and UKALL 2011. The threshold of the model chosen during the model development process was 1.7, which was based on size and relapse rate of the risk groups. The cut-off of the D8BM% model was equivalent to lymphoblast level of approximately 41% in the bone marrow.

In this thesis, T-cell ALL patients were categorised into marrow status groups (M1, M2 and M3) and early response groups (slow early response and rapid early response), using BM blast percentage measured at day 8. There was a statistically significant difference between patients with M1 status and M2 status in terms of relapse-free survival and overall survival in UKALL 97/99 and UKALL 2003, but not in UKALL 2011. Inconsistent results were also seen between those with M2 and M3 statuses in these three trials. Additionally, M2 and M3 groups did not show a clear separation within the standard-risk group defined by the D8BM% model (Figure 4.16.). On the other hand, slow early responders (blast level ≥25%) had significantly worse clinical outcomes than rapid early responders (blast level <25%) in all trials. However, they did not have consistent clinical outcomes in the standard-risk group (Figure 4.16.).

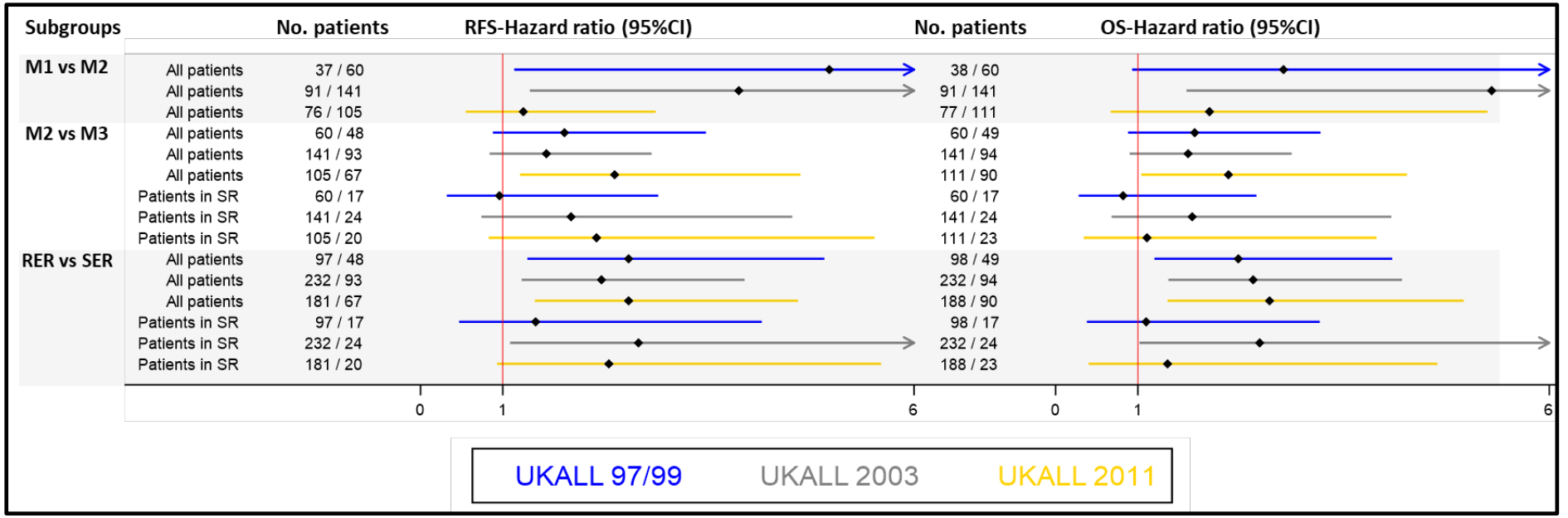

Figure 4.16. Prognostic significance of traditional marrow classification in UKALL 97/99, UKALL 2003, and UKALL 2011. T-cell ALL patients were grouped according to bone marrow blast percentage at day 8, and each colour represents different UKALL trials (M1, blast level <5%; M2, blast level of 5%-25%; M3, blast level ≥25%; RER, rapid early response (blast level <25%); SER, slow early response (blast level ≥25%); SR, standard-risk; RFS, relapse-free survival; OS, overall survival)

Additionally, as mentioned in Section 2.1.2., T-cell ALL patients in UKALL2003 and UKALL2011 were treated according to NCI-Rome risk classification, chromosomal abnormalities, and marrow responses at day 8 or day 15, using a threshold of ≥25% for treatment escalation. The equivalent threshold in the D8BM% model was around 41%. In UKALL2003, only 4 out of 24 patients were treated with regimen A among those with 25%-41% blasts at day 8. Although the number of patients received regimen A was very small, there was no statistically significant difference between those who received regimen A and those who received regimen C in EFS (25% vs 75%, *p*-value=0.10), OS (50% vs 80%, *p*-value=0.29). However, a significant difference was found in RFS (75% vs 16%, *p*-value=0.02). On the other hand, in UKALL2011, among patients with 25%-41% blast at day 8 (n=23), 4 patients were treated with regimen B and 1 patient was trial off. In this subset of patients, no significant differences observed between those received regimen B and those who received regimen C in terms of EFS (67% vs 68%, *p*-value=0.91), RFS (33% vs 38%, *p*-value=0.79), and OS (100% vs 92%, *p*-value=0.53).

4.3.8. The D8BM% Model in TARGET-ALL Dataset

The aims of TARGET-ALL project are to identify genetic alterations, which drive development and progression of ALL. The TARGET-ALL data for 265 T-cell ALL patients treated with AALL0434 treatment protocol was publicly available (Table 4.15), and it was used to validate the D8BM% model. It is important note that this data was generated for genetic analysis, so it was a selected cohort from AALL0434 trial (Dunsmore *et al.*, 2020). The median follow-up time of this cohort was 6 years.

Although the size of the standard-risk and high-groups defined by the D8BM% model was consistent (around 80% and 20%) in the TARGET-ALL cohort, there was no significant differences observed in sex, age, WBC at diagnosis, and CNS status between the risk groups (Table 4.16). In contrast, proportion of T-cell ALL patients with positive MRD (≥0.01) was higher in the high-risk group compared to the standard-risk group (75% vs 26%, p <0.001). Despite this MRD correlation, event-free survival (EFS) and overall survival (OS) rates did not statistically differ between the groups (Table 4.16). The reason of non-significant differences in clinical outcomes could be that 1) TARGET-ALL dataset was a selected cohort for genetic investigation; 2) EFS and OS rates of this cohort was higher than the UKALL cohorts, so differences in therapy might have an impact on prognostic ability of the model.

TARGET (n=265)	Total no	Values	
Sex	265	Male	202 (76%)
		Female	63 (24%)
Age, years	265	Median (range)	9.36 (1.0-30.0)
WBC, 10^9/L	265	Median (range)	99.3 (0.7-911.8)
CNS involvement	264	Yes	18 (7%)
		No	246 (93%)
NCI-Rome classification	265	Standard-risk	45 (17%)
		High-risk	220 (83%)
BM blast % at day 8	263	Median (range)	5 (0-95)
BM blast % at day 29	264	Median (range)	0 (0-10)
EOI MRD positive (≥0.01)	265	Yes	93 (35%)
		No	172 (65%)
Complete remission	262	Yes	262 (100%)
EFS at 5 years	257	% (95% CI)	91 (87-94)
OS at 5 years	262	% (95% CI)	96 (92-98)

Table 4.15. Clinical characteristics and outcomes of the patients treated with AALL0434 protocol in the TARGET-ALL cohort (WBC, white blood cell count; CNS, central nervous system; BM, bone marrow; EOI MRD, end of induction minimal/measurable residual disease; EFS, event-free survival; OS, overall survival).

D8BM% model	SR (n=215, 82%)	HR (n=48, 18%)	*p*-value
Sex (male, n (%))	164 (76)	38 (79)	0.67
Age (years, median (range))	8.9 (1.0-30.0)	10.0 (1.1-22.5)	0.16
WBC (10^9/L, median (range))	96.1 (0.7-769.3)	123.4 (2.2-911.8)	0.29
CNS involvement (Yes, n (%))	15 (7)	2 (4)	0.75
BM blast % at day 8 (median (range))	3 (0-40)	64 (42-95)	**<0.0001**
BM blast % at day 29 (median (range))	0 (0-3)	0 (0-10)	0.05
MRD % at EOI (≥0.01, n (%))	56 (26)	36 (75)	**<0.001**
Complete remission (Yes, n (%))	212 (100)	48 (100)	-
EFS at 5 years (%, 95% CI)	91 (86-94)	93 (81-98)	0.66
OS at 5 years (%, 95% CI)	95 (91-97)	98 (86-99)	0.66

Table 4.16. Clinical characteristics and outcomes of patients in the risk groups identified by the D8BM% model in TARGET-ALL cohort (WBC, white blood cell count; CNS, central nervous system; BM, bone marrow; MRD, minimal/measurable residual disease; EOI, end of induction; EFS, event-free survival; OS, overall survival; SR, standard-risk; HR, high-risk).

4.4. Discussion

The aims of Chapter 4 were to develop and validate a baseline prognostic model using available variables in the UKALL trials. UKALL VIII was excluded from the model development/validation processes because of its small sample size, and infants (<1 years old) in all trials were not included in the processes because of insufficient case number.

Models developed in UKALL X+XI, containing WBC at diagnosis and hepatomegaly, did not have statistically significant prognostic impact in overall survival, and they had statistically nonsignificant prognostic values in UKALL 97/99 (Table 4.7). Also, these factors did not have consistent prognostic importance in the UKALL trials (Figure 3.2.). The main reason why the models did not have satisfactory results in validation cohort (UKALL 97/99) could be the differences in the treatment protocols/stratifications used in UKALL X, UKALL XI and UKALL 97/99 because treatment revisions in these trials would have an impact on prognostic abilities of WBC at diagnosis and hepatomegaly (Vora *et al.*, 2006; Mitchell *et al.*, 2005). In the literature, prognostic significance of WBC at diagnosis is still controversial, and organomegalies have not prognostically been important in T-cell ALL (Shuster *et al.*, 1990; Burns *et al.*, 2021; Wei *et al.*, 2015). Therefore, in the next phase, UKALL 97/99 was used as a discovery cohort for model development, and UKALL 2003 and UKALL 2011 were used as validation cohorts. The size of risk groups defined by the D8BM% model (RFS-model), containing BM blast percentage at day 8, was consistent in the validation cohorts (Figure 4.7.), and the high-risk group had inferior clinical outcomes in comparison to the standard-risk group in the consecutive three trials (Figure 4.9.). Also, the high-risk group had more positive MRD cases (Figure 4.8.), and high MRD levels were frequently observed in this group (Table 4.13., Table 4.14.). More importantly, the D8BM% model separated T-cell ALL patients within the non-refractory group (MRD-IR group) in the validation cohorts, and the standard-risk group had a significantly lower relapse rate compared to those in the high-risk group within the MRD-IR group (Figure 4.11., Figure 4.13.). The discriminative ability of the D8BM% model was as good as the ability of MRD stratification at the end of induction in UKALL 2003 and UKALL 2011 (Figure 4.12., Figure 4.14). Additionally, this prognostic model was robust across different subgroups (sex, age, WBC, MRD, genetics) within UKALL 97/99, UKALL 2003, and UKALL 2011 (Figure 4.15.). In terms of calibration, the D8BM% model lost its significance in UKALL 2003, but not in UKALL 2011 (Figure 4.10.). The difference of calibration accuracy between the validation cohorts could be due to subjective evaluation of marrow blasts measured by light microscope.

Although thresholds of traditional marrow classification (5% and 25%) have been proposed for ALL patients and they are possibly arbitrary cut-offs, the threshold of the D8BM% model (approximately 41%), developed and validated specifically for T-cell ALL patients, was not arbitrary because it was defined according to size and relapse rate of the risk groups (Figure 4.6.). Traditional marrow classification (M1, M2, M3) did not have consistent prognostic values across UKALL 97/99, UKALL 2003 and UKALL 2011. However, patients with slow early response (blast <25%) had inferior clinical outcomes compared to rapid responders (blast ≥25%), and there was no consistent significant difference between slow and rapid responders in the standard-risk group, which was identified by the D8BM% model (Figure 4.16.). Other research groups reported that T-cell ALL patients with slow early responders experienced inferior clinical outcomes in comparison to rapid early responders (Gaynon *et al.*, 1997; Griffin *et al.*, 2000; Melchior *et al.*, 2012; Möricke *et al.*, 2013). Notably, although the size of patients who had 25%-41% blast at day 8 and received less intensive therapy was very small in UKALL2003 and UKALL2011, there was no consistent significant difference in terms of clinical outcomes. This might indicate that this subset of patients does not necessarily need intensive treatment.

Proliferative subtype (mainly *TLX1* rearrangements and *NKX2* rearrangements)*, NOTCH1* mutations, and *CDKN2A/B* deletions were enriched in the standard-risk group, whereas immature subtype and *JAK1/3* mutations were more frequent in the high-risk group (Table 4.12.). A study found that high expression of *TLX1* was related to good prognosis in childhood T-cell ALL (Ferrando *et al.*, 2002). Another report indicated that paediatric T-cell ALL patients with *TLX1* rearrangements showed a prognostic advantage compared to negative-*TLX1* cases (Kees *et al.*, 2003). Similarly, T-cell ALL patients with *NOTCH1* mutations tended to have good early treatment response (Ferrando, 2010) and were less frequent in induction failure group, compared to treatment-response group (O'Connor *et al.*, 2023). Jenkinson and colleagues reported that those with mutations in *NOTCH1* had better survival rate, but no significant difference in event-free survival and relapse rates in UKALL 2003, in comparison to patients without *NOTCH1* mutations (Jenkinson *et al.*, 2013). On the other hand, other studies reported that *JAK* mutations were not associated with adverse clinical outcomes (Vicente *et al.*, 2015), and *CDKN2A/B* deletions did not have the same consistent prognostic direction (Mirji *et al.*, 2016; Krieger *et al.*, 2010). Another study demonstrated that the presence of *JAK* mutations was positively correlated with *CDKN2A/B* deletions and was linked to poor clinical outcome (Mullighan *et al.*, 2009). Despite the fact

that some researchers concluded that ETP-ALL (a subgroup of immature subtype in this thesis) was associated with worse chemotherapy response and clinical outcomes compared to typical T-cell ALL (Inukai *et al.*, 2012; Coustan-Smith *et al.*, 2009; Allen *et al.*, 2013), other people showed that there was no statistically significant difference in clinical outcomes (Wood *et al.*, 2014; Patrick *et al.*, 2014a). Overall, although the standard-risk and high-risk groups have had more patients with low-risk genetic features (*NOTCH1* mutations, *TLX1* rearrangements) and high-risk feature (ETP-ALL), respectively, the literature is not consistent in terms of prognostic value of the known genetic abnormalities/subtypes.

The limitations of this chapter were (1) exclusion of BM blast percentage at day 15 due to high level of missing values; (2) imputation method used to replace missing values with estimated values in the marrow variables in combined UKALL X+XI and UKALL 97/99 datasets; (3) relatively small sample size of UKALL 97/99 after excluding patients in HR1 treatment group; (4) small sample size of patient subsets of interest and available genetic abnormalities in UKALL97/99 and UKALL 2003; (5) unavailability of MRD in early trials. Although there were limitations for the retrospective analyses in this chapter, the prognostic model was validated in the following trials where patients were differently randomised and stratified into treatment groups.

In summary, MRD is the only strong prognostic factor and widely used predictor for risk stratification in T-cell ALL, and early treatment response variables are associated with good/poor prognosis. Also, the D8BM% model correlated with MRD at the end of induction. A report by Basso and coworkers demonstrated positive correlation between early marrow response at day 15 and MRD at day 15 in ALL (Basso *et al.*, 2009). Therefore, this model, which has been validated in two independent trials, can stratify children/young adults with T-cell ALL to treatment groups at very early stage of the therapy (at day 8), and can make EOI MRD stratification more efficient for treatment allocation at day 28 as shown in Figure 4.17, which chronologically shows the D8BM% model at day 8 and EOI MRD stratification at day 28. Another beneficial side of using the D8BM% model is that measuring the lymphoblast level is not as expensive as measuring MRD; hence, it can be used in countries with limited financial resources. However, a gold standard method to measure the lymphoblast level in bone marrow should be established.

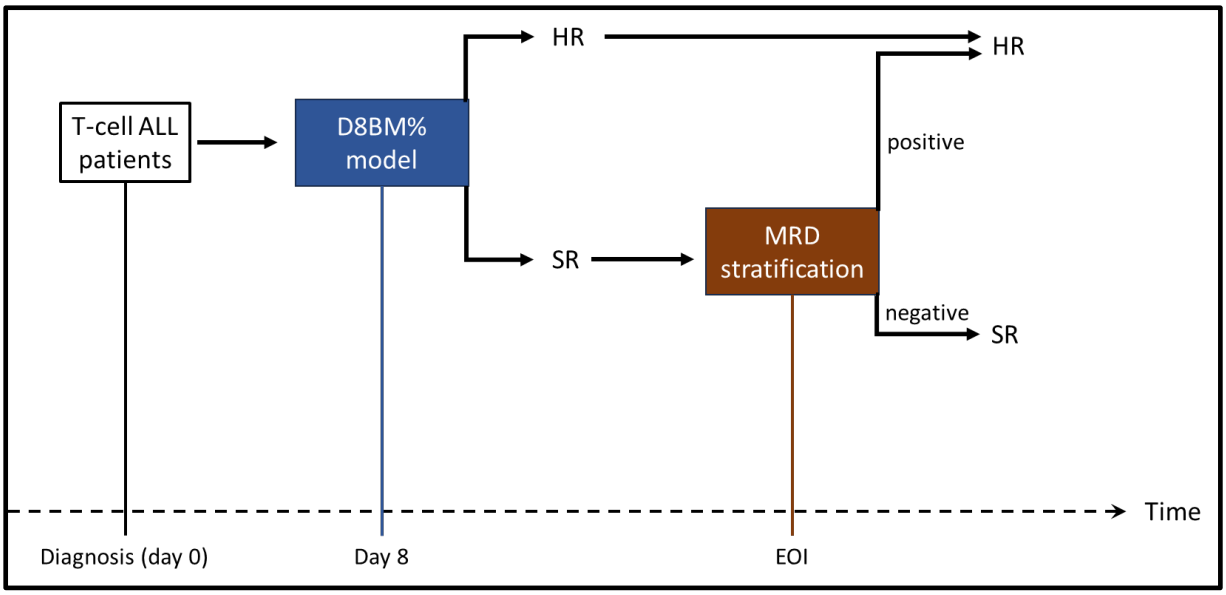

Figure 4.17. Possible clinical usefulness of the D8BM% model with MRD stratification at the end of induction. The prognostic model based on marrow measurement at day 8 and MRD stratification at EOI (day 28) were shown in chronological order (SR, standard-risk; HR, high-risk; EOI, end of induction).

Chapter 5. Dynamic Prognostic Models for Risk Stratification in T-cell Acute Lymphoblastic Leukaemia

5.1. Introduction

Both static and dynamic models can be used for the prediction of clinical outcomes with different methods. Static models are based on key prognostic factor(s) measured at a single-time point, whereas dynamic prediction models indicate a trend of key prognostic factor(s) over time using data at multiple time points. Data measured/collected during therapy or follow-up time are required for the development of dynamic prognostic models; therefore, static prognostic models, which do not need longitudinal data for the variable(s) of interest, are commonly used in cancer research (Halabi, Li and Luo, 2019). As discussed in Chapter 1, robust and validated prognostic models are needed for the prediction of T-cell ALL prognosis, and to our knowledge, there is currently no dynamic prognostic model for T-cell ALL.

In the previous chapter, a static prognostic model (D8BM% model), consisting of early treatment marrow response measured at day 8, has been developed and validated in two independent UKALL trials, and it has added a prognostic value to minimal/measurable residual disease (MRD) at the end of induction. Hence, using early marrow and/or MRD responses at multiple time points for the development of dynamic prognostic models can provide more information about prediction of future clinical events.

The aims of this chapter are to develop and validate dynamic prognostic models for risk prediction of clinical events in T-cell ALL using data from UKALL 2003 and UKALL 2011.

The objectives of the fifth chapter are to:

1) Develop and validate dynamic prognostic models with marrow and MRD values using time series regression and Area Under the Curve (AUC) method.

2) Compare clinical characteristics and outcomes of risk groups defined by the dynamic models with the static model (D8BM% model) and MRD stratification.

3) Investigate dynamics of MRD responses at different time points (week 5, week 9, and week 14) for patients with high-risk features.

5.2. Data and Methods

UKALL 2003 (n=386) and UKALL 2011 (n=321) were used to develop and validate dynamic prognostic models. In UKALL 2003, missing diagnostic bone marrow (BM) values (n=93) were imputed with random values ranging from 90% to 99%. Measurement of BM blast at diagnosis was not available in UKALL 2011; therefore, this variable was randomly generated within a range of 90 to 99%.

Time series regression (Nonlinear Mixed-Effects (NLME) regression) was one of the methods used to develop a dynamic model in UKALL 2003. Nonlinear mixed-effects regression technique has become popular in medical research recently, for example, in pharmacokinetics and HIV (human immunodeficiency virus) dynamics. This method combines fixed effects (factors of interest, such as predictors) and random effects (factors of interest showing variations within subgroups, such as patient) to analyse the repeated measurements (Davidian and Giltinan, 2003). The regression model defines a nonlinear relationship between explanatory (predictor) variable(s) and response (outcome) variable(s); therefore, the NLME model with an exponential decay function was used to examine the relationship between trend of treatment responses of T-cell ALL patients (marrow responses and MRD response at the end of induction (EOI)) and clinical events. All marrow and MRD values were re-scaled within a range of 0-10 (MRD values of >100% were considered as maximum, so it was re-scaled to 10), and the actual measurement days were used in the regression modelling process. In UKALL 2003, patients in the standard-risk group and high-risk group according to NCI-Rome classification were treated with regimen A and regimen B, respectively. In addition to the treatment groups, early marrow responses (BM blast percentages at day 8 and day 15) were evaluated for a need of intensified treatment (regimen C). In UKALL 2011, all T-cell ALL patients were assigned to the NCI-Rome high-risk group and allocated to regimen B. The same early marrow classification was applied to these patients in the trial (Figure 1.6.; Table 3.1.). As a result, in this chapter, these early time points (around day 8 and/or around day 15) were the time points when early treatment response data were collected in both trials. Day 0 was accepted as the time of measuring diagnostic marrow percentage. Marrow measurements from day 6 to day 10 were considered as marrow response at day 8, while marrow percentages measured from day 12 to day 19 were evaluated as marrow response at day 15. In case of missing measurement times, day 8 and day 15 were assigned to patients treated with regimen B or regimen A, respectively. Day 28 was considered to be the MRD measurement time for patients with MRD values in

UKALL 2003, whereas it was day 29 in UKALL 2011. Marrow values (percentages) were determined using morphological assessment (light microscopy), while MRD values were based on Polymerase Chain Reaction (PCR).

The bone marrow and MRD values were the inputs to the time series regression model (NLME model), and coefficient 1 (estimated baseline level), coefficient 2 (estimated maximum level) and coefficient 3 (estimated decline rate) were the outputs from the regression model (Figure 5.1.). Coefficient 3 indicates a trend of response to treatment. T-cell ALL patients with high coefficient 3 have quick decline trends, meaning a rapid treatment response, whereas negative or low coefficient 3 demonstrates no response or slow treatment response, respectively.

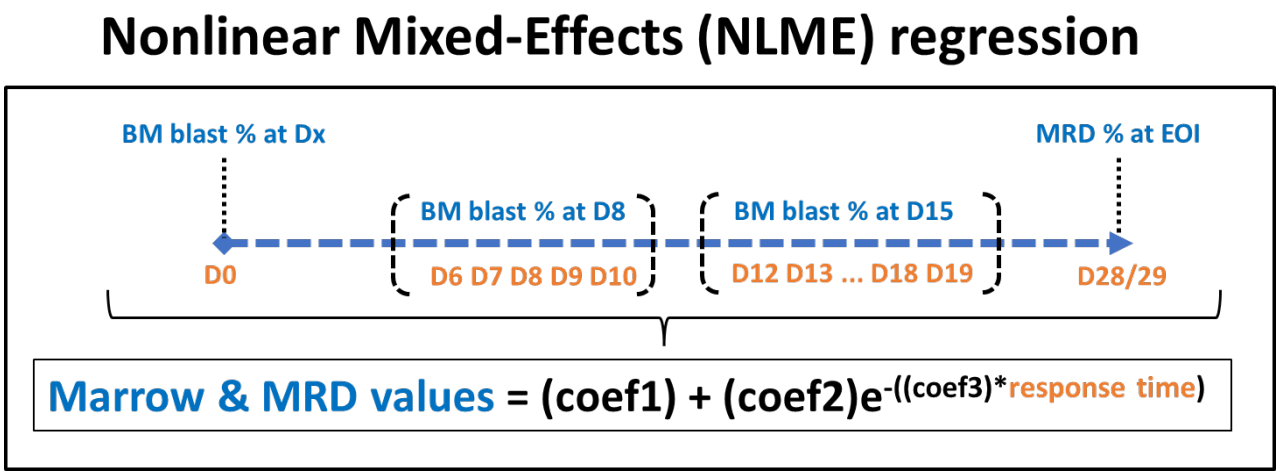

Figure 5.1. Inputs (blue and orange colours) and outputs (black colour) of a Nonlinear Mixed-Effects (NLME) model for the identification of treatment response trend in UKALL 2003 and UKALL 2011 (BM, bone marrow; Dx, diagnosis; D, day; MRD, minimal/measurable residual disease; coef, coefficient).

Another method used for the dynamic model development was the Area Under the Curve (AUC). This method measures the area under the curves of marrow and/or MRD variables for each patient, using a mathematical formula (trapezoid formula). The sum of all these areas gives an overall score that represents a trend score of treatment response (Figure 5.2.). The same rules/criteria in the NLME model development (patient selection, actual measurement days, re-scaling the values) were applied to the AUC model for calculation of the areas. Patients with low AUC scores tend to have quicker treatment responses than those with high AUC scores.

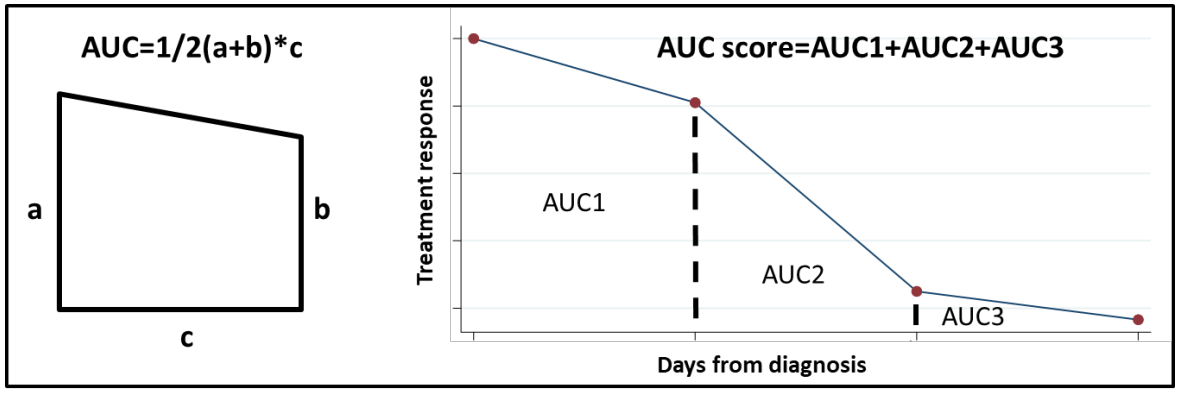

Figure 5.2. The calculation of the AUC model score using the trapezoid formula (a, the highest marrow value; b, the lowest marrow value; c=time in days; AUC, Area Under the Curve).

T-cell ALL patients treated with intensive therapy (regimen C) on UKALL 2011 had MRD responses at multiple time points (week 5, week 9 and week 14). The dynamics of MRD responses over time was therefore investigated for these patients in UKALL 2011. MRD values (percentages) were used for further investigation of MRD dynamics, using the AUC method.

Comparison tests, as mentioned in Chapter 2, were performed to compare clinical characteristics of patients in the risk groups defined by the static and dynamic models. Accuracy of the NLME model was checked by correlation and R-squared scores, using predicted and observed values. The Kolmogorov-Smirnov test was used to compare distributions of the model scores in both trials. Time-to-event analyses (Cox regression and Kaplan-Meier curves) were performed to determine prognosis of the risk groups. The Area Under the Receiver Operating Characteristic (AUROC) method was used for the comparison of discriminative abilities of the models in terms of relapse status.

5.3. Results

5.3.1. Development and Validation of the Dynamic Models for Early Response in UKALL 2003 and UKALL 2011

5.3.1.1. Distribution Patterns of the NLME and AUC Trend Scores

T-cell ALL patients with marrow values at two or more time points (at diagnosis and day 8 and/or day 15) and with complete MRD values at the end of induction were included in the dynamic model development processes for early response trend in UKALL 2003 (n=237) and UKALL 2011 (n=173). NLME trend score (coefficient 3) and AUC trend score were generated from the NLME model and AUC model, respectively, using re-scaled marrow and MRD values. The distribution patterns of the trend scores from these two dynamic models are shown in Figure 5.3. Although the distributions of NLME trend score were different between the trials (p-value <0.001), the difference in the distributions of AUC trend score was not statistically significant (p-value=0.17). BM blast percentage at day 15 was not available in UKALL 2011, so the dynamic models were validated without early marrow response at day 15 in this trial. The lack of this early marrow response could affect the distributional difference of the NLME model score, but interestingly not the AUC model score. It means that the NLME model can be more sensitive to changes in response values than the AUC model; therefore, this can be the main reason for the distributional difference of the NLME model in UKALL 2003 and UKALL 2011.

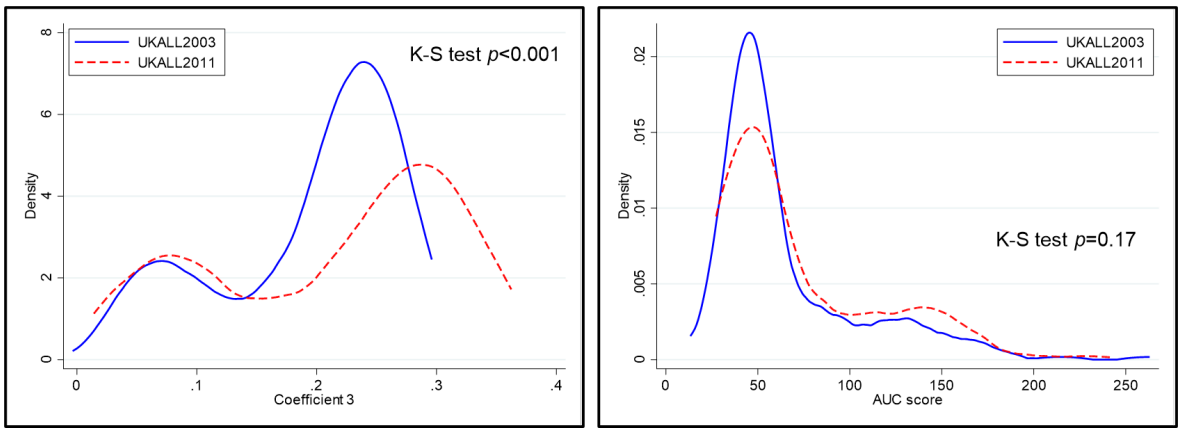

Figure 5.3. The distributions of the NLME trend score (coefficient 3) and AUC trend score in UKALL 2003 and UKALL 2011 (K-S, Kolmogorov-Smirnov; NLME, Nonlinear Mixed-Effects; AUC, Area Under the Curve).

5.3.1.2. Accuracy of Predicted Values from the NLME Model

To check accuracy of the NLME model, the predicted (estimated) response values from the model were compared to observed (actual) response values in the trials. Observed and predicted values were highly correlated in UKALL 2003 (correlation=0.98, p-value <0.0001)

and UKALL 2011 (correlation=0.99, *p*-value <0.0001). In addition to the positive correlation, R-squared score (R^2) was high in these trials, which means that the majority of the predicted response values were close to the observed response values (R^2=97% in UKALL 2003 and R^2=98% in UKALL 2011) (Figure 5.4.)

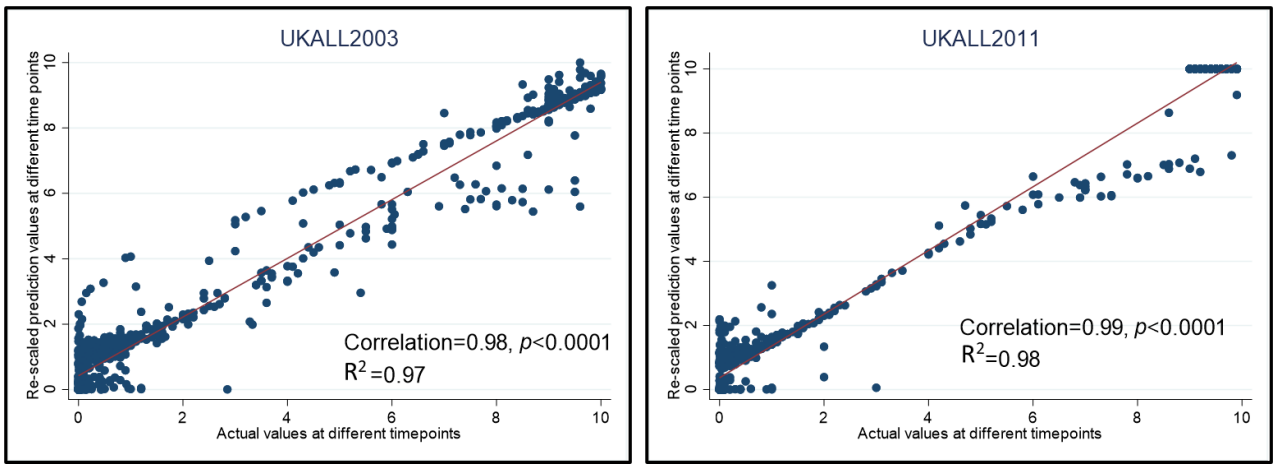

Figure 5.4. Correlation between the observed and predicted response values in UKALL 2003 and UKALL 2011 (R^2, R-squared score).

5.3.1.3. Identification of Optimal Thresholds for the Dynamic Models in UKALL 2003

The trend scores from the NLME and AUC models were used to stratify T-cell ALL patients into two risk groups: standard-risk (SR) and high-risk (HR). The optimal thresholds were selected according to proportion of patients in each risk group in UKALL 2003 (~80% for SR vs ~20% for HR) and relapse rates (at least doubled relapse rates between SR and HR). Therefore, cut-offs of 0.11 and 90 were used in the NLME and AUC models for risk stratification, respectively. The numbers of patients within the SR and HR groups defined by the NLME model were 192 (81%) and 45 (19%), respectively, while there were 190 (80%) patients in the SR group and 47 (20%) patients in the HR group identified by the AUC model (Figure 5.5.; Table 5.1).

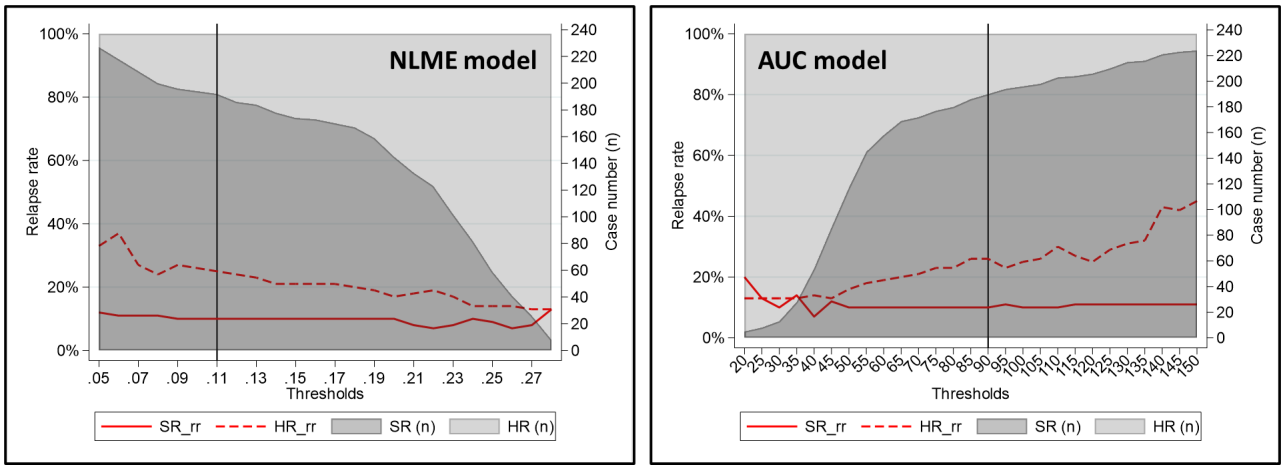

Figure 5.5. Threshold analysis for the NLME and AUC models in UKALL 2003. Different thresholds of the NLME and AUC scores are shown on the x-axis. The y-axis on the left side indicates relapse rates, and the y-axis on the right side shows the number of patients. Solid and dashed red lines represent relapse rates of the standard-risk and high-risk groups, respectively. The line where grey colours (dark grey and bright grey) cross in the graph demonstrates the number/proportion of patients in the standard-risk group (NLME, Nonlinear Mixed-Effects; AUC, Area Under the Curve; SR, standard-risk; HR, high-risk; rr, relapse rate; n, patient number).

5.3.1.4. Clinical Characteristics and Outcomes of the Risk Groups in UKALL 2003

In UKALL 2003, treatment responses during induction phase can be seen in Figure 5.6. The graphs show that the standard-risk group (n=192, NLME trend score ≥0.11) had quicker response (sharper decline curves) than the high-risk group (n=45, NLME trend score <0.11) in the NLME model. Similarly, patients in the standard-risk group categorised by the AUC model (n=190, AUC trend score <90) had a rapid response in comparison to those in the high-risk group (n=47, AUC trend score ≥90). The risk groups from these dynamic models shared similar treatment trend responses during induction phase in UKALL 2003.

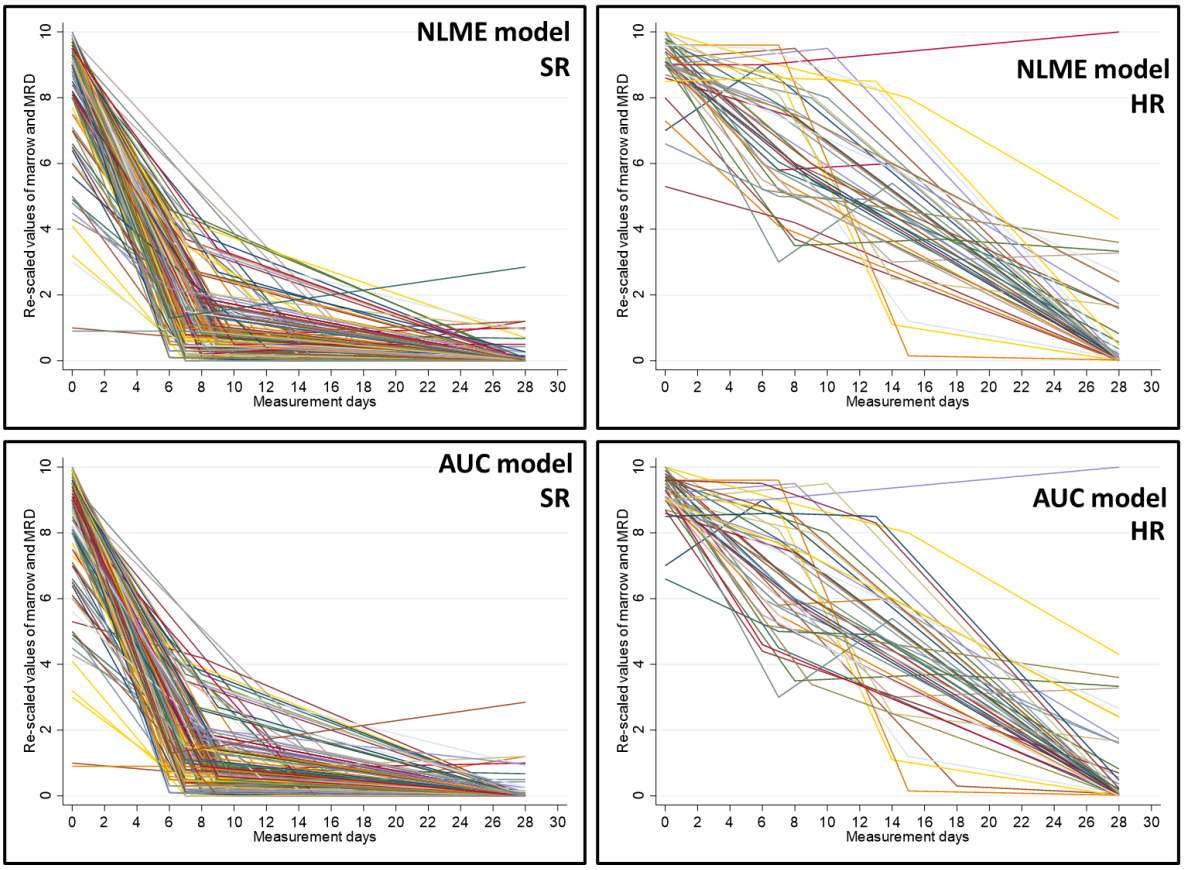

Figure 5.6. Treatment response trends of patients during induction phase in the standard-risk (SR) and high-risk (HR) groups defined by the NLME and AUC models in UKALL 2003 (NLME, Nonlinear Mixed-Effects; AUC, Area Under the Curve; MRD, minimal/measurable residual disease).

In UKALL 2003, comparison of clinical characteristics of the risk groups identified by the dynamic models revealed that patients in the standard-risk group were younger than patients in the high-risk group (median in years; 9.2 vs 12.9, *p*-value=0.003 in the NLME model; 9.2 vs 12.9, *p*-value=0.006 in the AUC model). Also, a smaller proportion of patients were treated with intensive therapy in the standard-risk groups (32% vs 89%, *p*-value <0.001 in the NLME model; 32% vs 87, *p*-value <0.001 in the AUC model). Notably, the risk stratification

was based on high-risk cytogenetics and slow early marrow response (≥25%) at day 8 or day 15 in the UKALL2003 protocol. Additionally, the standard-risk groups had lower marrow percentages at day 8 (7% vs 61%, *p*-value <0.0001 in the NLME model; 7% vs 60%, *p*-value <0.0001 in the AUC model) and at day 28 (1% vs 2%, *p*-value <0.003 in the NLME model; 1% vs 2%, *p*-value <0.003 in the AUC model), and a smaller proportion of patients with positive MRD (MRD ≥0.01%) (54% vs 87%, *p*-value <0.001 in the NLME model; 54% vs 85%, *p*-value <0.001 in the AUC model). There was no statistically significant difference in sex, white blood cell count (WBC) at diagnosis or central nervous system (CNS) involvement between the risk groups (Table 5.1.).

NLME model	SR (n=192, 81%)	HR (n=45, 19%)	*p*-value
Sex (male, n (%))	142 (74)	37 (82)	0.25
Age (years, median (range))	9.2 (1.4-24.3)	12.9 (3.2-23.3)	**0.003**
WBC (10^9/L, median (range))	104.3 (2.5-881.0)	144.6 (0.5-777.0)	0.18
CNS involvement (Y, n (%))	10 (5)	2 (4)	1
Regimen C (Y, n (%))	62 (32)	40 (89)	**<0.001**
BM blast % at d8 (median (range))	7 (0-46)	61 (30-96)	**<0.0001**
BM blast % at d28 (median (range))	1 (0-13)	2 (0-70)	**0.003**
MRD % at EOI (≥0.01, n (%))	103 (54)	39 (87)	**<0.001**

AUC model	SR (n=190, 80%)	HR (n=47, 20%)	*p*-value
Sex (male, n (%))	142 (75)	37 (79)	0.60
Age (years, median (range))	9.2 (1.4-24.3)	12.9 (3.2-23.3)	**0.006**
WBC (10^9/L, median (range))	98.9 (2.5-881.0)	148.0 (0.5-777.0)	0.06
CNS involvement (Y, n (%))	10 (5)	2 (4)	1
Regimen C (Y, n (%))	61 (32)	41 (87)	**<0.001**
BM blast % at d8 (median (range))	7 (0-43)	60 (30-96)	**<0.0001**
BM blast % at d28 (median (range))	1 (0-8)	2 (0-70)	**0.003**
MRD % at EOI (≥0.01, n (%))	102 (54)	40 (85)	**<0.001**

Table 5.1. Clinical characteristics of patients in the risk groups from the NLME and AUC models in UKALL 2003 (WBC, white blood cell count; CNS, central nervous system; BM, bone marrow; d8, day 8; d28, day 28; MRD, minimal/measurable residual disease; EOI, end of induction; Y, yes; NLME, Nonlinear Mixed-Effects; AUC, Area Under the Curve; SR, standard-risk; HR, high-risk).

Of 237 patients, five cases in the standard-risk group defined by the NLME model moved to the high-risk group defined by the AUC model, whereas three cases moved from the NLME high-risk group to the AUC standard-risk group (Figure 5.7.). This subset of patients moving to the high-risk group (n=5) consisted of two females and three males, and they all achieved complete remission at the end of induction. One of these cases had high MRD level (≥5%) and relapse, but there were no deaths in this group during the follow-up time. On the other hand, of three patients moving to the AUC standard-risk group, all were males, achieved

complete remission and had no any poor clinical features (high MRD, second tumour, relapse or death).

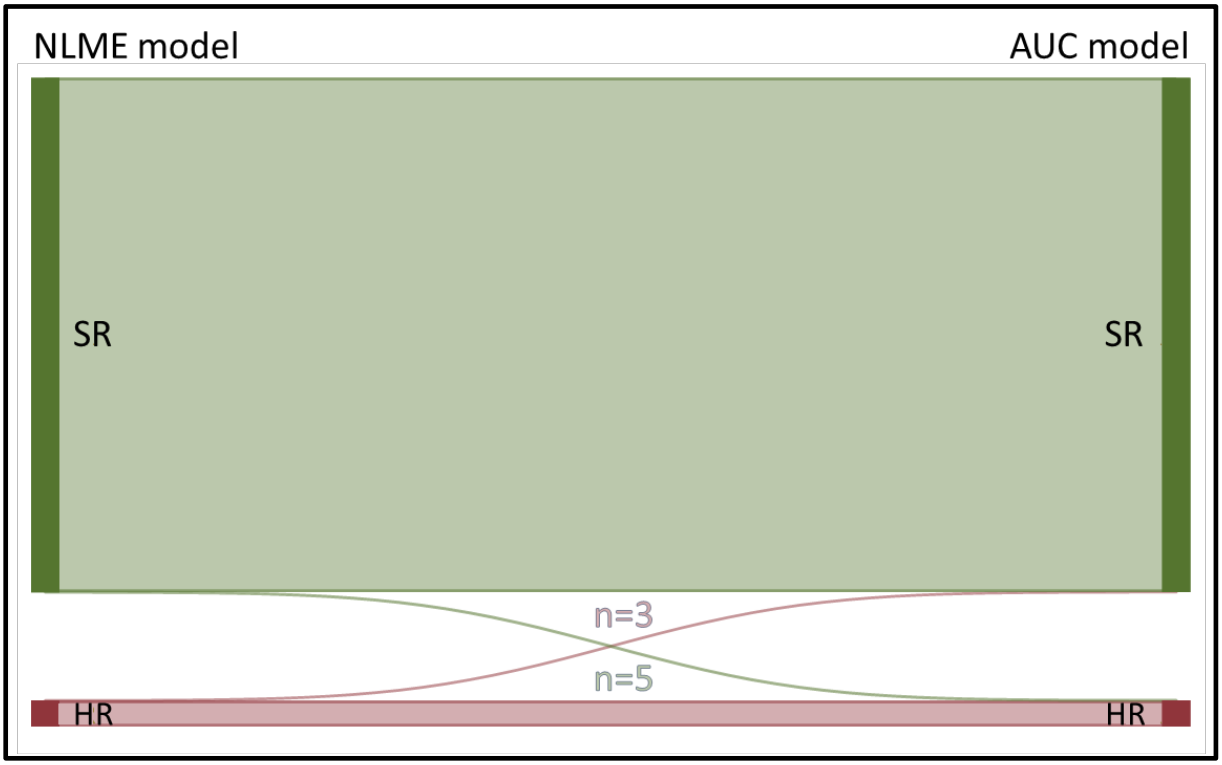

Figure 5.7. Patients moving to different risk groups identified by the NLME and AUC models. The numbers of patients shifting to the standard-risk group or high-risk group were three and five, respectively (NLME, Nonlinear Mixed-Effects; AUC, Area Under the Curve; SR, standard-risk; HR, high-risk).

Clinical outcomes of the risk groups from the NLME model demonstrated that those in the standard-risk group had higher event-free survival rate (86% vs 68%, *p*-value=0.01) and overall survival rate (91% vs 75%, *p*-value=0.008) as well as lower relapse rate (10% vs 25%, *p*-value=0.03) than patients in the high-risk group at five years. The AUC model showed similar clinical results indicating that the standard-risk group was superior to the high-risk group in event-free survival (86% vs 68%, *p*-value=0.006), relapse rate (10% vs 26%, *p*-value=0.01) and overall survival (91% vs 76%, *p*-value=0.01) at five years (Figure 5.8).

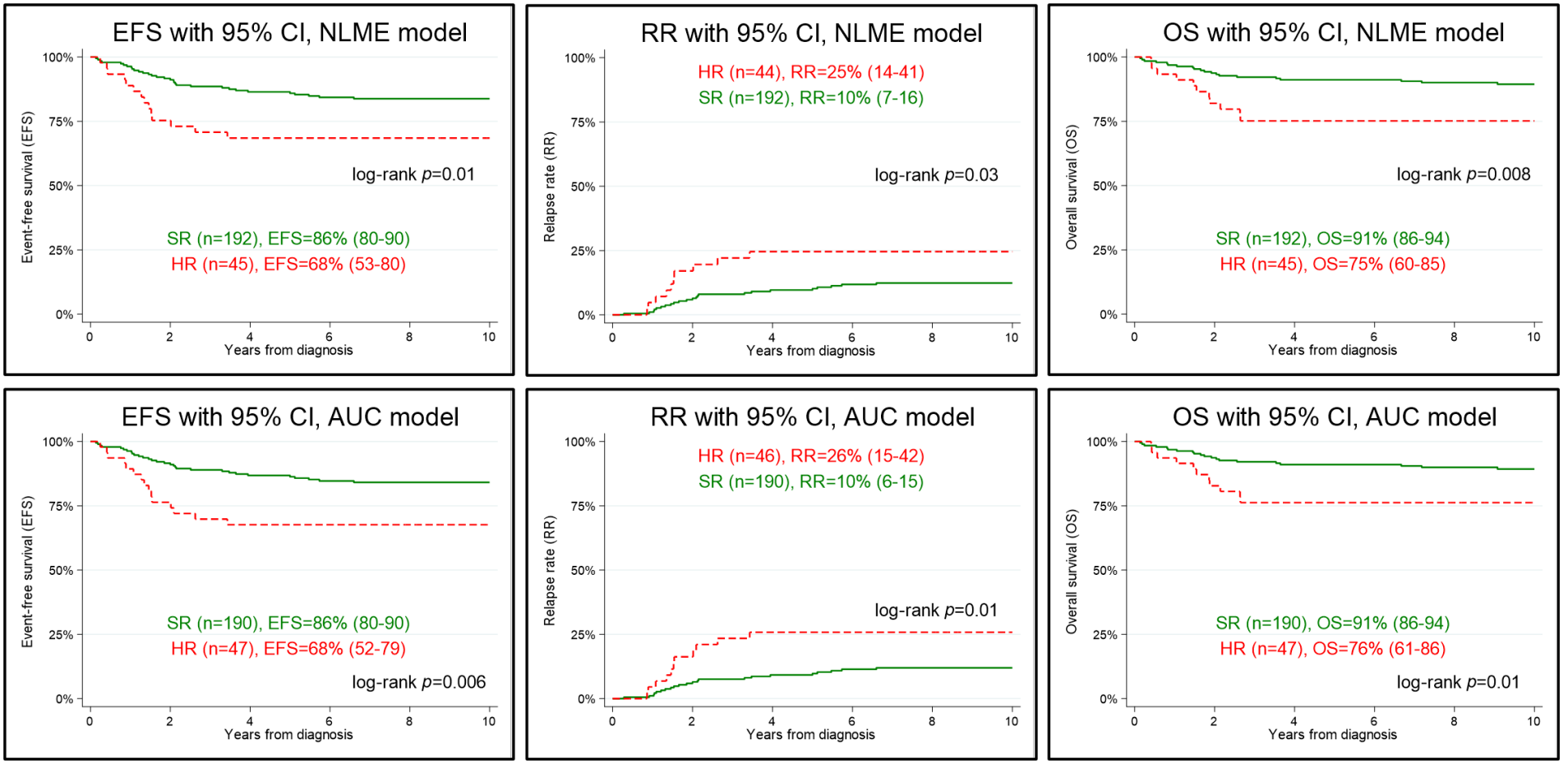

Figure 5.8. Clinical endpoints of the risk groups defined by the dynamic prognostic models in UKALL 2003. The 95% CIs were included next to the rates in the graphs (NLME, Nonlinear Mixed-Effects; AUC, Area Under the Curve; SR, standard-risk; HR, high-risk; EFS, event-free survival; RR, relapse rate; OS, overall survival; CI, confidence interval).

5.3.1.5. Clinical Characteristics and Outcomes of the Risk Groups in UKALL 2011

The same thresholds, which were used to generate the risk groups in UKALL 2003, were applied to patients in UKALL 2011. Treatment response patterns of the risk groups from the NLME and AUC models were similar in UKALL 2011, meaning that the standard-risk groups experienced quicker treatment response than the high-risk groups (Figure 5.9).

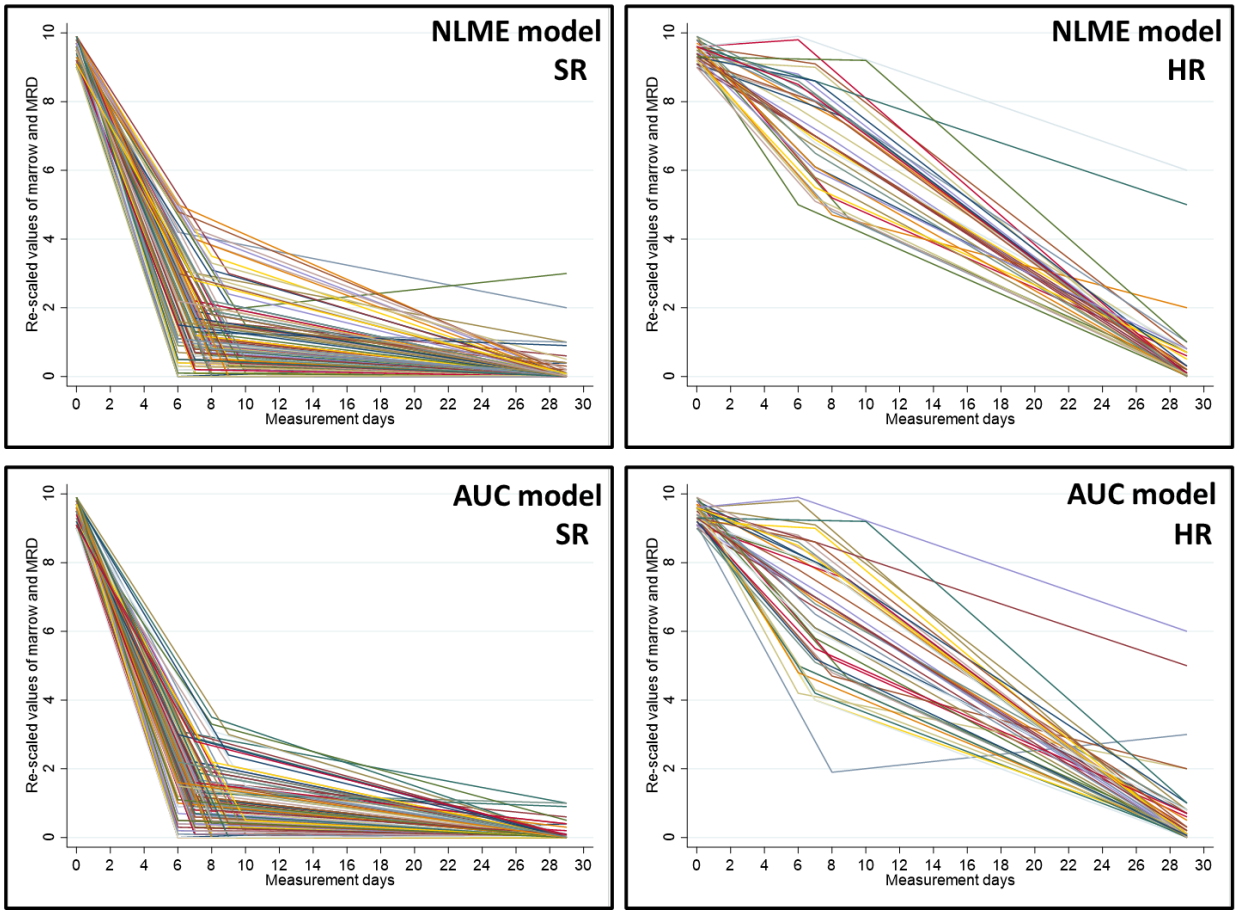

Figure 5.9. Treatment responses of the risk groups identified by the NLME and AUC models in UKALL 2011 (NLME, Nonlinear Mixed-Effects; AUC, Area Under the Curve; SR, standard-risk group; HR, high-risk group; MRD, minimal/measurable residual disease).

The standard-risk and high-risk groups had 134 (77%) vs 39 (23%) patients in the NLME model and 126 (73%) vs 47 (27%) in the AUC model, respectively. The standard-risk group was younger (median in year, 9.0 vs 12.0, *p*-value=0.03) and had a smaller proportion of intensive treatment (regimen C, 63% vs 94%, *p*-value=0.001) compared to the high-risk group defined by the NLME model. Those in the standard-risk group also demonstrated lower BM blast percentage at day 8 (9% vs 70%, *p*-value <0.0001), day 28 (1% vs 2%, *p*-value=0.02) and a smaller proportion of cases with positive MRD at the end of induction (MRD ≥0.005%, 64 % vs 95%, *p*-value <0.001) than patients in the high-risk group. There

was no statistically significant difference between the risk groups in terms of sex, WBC at diagnosis, or CNS involvement (Table 5.2.).

In the AUC model, the standard-risk group had lower BM blast percentages at day 8 (median, 8% vs 69%, *p*-value <0.0001) and day 28 (median, 1% vs 2%, *p*-value=0.02) than the high-risk group. Additionally, the proportion of cases with positive MRD at the end of induction was lower in the standard-risk group (63% vs 94%, *p*-value <0.001), and the high-risk group received more regimen C treatment (95% vs 61%, *p*-value <0.001) (Table 5.2.).

NLME model	SR (n=134, 77%)	HR (n=39, 23%)	*p*-value
Sex (male, n (%))	103 (77)	30 (77)	0.99
Age (years, median (range))	9.0 (1.0-24.0)	12.0 (1.0-24.0)	**0.03**
WBC (10^9/L, median (range))	91.9 (1.3-883.0)	165.0 (0.8-1000.0)	0.20
CNS involvement (Y, n (%))	11 (9)	2 (6)	0.73
Regimen C (Y, n (%))	80 (63)	31 (94)	**0.001**
BM blast % at d8 (median (range))	9 (0-50)	70 (46-99)	**<0.0001**
BM blast % at d28 (median (range))	1 (0-15)	2 (0-68)	**0.02**
MRD % at EOI (≥0.005, n (%))	86 (64)	37 (95)	**<0.001**

AUC model	SR (n=126, 73%)	HR (n=47, 27%)	*p*-value
Sex (male, n (%))	98 (78)	35 (74)	0.65
Age (years, median (range))	9.0 (1.0-24.0)	11.0 (1.0-24.0)	0.08
WBC (10^9/L, median (range))	87.5 (1.3-883.0)	174.9 (0.8-1000.0)	0.11
CNS involvement (Y, n (%))	11 (9)	2 (5)	0.52
Regimen C (Y, n (%))	75 (61)	36 (95)	**<0.001**
BM blast % at d8 (median (range))	8 (0-35)	69 (19-99)	**<0.0001**
BM blast % at d28 (median (range))	1 (0-15)	2 (0-68)	**0.02**
MRD % at EOI (≥0.005, n (%))	79 (63)	44 (94)	**<0.001**

Table 5.2. Clinical characteristics of patients in the risk groups identified by the dynamic models in UKALL 2011 (WBC, white blood cell count; CNS, central nervous system; BM, bone marrow; d8, day 8; d28, day 28; MRD, minimal/measurable residual disease; EOI, end of induction; Y, yes; NLME, Nonlinear Mixed-Effects; AUC, Area Under the Curve; SR, standard-risk; HR, high-risk).

In UKALL 2011, there was no patient moving from the high-risk group to the standard-risk group between the NLME and AUC models. However, 8 out of 173 patients moved from the standard-risk group defined by the NLME model to the high-risk group defined by the AUC model (Figure 5.10.). In this subset of patients (n=8), there were three females and five males, and two of them did not achieve complete remission because of high MRD levels (≥5%). In addition, one relapse and one death were observed among these patients during the follow-up time.

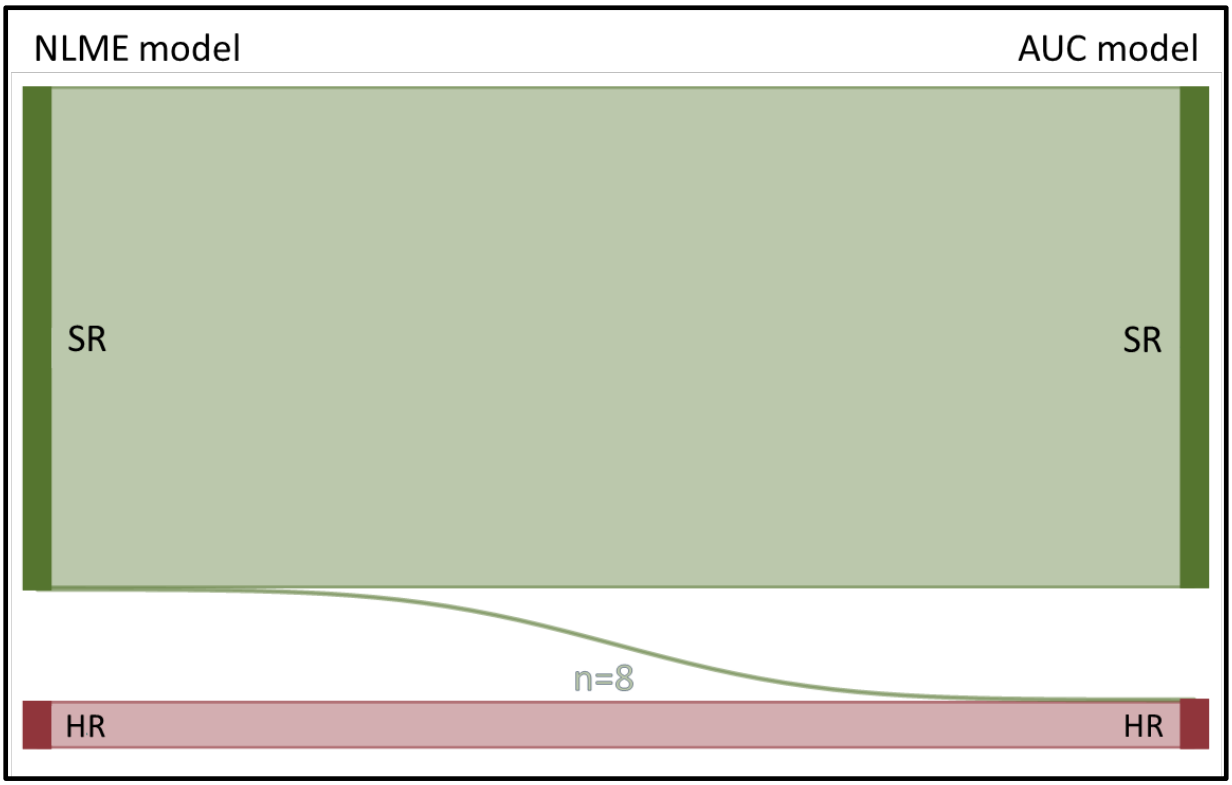

Figure 5.10. Patients shifting from the standard-risk group to the high-risk group between the dynamic models in UKALL 2011. Only eight patients ended in the AUC high-risk group between the models (NLME, Nonlinear Mixed-Effects; AUC, Area Under the Curve; SR, standard-risk; HR, high-risk).

In clinical perspective, the standard-risk groups were superior to the high-risk groups in 5-year event-free survival (84% vs 59%, *p*-value=0.0004 in the NLME model; 85% vs 59%, *p*-value=0.0003 in the AUC model), relapse rate (13% vs 36%, *p*-value=0.001 in the NLME model; 13% vs 35%, *p*-value=0.002 in the AUC model) and overall survival rate (92% vs 67%, *p*-value=0.0008 in NLME model; 92% vs 69%, *p*-value=0.0007 in the AUC model) (Figure 5.11.).

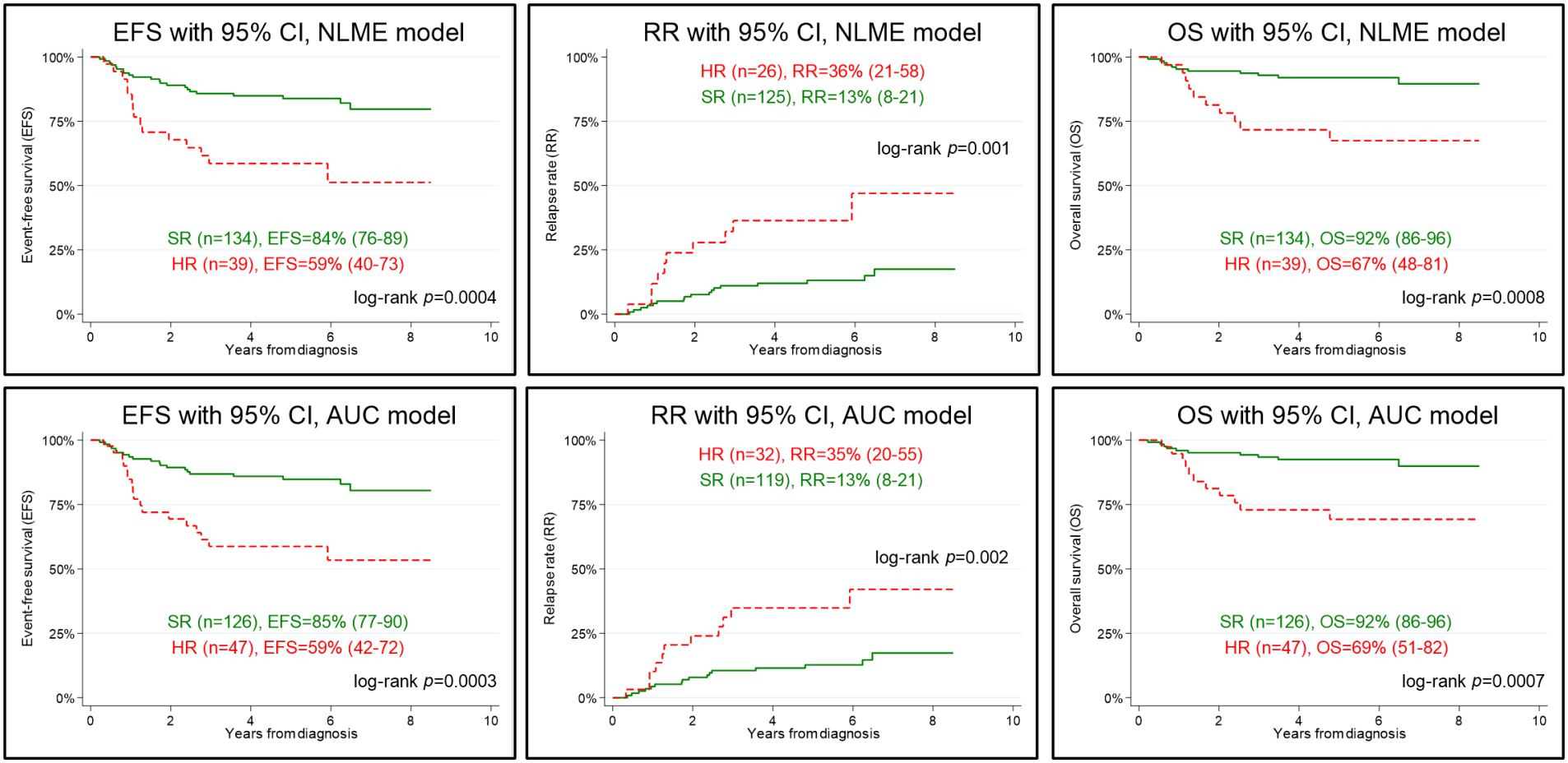

Figure 5.11. Event-free survival (EFS), relapse rate (RR) and overall survival (OS) of the risk groups in UKALL 2011. Percentages and 95% confidence intervals (CI) in the graphs correspond to 5-year event-free survival, relapse and overall survival rates. The 95% CIs were included next to the rates in the graphs (NLME, Nonlinear Mixed-Effects; AUC, Area Under the Curve; SR, standard-risk; HR, high-risk).

5.3.1.6. Comparison of the Dynamic Models with the D8BM% Model and EOI MRD

Most patients in the risk groups identified by the NLME and AUC models stayed in the same groups defined by the D8BM% model (Table 5.3.). Of 231 cases, only seven patients moved to opposite risk groups between the NLME and D8BM% models in UKALL 2003 (SR group to HR group (n=4), one female and three males, no remission failure, one high MRD level (≥5%), one relapse, no death; HR group to SR group (n=3), all males, no remission failure, one high MRD level (≥5%), one relapse, no death). On the other hand, six patients in the AUC model shifted to the opposite risk groups (SR group to HR group (n=3), all males, no remission failure, no high MRD level (≥5%), no relapse, no death; HR group to SR group (n=3), all males, no remission failure, one high MRD level (≥5%), one relapse, no death). In UKALL 2011, of 173 patients, there was no case moved from the NLME high-risk group to the D8BM% standard-risk group or from the AUC standard-risk group to the D8BM% high-risk group. However, eight patients shifted to the opposite risk groups between the dynamic models and the D8BM% model (the NLME SR group to the D8BM% HR group (n=5), three females and two males, one remission failure, one high MRD level (≥5%), one relapse, one death; the AUC HR group to the D8BM% SR group (n=3), all males, one remission failure, one high MRD level (≥5%), no relapse, no death).

UKALL 2003		
D8BM% model	NLME model	
	SR (n, %)	**HR (n, %)**
SR	182 (98)	3 (7)
HR	4 (2)	42 (93)

UKALL 2011		
D8BM% model	NLME model	
	SR (n, %)	**HR (n, %)**
SR	129 (96)	0 (0)
HR	5 (4)	39 (100)

UKALL 2003		
D8BM% model	AUC model	
	SR (n, %)	**HR (n, %)**
SR	182 (98)	3 (7)
HR	3 (2)	43 (93)

UKALL 2011		
D8BM% model	AUC model	
	SR (n, %)	**HR (n, %)**
SR	126 (100)	3 (6)
HR	0 (0)	44 (94)

Table 5.3. The number of patients moving to the opposite risk groups between the dynamic models and D8BM% model in UKALL 2003 and UKALL 2011 (NLME, Nonlinear Mixed-Effects; AUC, Area Under the Curve; SR, standard-risk; HR, high-risk).

A correlation between the risk groups, which were assigned by the dynamic models, and EOI MRD levels was found in both UKALL 2003 and UKALL 2011. The standard-risk groups in the NLME and AUC models had more patients with negative MRD values (<0.01% or

<0.005%) than the high-risk groups, whereas cases with high MRD (≥5%) were frequently observed in the high-risk groups (Table 5.4.). Similar to the interaction between the D8BM% model-assigned risk groups and MRD-intermediate risk (MRD-IR) group in terms of relapse rates (Figure 4.11.), the NLME and AUC models separated patients within the MRD-IR group in UKALL 2003 (SR vs HR, 6% vs 26%, *p*-value=0.02 in the NLME model; SR vs HR, 4% vs 30%, *p*-value=0.001 in the AUC model) and in UKALL 2011 (SR vs HR, 16% vs 39%, *p*-value=0.008 in the NLME model; SR vs HR, 15% vs 38%, *p*-value=0.009 in the AUC model) (Table 5.4.).

UKALL 2003				UKALL 2011			
EOI MRD	NLME model		5-year RR (%) (*p*-value)	EOI MRD	NLME model		5-year RR (%) (*p*-value)
	SR (n, %)	**HR (n, %)**	**SR vs HR**		**SR (n, %)**	**HR (n, %)**	**SR vs HR**
<0.01%	89 (46)	6 (13)	10 vs 0 (0.37)	**<0.005%**	48 (36)	2 (5)	9 vs 0 (0.69)
0.01%-5%	94 (49)	25 (56)	6 vs 26 (0.02)	**0.005%-5%**	79 (59)	24 (62)	16 vs 39 (0.008)
≥5%	9 (5)	14 (31)	56 vs 35 (0.42)	**≥5%**	7 (5)	13 (33)	-

UKALL 2003				UKALL 2011			
EOI MRD	AUC model		5-year RR (%) (*p*-value)	EOI MRD	AUC model		5-year RR (%) (*p*-value)
	SR (n, %)	**HR (n, %)**			**SR (n, %)**	**HR (n, %)**	**SR vs HR**
<0.01%	88 (46)	7 (15)	10 vs 0 (0.33)	**<0.005%**	47 (37)	3 (6)	9 vs 0 (0.62)
0.01%-5%	94 (49)	25 (53)	4 vs 30 (0.001)	**0.005%-5%**	74 (59)	29 (62)	15 vs 38 (0.009)
≥5%	8 (4)	15 (32)	63 vs 32 (0.22)	**≥5%**	5 (4)	15 (32)	-

Table 5.4. Correlation and interation between the risk groups defined by the dynamic models and MRD levels at the end of induction in UKALL 2003 and UKALL 2011 (NLME, Nonlinear Mixed-Effects; AUC, Area Under the Curve; SR, standard-risk; HR, high-risk; EOI MRD, minimal/measurable residual disease at the end of induction; RR, relapse rate).

Discrimination abilities of EOI MRD stratification (MRD <0.01% in UKALL 2003 and MRD <0.005% in UKALL 2011), the D8BM% model and the dynamic models were evaluated using the AUROC method in both trials. This analysis revealed that there was no statistically significant difference between any of the stratification models in UKALL 2003 (AUROC=0.54 for MRD stratification vs AUROC=0.56 for the D8BM% model vs AUROC=0.57 for the NLME model vs AUROC=0.58 for the AUC model, *p*-value=0.51) and in UKALL 2011 (AUROC=0.61 for MRD stratification vs AUROC=0.63 for the D8BM% model vs AUROC=0.62 for the NLME model vs AUROC=0.62 for the AUC model, *p*-value=0.54) (Figure 5.12.).

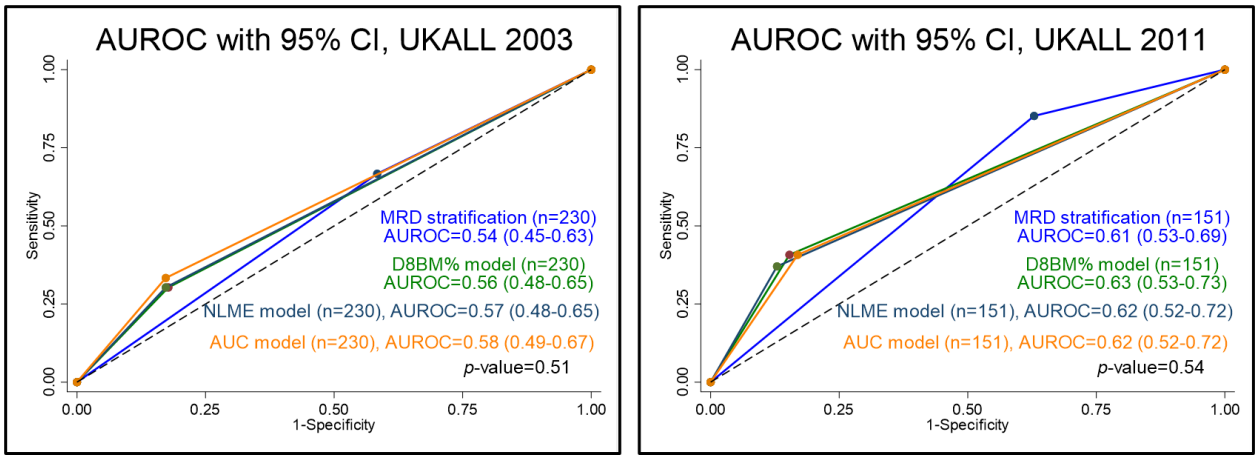

Figure 5.12. Evaluation of the discriminative abilities of the static and dynamic models in UKALL 2003 and UKALL 2011 in terms of relapse status. The 95% CIs were included next to the scores in the graphs (NLME, Nonlinear Mixed-Effects; AUC, Area Under the Curve; MRD, minimal/measurable residual disease; AUROC, Area Under the Receiver Operating Characteristic; CI, confidence interval).

5.3.2. Dynamics of EOI-EOC MRD Values in UKALL 2011

In UKALL 2011, MRD levels at week 9 and week 14 were measured in the marrow samples from patients who received regimen C, so the dynamics of MRD at multiple time points was investigated in this patient group with high-risk features (n=100). T-cell ALL patients with complete MRD values at three different time points (week 5, week 9 and week 14) were included in the process of the development of dynamic MRD model (end of induction-end of consolidation minimal/measurable residual disease model (EOI-EOC MRD model)), using the AUC method. The score of the EOI-EOC MRD model was calculated using MRD percentages, and Figure 5.13. demonstrates the distribution of the model score.

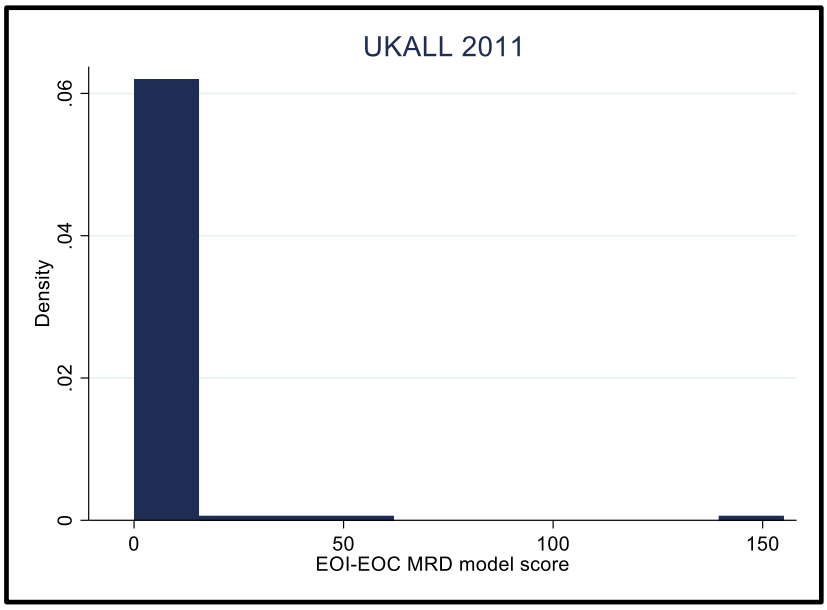

Figure 5.13. Distribution of the score of EOI-EOC MRD model in UKALL 2011 (EOI-EOC MRD, end of induction-end of consolidation minimal/measurable residual disease).

EOI-EOC MRD model score of 4 was used as the optimal threshold to categorise patients into high-risk group (score <4) and very high-risk group (score ≥4) based on the proportion of patients and relapse rates of the risk groups (Figure 5.14).

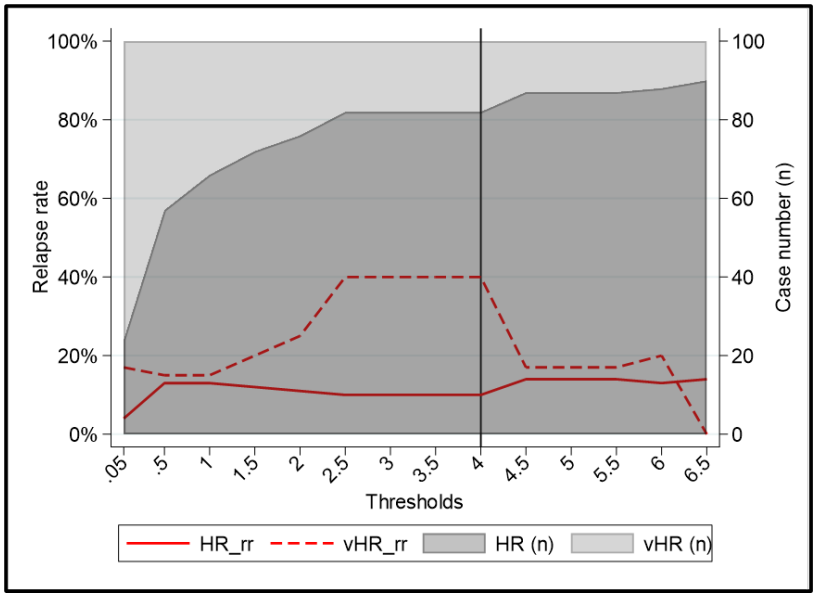

Figure 5.14. Threshold analysis of EOI-EOC MRD model score to identify optimal cut-off in UKALL 2011. The thresholds from the model scores are shown on the x-axis. The y-axis on the left side indicates relapse rates, and the y-axis on the right side shows the number of patients. Solid and dashed red lines represent relapse rates of the high-risk group and very high-risk group, respectively. The line where grey colours (dark grey and bright grey) cross in the graph demonstrates the number/proportion of patients in the high-risk group (HR, high-risk; vHR, very high-risk; rr, relapse rate; n, patient number).

As expected, the high-risk group (n=82, 82%) had more rapid MRD response (sharper declines) than the very high-risk group (n=18, 18%) (Figure 5.15; Table 5.5.). Interestingly, patients in the very high-risk group experienced higher BM blast percentage at day 8 compared to those in the high-risk group (median, 11% vs 46%, *p*-value=0.02) (Table 5.5.).

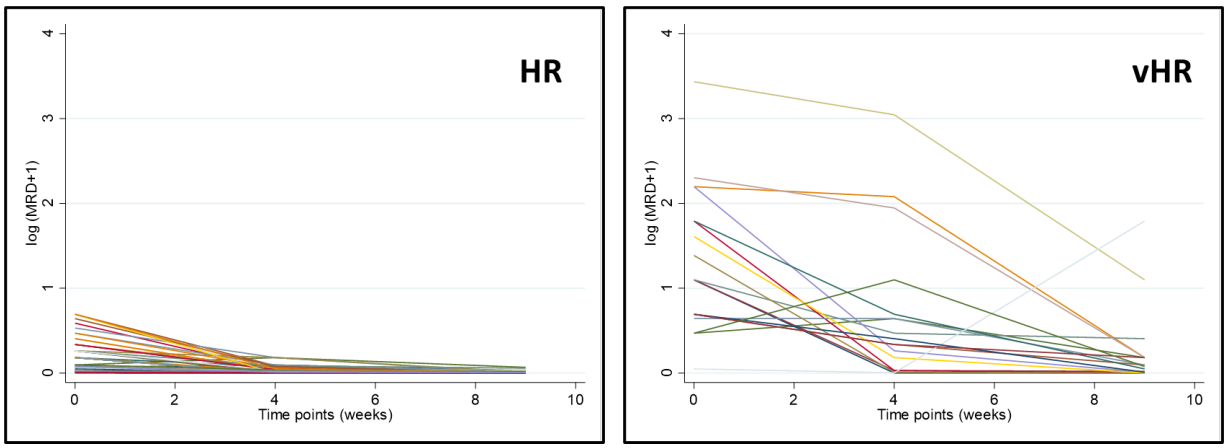

Figure 5.15. MRD responses in log scale in the risk groups defined by the EOI-EOC MRD model in UKALL 2011. The reference time point is considered as week 5 (week 0 in the graphs), and other time points are adjusted to this reference time point (week 9 and week 14 are week 4 and week 9 in the graphs, respectively) (HR, high-risk; vHR, very high-risk; MRD, minimal/measurable residual disease; EOI-EOC MRD, end of induction-end of consolidation minimal/measurable residual disease).

EOI-EOC MRD model	HR (n=82, 82%)	vHR (n=18, 18%)	*p*-value
Sex (male, n (%))	68 (83)	13 (72)	0.29
Age (years, median (range))	9.5 (1.0-24.0)	8.0 (1.0-20.0)	0.56
WBC (10^9/L, median (range))	96.0 (7.0-1000.0)	73.0 (1.2-624.0)	0.19
BM blast % at d8 (median (range))	11 (0-98)	46 (1-94)	**0.02**
BM blast % at d28 (median (range))	1 (0-15)	2 (0-6)	0.07
MRD % at week 5 (median (range))	0.07 (0.02-1.0)	2 (0.05-30)	**<0.0001**
MRD % at week 9 (median (range))	0.01 (0-0.2)	0.45 (0-20)	**<0.0001**
MRD % at week 14 (median (range))	0.001 (0-0.07)	0.09 (0-5)	**<0.0001**

Table 5.5. Clinical characteristics of patients in the high-risk and very high-risk groups in UKALL 2011 (WBC, white blood cell count; BM, bone marrow; MRD, minimal/measurable residual disease; EOI-EOC MRD, end of induction-end of consolidation minimal/measurable residual disease; HR, high-risk; vHR, very high-risk; d, day).

Clinical outcomes of the risk groups demonstrated that the very high-risk group was inferior to the high-risk group in event-free survival (59% vs 90%, *p*-value=0.014), relapse rate (40% vs 10%, *p*-value=0.04) and overall survival (70% vs 97%, *p*-value=0.0007) (Figure 5.16).

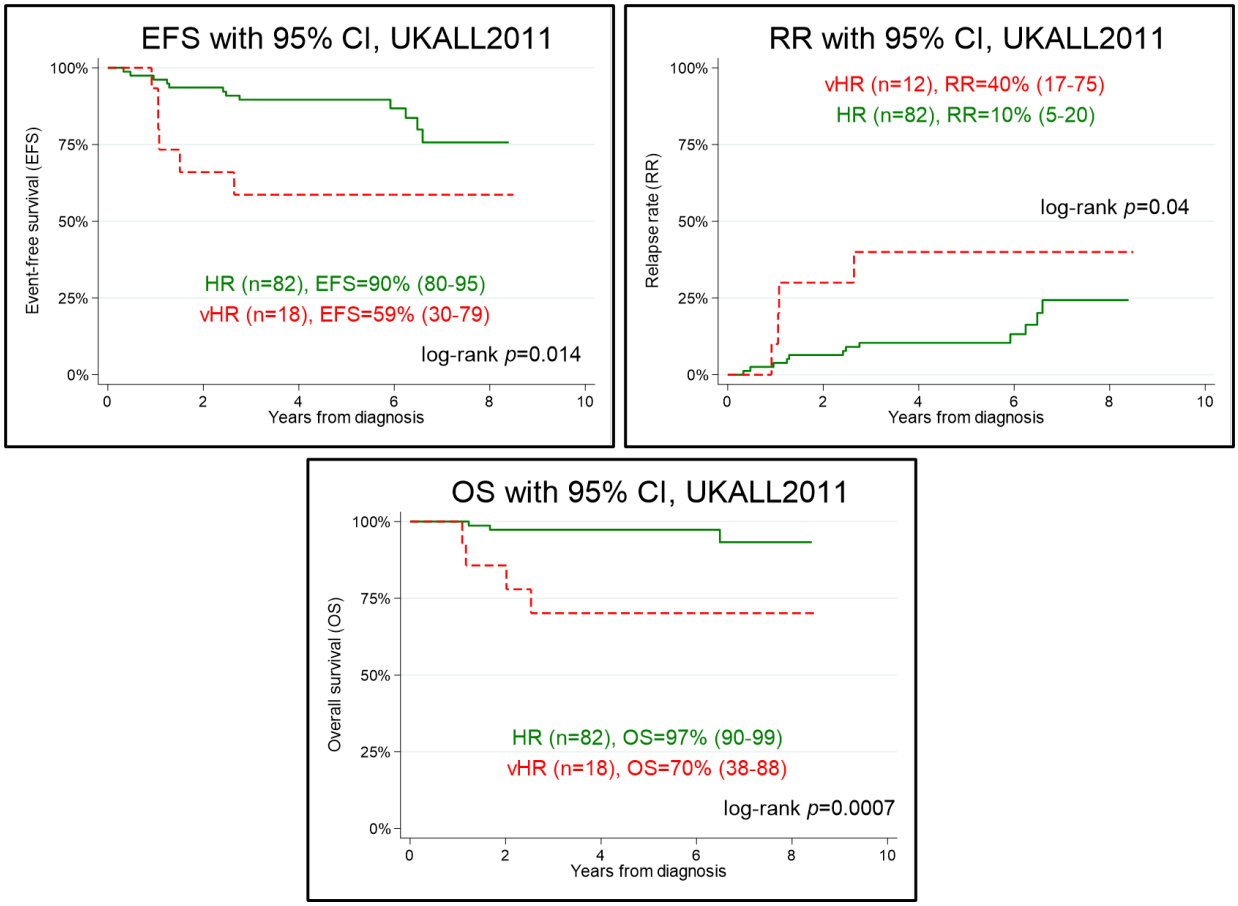

Figure 5.16. Clinical outcomes of the risk groups defined by the EOI-EOC MRD model in UKALL2011. The 95% CIs were included next to the rates in the graphs (HR, high-risk; vHR, very high-risk; EFS, event-free survival; RR, relapse rate; OS, overall survival; CI, confidence interval).

In comparison to the high-risk group, the very high-risk group had higher proportion of patients in the D8BM% high-risk group (53% vs 23%), in the NLME high-risk group (47% vs 20%) and in the AUC high-risk group (53% vs 22%). In addition, the very high-risk group captured most cases with MRD values of ≥0.01% at the end of consolidation (78% vs 13%) (Table 5.6.).

UKALL 2011			
Models/MRD groups		EOI-EOC MRD model	
	Subgroups	**HR (n, %)**	**vHR (n, %)**
D8BM% model	**SR**	56 (77)	8 (47)
	HR	17 (23)	9 (53)
EOC MRD	**<0.01%**	71 (87)	4 (22)
	≥0.01%	11 (13)	14 (78)
NLME model	**SR**	48 (80)	8 (53)
	HR	12 (20)	7 (47)
AUC model	**SR**	47 (78)	7 (47)
	HR	13 (22)	8 (53)

Table 5.6. Distribution of patients in the risk groups defined by different prognostic models and MRD stratification at the end of consolidation in UKALL 2011 (NLME, Nonlinear Mixed-Effects; AUC, Area Under the Curve; EOI-EOC MRD, end of induction-end of consolidation minimal/measurable residual disease; MRD, minimal/measurable residual disease; EOC, end of consolidation).

AUROC analysis showed that the EOI-EOC MRD model based on post-induction MRD values distinguished the relapsed and non-relapsed patients with 57% (AUROC=0.57 (0.46-0.69)). However, there was no statistically significant difference between MRD stratification with the threshold of 0.01% at the end of consolidation and the EOI-EOC MRD model (0.56 vs 0.57, p=0.72) (Figure 5.17.)

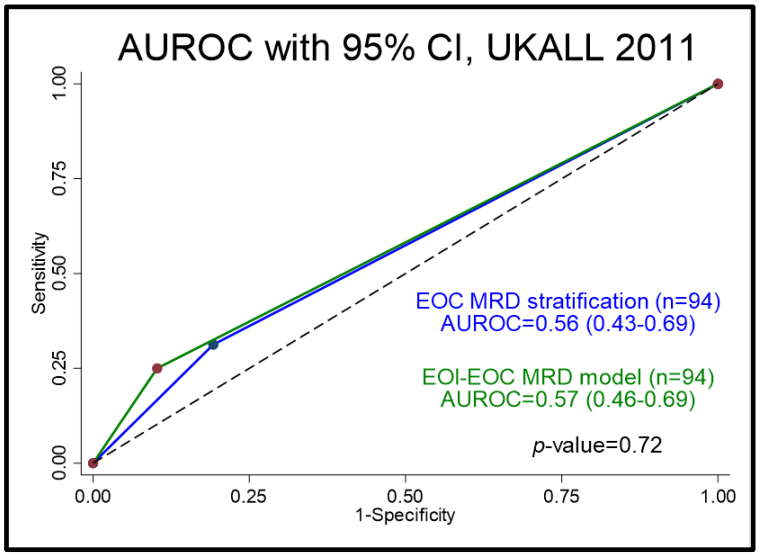

Figure 5.17. Discriminative abilities of the EOI-EOC MRD model and MRD stratification with the cut-off of 0.01% at the end of consolidation. The 95% CIs were included next to the scores in the graphs (EOI-EOC MRD, end of induction-end of consolidation minimal/measurable residual disease; MRD, minimal/measurable residual disease; EOC, end of consolidation; AUROC, Area Under the Receiver Operating Characteristic).

5.4. Discussion

Prognostic models are used to stratify patients into different treatment groups in oncology (Craddock *et al.*, 2022), and static prognostic models were developed for paediatric patients with acute lymphoblastic leukaemia (ALL) (DelRocco *et al.*, 2024; Enshaei *et al.*, 2020). To date, there are no robust static or dynamic models validated in T-cell ALL cohorts, other than single-time point MRD stratification. Chapter 4 has showed that early treatment response at a single-time point (day 8) can categorise patients into risk groups and add prognostic value to MRD at the end of induction (EOI). In this chapter, early treatment responses (marrow and/or MRD responses) were used to build dynamic prognostic models using Nonlinear Mixed-Effects (NLME) regression and Area Under the Curve (AUC). To our knowledge, dynamic models developed using the NLME and AUC methods were first investigated in this thesis for T-cell ALL patients.

In UKALL 2003 and UKALL 2011, the NLME and AUC models with optimal thresholds separated T-cell ALL patients into the standard-risk group and high-risk group, depending on the trend of treatment response during the induction phase, and the risk groups showed distinct clinical outcomes (Figure 5.6., Figure 5.8., Figure 5.9., Figure 5.11). These two models identified similar number of cases within each risk group, and only a few patients moved to the opposite risk groups between the NLME and AUC models (Figure 5.7., Figure 5.10.). Despite the fact that the induction treatment protocols and MRD stratifications were revised in UKALL 2011 (Goulden, 2013), the size of the risk groups only slightly changed compared to those in UKALL 2003 (Table 5.1., Table 5.2.). Additionally, the availability of a marrow response variable could be another reason for the slight difference as BM blast percentage at day 15 was not measured in UKALL 2011; therefore, the dynamic models were validated using the BM blast percentages at diagnosis and day 8, and MRD at the end of induction in UKALL 2011. As a result, despite different treatment protocols and the lack of marrow response data, the dynamic models were successfully applied to treatment response variables in both trials. The NLME model can be more robust for complete data, whereas incomplete data can be more suitable for the AUC model because the NLME model is more sensitive to small changes than the AUC model, as seen in Figure 5.3. showing the distributions of the trend scores in the trials. On the other hand, when comparing patients in the risk groups defined by the dynamic models and the D8BM% model, most patients stayed in the same risk groups in UKALL 2003 and UKALL 2011 (Table 5.3.), indicating that the

D8BM% model with a single-time point measurement was strong enough to identify the majority of patients in the risk groups at very early stage of treatment.

In UKALL 2011, the dynamics of MRD values, which were measured in the bone marrow samples taken from T-cell ALL patients in the regimen C treatment group, were examined using the AUC method. EOI-EOC MRD model consisted of MRD values at different time points (week 5, week 9 and week 14). The EOI-EOC MRD model classified patients into the high-risk and very high-risk groups according to the trend of MRD response (Figure 5.14.), and the high-risk group was superior to the very high-risk group in clinical outcomes (Figure 5.16.). However, when comparing the EOI-EOC MRD model with a single-time point MRD at the end of consolidation, there was no statistically significant difference in terms of ability to discriminate relapse status (Figure 5.17). Similar to this result, Popov and colleagues recently published a paper demonstrating that a combination result of MRD values at EOI and EOC provided very limited additional prognostic information in B-cell ALL (Popov *et al.*, 2023b). Hence, MRD measured at a single-time point can be sufficient to identify the risk groups in T-cell ALL. An interesting observation was also seen in comparison analysis of clinical characteristics. Patients in the very high-risk group had higher BM blast percentage at day 8 than those in the high-risk group (Table 5.5.). This observation indicated that early marrow response at day 8 was associated with the trend of MRD responses measured until the end of consolidation. This is further evidence that early treatment response is a significant prognostic indicator at very early stage of the therapy in T-cell ALL. These findings can be explained by the fact that MRD levels at EOI or EOC are more frequently undetectable compared to marrow levels measured morphologically at early time points. Popov and colleagues, investigating MRD responses at day 15 and day 36 in B-cell ALL (n=507), reported that early MRD response at day 15 had significant prognostic value in B-cell ALL (Popov *et al.*, 2023a). On the other hand, although a positive correlation was observed between MRD dynamics at later timepoints and BM blast percentage at day 8, EOC MRD with the threshold of 0.01% identified higher proportion of patients in HR and vHR, compared to the D8BM% model (Table 5.6.).

The main advantage of dynamic models over static models is that dynamic models account for changes over time, using more information collected at multiple time points. Also, these models are not based on fixed values, so they provide updated information about data (Jenkins *et al.*, 2018). In clinical settings, collecting data at multiple time points can be challenging, and missing time points can affect the outputs of dynamic models; therefore,

using dynamic models in the real world can be limited and complicated. The dynamic models developed in this chapter required response data collected at minimum of two different time points, so more time and effort are needed for the dynamic models, unlike static models such as the D8BM% model or EOI MRD stratification.

The main limitation of developing/validating the dynamic models was about different laboratory techniques, which were used for measurements of early treatment response variables. Early marrow responses were measured by morphological assessment (light microscopy), meaning a subjective evaluation of the bone marrow samples, while PCR was used to measure MRD levels. Although the dynamic models have been validated in UKALL 2011, the models should be evaluated/validated in different T-cell ALL cohorts containing response values generated by the same laboratory technique (e.g. flow cytometry) because treatment response values detected by different laboratory methods have been used as inputs for the dynamic models in this chapter.

In conclusion, the findings in this chapter showed that the NLME and AUC methods were effectively used for the repeated measurements (treatment responses), and the dynamic models, which were developed using these methods, stratified T-cell ALL patients into the treatment risk groups in UKALL 2003 and UKALL 2011. However, the advantage of including trend/change of the treatment responses over time in the dynamic models could not add significant discriminative or prognostic value to the static model (D8BM% model) or EOC MRD stratification. Therefore, static models/stratifications that require less time and effort are sufficient to stratify T-cell ALL patients for treatment allocation. On the other hand, the reduction rate or trend of treatment response variables during induction phase were prognostic, highlighting again that early treatment response had prognostic significance in T-cell ALL.

Chapter 6. Investigation of T-cell Acute Lymphoblastic Leukaemia Genetics Using Bioinformatics

6.1. Introduction

In recent years, gene expression profiling techniques have provided useful information about global patterns of gene activity. There are many techniques (e.g. microarrays, Quantitative Reverse Transcription-Polymerase Chain Reaction (qRT-PCR), RNA-sequencing (RNA-seq), etc.) for the determination of expression profiles (Richards, 2005). In T-cell ALL, biological subtypes and genetic alterations have been identified by these methods. For example, Liu and colleagues identified eight different subtypes (TAL1, TAL2, TLX1, TLX3, HOXA, LMO1/LMO2, LMO2/LYL1, and NKX2-1) based on RNA-seq analysis and investigated driver genes and genetic alterations in T-cell ALL (Liu *et al.*, 2017b). Another study reported that 15 T-cell ALL subtypes (TAL1 DP-like, TAL1 αβ-like, NKX2-1, NKX2-5, LMO2 γδ-like, HOXA9 TCR, TLX1, TLX3, TME-enriched, STAG2/LMO2, SPI1, MLLT10, ETP-like, BLC11B, KMT2A) were identified using Whole Transcriptome Sequencing (WTS), taking into account both coding and noncoding RNA regions (Pölönen *et al.*, 2024). Although these methods have improved our knowledge of T-cell ALL genetics and biology, the availability of robust statistical techniques for the analyses and computational facilities (storage and processing of data) are the main limitations of using these technologies (Jain, 2011).

TALLSorts is a bioinformatic tool that uses bulk RNA-seq data to classify T-cell ALL subtypes. Key genes, which are selected by logistic regression with one vs rest strategy, are used for the identification of the subtypes in this algorithm. The TALLSorts algorithm identifies eight T-cell ALL subtypes based on deregulation of *BCL11B* (BCL11B subtype), presence of *KMT2A* or *MLLT10* gene fusions (HOXA-KMT2A or HOXA-MLLT10 subtypes), overexpression of *NKX2* (NKX2 subtype), deregulation of *TAL1/2* or *LMO1/2* (TAL/LMO subtype), overexpression of *TLX1* (TLX1 subtype) or *TLX3* (TLX3 subtype), and other T-cell ALL subtypes (Diverse subtype). This classifier accepts count matrix from RNA-seq data aligned to the hg38 reference genome (Gu *et al.*, 2023).

DESeq2 is an R package used to analyse bulk RNA-seq data for the identification of differentially expressed genes between different groups (Love, Huber and Anders, 2014). The DESeq2 vignette clearly explains its workflow from inputting RNA-seq data to generating/visualising results (Love, Anders and Huber, 2024). This algorithm performs estimation of size factors (to normalise sequencing depth and RNA composition), estimation of gene-wise dispersion (to generate more accurate estimates of variability (spread) for each gene within groups), negative binomial generalised linear model fitting (to fit a model curve to dispersion estimates for each gene) and the Wald test (to statistically compare two groups for differentially expressed genes) (Figure 6.1.).

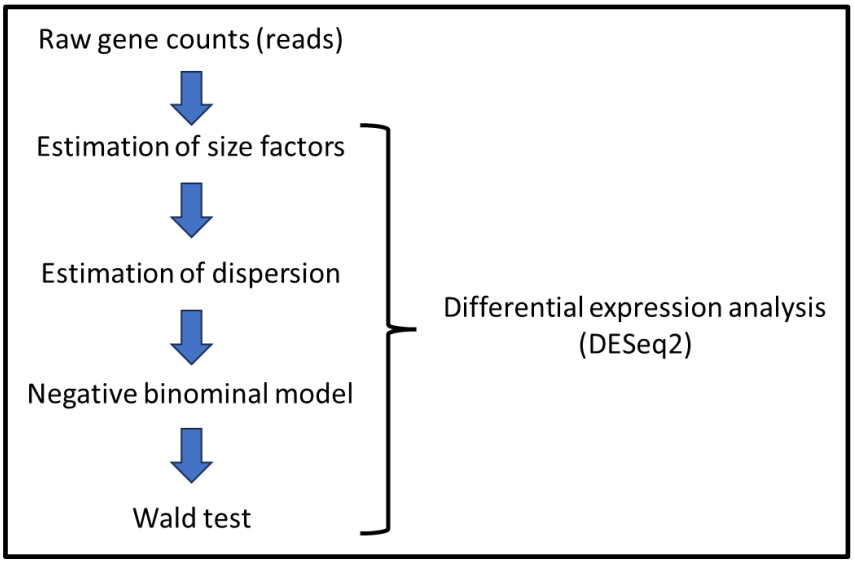

Figure 6.1. Main steps of differential expression analysis using the DESeq2 algorithm.

Clustering is a technique to categorise data points based on similarity within clusters and dissimilarity between clusters. Supervised clustering, for example DESeq2, uses labelled data to train the algorithm for the separation of the clusters, while unsupervised clustering categorises the clusters using unlabelled data, meaning that there is no data training for classification. Although there are many unsupervised methods in the literature, ConsensusClusterPlus algorithm has become popular in cancer research, specifically in cancer genomics (Thongprayoon *et al.*, 2021; Ren *et al.*, 2022; Hu *et al.*, 2022), and it is an effective technique, which utilises different algorithms along with resampling and distance matrix methods (Monti *et al.*, 2003). The ConsensusClusterPlus starts with subsampling a proportion of samples/features. Next, each subsample is partitioned into user-specified k groups by the clustering algorithm. This step is repeated for a user-defined number. Later, the algorithm calculates consensus values (the proportion of clustering runs where items are together) and classifies samples/features into k clusters on the basis of "1-consensus value". The final k clusters are the consensus clusters, meaning the final output of the algorithm (Figure 6.2.).

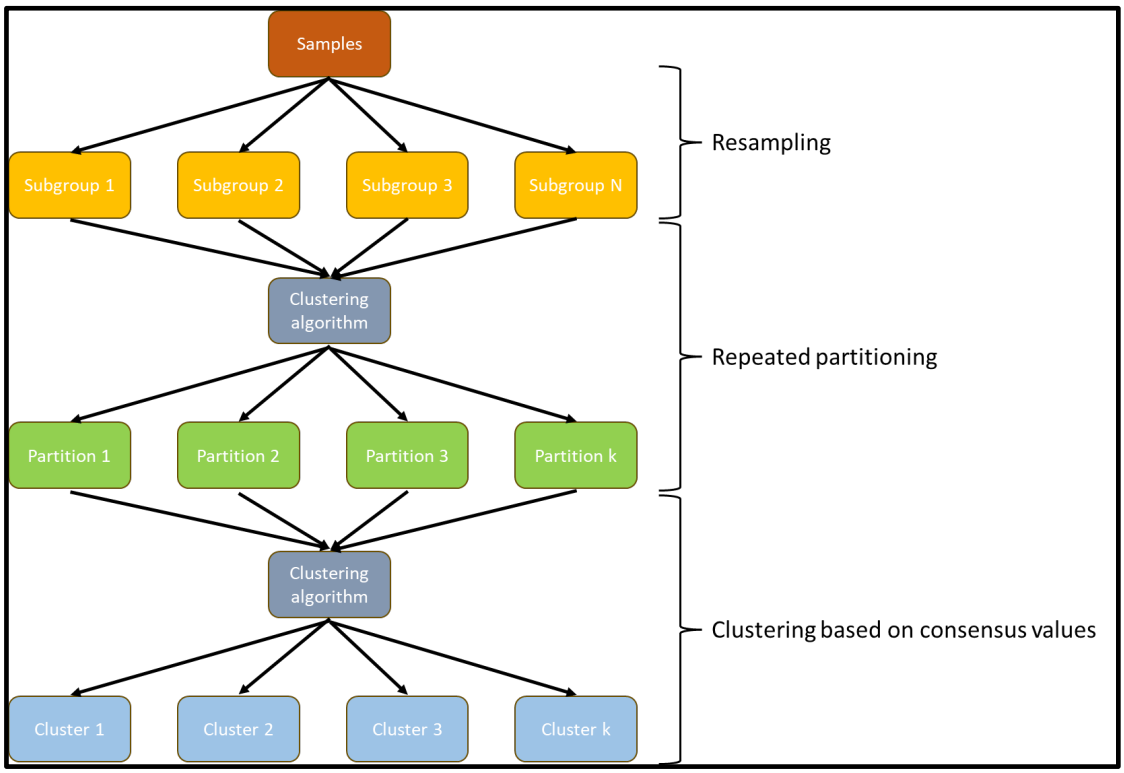

Figure 6.2. Overview of an unsupervised clustering method using the ConsensusClusterPlus algorithm.

The aims of Chapter 6 are to evaluate predictive performance of the genetic classifier (TALLSorts), to identify genes related to T-cell ALL, and to discover relevant subtypes in T-cell ALL.

The objectives of this chapter are to:

1) Externally evaluate the TALLSorts algorithm to check its ability to identify T-cell ALL subtypes using RNA-seq data from the UKALL trials and check distribution of the predicted subtypes by the classifier in the risk groups defined by the D8BM% model.

2) Identify differentially expressed genes between the standard-risk and high-risk groups identified by the D8BM% model using the differential expression analysis (DESeq2) and find overlapping genes between the genetic classifier and differential expression analysis.

3) Determine potential genetic clusters (subgroups) using the unsupervised clustering analysis (ConsensusClusterPlus) and examine the proportion of the risk groups defined by the D8BM% model and clinical characteristics of patients within the clusters.

6.2. Data and Methods

6.2.1. Processing FASTQ Files

Bulk RNA-seq data (paired-end FASTQ files) from 126 T-cell ALL patients were used in this chapter. They were sequenced in 2018 (n=29) and 2024 (n=97). Patients were treated with either UKALL 2003 treatment protocol (n=87) or UKALL 2011 treatment protocol (n=39). Quality control was performed on FASTQ files using "FastQC" (version 0.12.1) and "multiqc" (version 1.0.dev0). Adapters and poor-quality reads were trimmed with "bbmap" (version 39.06) and "homer" (version 4.11.1). "Salmon" (version 1.9.0) was used for alignment to the hg38 reference genome (GRCh38.108.idx) for the estimated read counts from trimmed FASTQ files. Outputs of Salmon software (quant files) were imported with the "tximport" package (version 1.30.0) to R for the purpose of creating a count matrix for the TALLSorts, DESeq2, and ConsensusClusterPlus algorithms (Figure 6.3.).

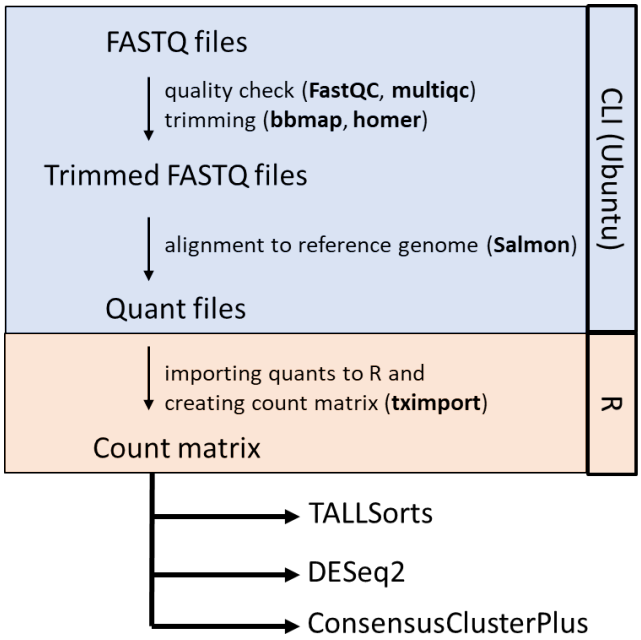

Figure 6.3. Processing RNA-seq data (FASTQ files) for further investigations using the TALLSorts, DESeq2 and ConsensusClusterPlus (CLI, Command Line Interface; R, R programming language).

6.2.2. Genetic Classifier (TALLSorts)

A count matrix was generated from the quant files of 126 samples to perform TALLSorts analysis for the classification of T-cell ALL subtypes. Gene fusions detected by the laboratory methods (FISH, MLPA, molecular/cytogenetics) were re-categorised into relevant subtypes defined by the TALLSorts algorithm: BCL11B subtype (*BCL11B* fusions), HOXA_KMT2A subtype (*KMT2A* fusions), HOXA_MLLT10 subtype (*MLLT10* fusions),

NKX2 subtype (*NKX2* fusions), TAL/LMO subtype (*TAL1 / TAL2 / LMO1 / LMO2 / LMO3 / SIL::TAL1* fusions), TLX1 subtype (*TLX1* fusions), and TLX3 subtype (*TLX3* fusions). Each sample was classified into the subtypes with the highest score, as estimated by the TALLSorts, among the eight possibilities, only if the score was >0.50; otherwise, it was labelled "unclassified subtype".

6.2.3. Differential Expression Analysis (DESeq2)

High-throughput technologies provide novel insights to better understand biological variation underlying different clinical conditions. Biological variations from gene/protein changes are identified in outputs from the sequencing technologies. Sometimes, the source of variations can be seen due to technical differences, such as differing experiment time, location, different instrumentation, etc. Such differences can change the data produced by the technologies and are termed "batch effect". Principial Component Analysis (PCA) is one of the unsupervised methods used to explore batch effect(s) via PCA plots. The technique performs dimensionality reduction to make the interpretation of high-dimensional data more accessible and efficient. It clusters samples with similar expression profiles together. PC1 has the most variation, while PC2 represents the second most variation in the data. Separate clusters in a PCA plot demonstrate a presence of batch effect(s) caused by factor(s) of interest (Goh, Wang and Wong, 2017).

Genes with low counts (total normalised counts in one sample <10) were removed from the analyses. After filtering out the genes without low counts, batch effect was checked by the PCA analysis with the plotPCA command from the "DESeq2" package. Figure 6.4. proved the existence of a batch effect caused by sequenced year of the RNA-seq data; therefore, the batch effect was removed with the "limma" package (version 3.58.1) and corrected by adding it to DESeq2 formula before the identification of differentially expressed genes between the risk groups.

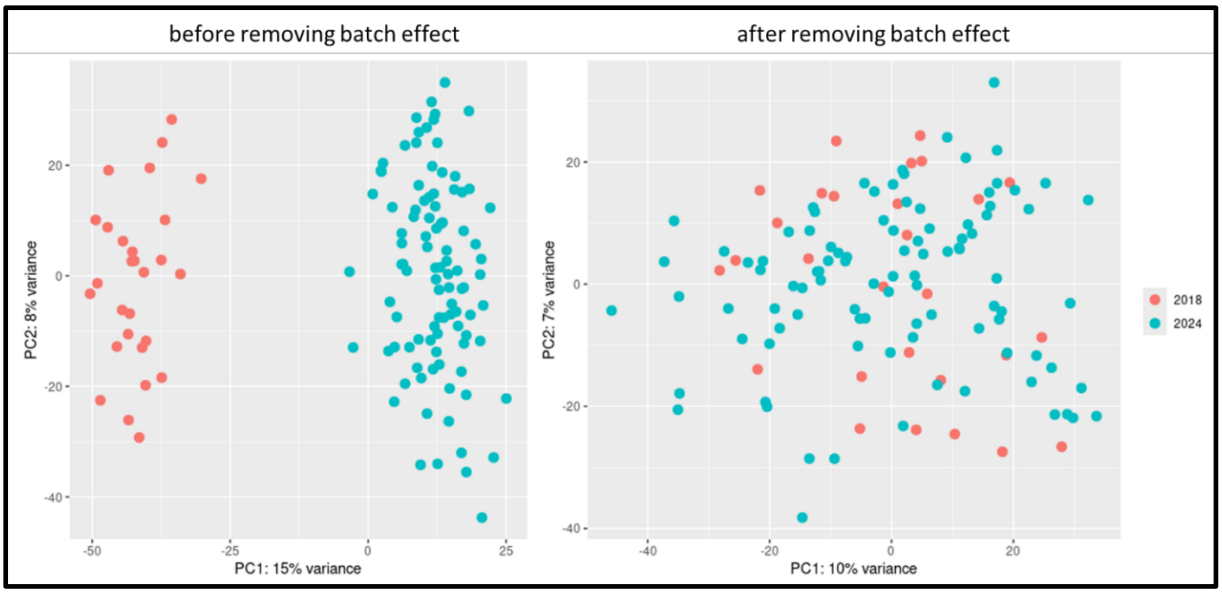

Figure 6.4. PCA plots for the identification of the batch effect in RNA-seq data. Each dot represents samples sequenced in 2018 (red) and 2024 (blue) (PC1, the first principal component; PC2, the second principal component).

The count matrix file was utilised to identify differentially expressed genes, using "DESeq2" (version 1.42.1). Comparison groups were defined as the standard-risk (n=92) and high-risk (n=33) identified by the D8BM% model for differential expression analysis. The fitted DESeq2 curve to estimates of dispersion (variability) for each gene and distributional assumption of the DESeq2 model were checked in Figure 6.5., indicating that the assumption was met (the mean of normalised counts was not equal to the variance of normalised counts in the negative binomial distribution).

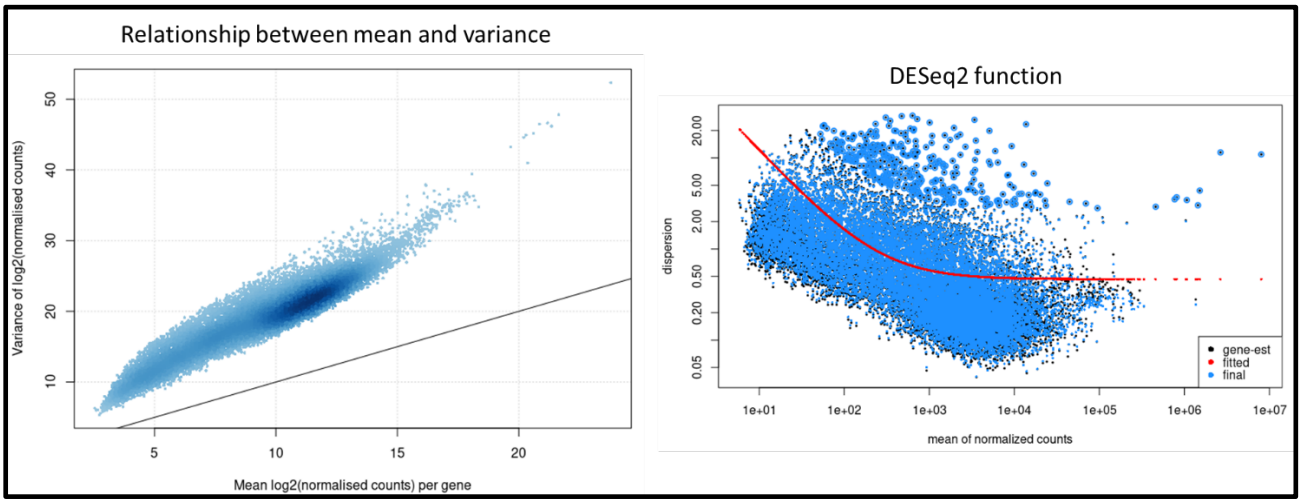

Figure 6.5. Assumption of the DESeq2 model (negative binomial distribution) depending on mean and variance, and its function. Each blue-colour point represents the expression level of different genes.

Statistically significant genes were filtered out using an adjusted *p*-value (adjp-value <0.05) and log2 fold change (log2FC ≤-2 and log2FC ≥2).

The "org.Hs.eg.db" package (version 3.18.0) was used for annotation of DESeq2 results with ENSEMBL identifier. To discover biologically relevant genes, gene set enrichment analysis was performed for commonalities in the differentially expressed genes, using "clusterProfiler" (version 4.10.1), and Gene Ontology (GO) terms was used to define biological functions of the genes.

6.2.4. Unsupervised Clustering Analysis (ConsensusClusterPlus)

The count matrix file containing information for 126 samples was used for the unsupervised clustering analysis (ConsensusClusterPlus). After removing batch effect/genes with low counts and transforming/normalising the counts, Median Absolute Deviation (MAD) was calculated for each gene across the samples, using the "matrixStats" package (version 1.3.0.). The top 2000 genes with the highest MAD scores were used for the feature selection. ConsensusClusterPlus, using the resampling method, partition around medoids (K-medoids) and Spearman correlation, was applied to the transformed/normalised counts for the identification of clusters based on similarity/dissimilarity of the samples. The following clustering settings were used in the ConsensusClusterPlus: maximum number of clusters=8; number of resampling=50; proportion of items in each iteration=0.8; clustering algorithm=partition around medoids (K-medoids); distance matrix=Spearman correlation). Later, distribution of the risk groups defined by the D8BM% model and the subtypes predicted by TALLSorts were checked within the clusters.

6.3. Results

6.3.1. Comparing Tested and Non-tested Cohorts for RNA-seq Analyses

The samples sequenced in 2018 and 2024 (tested cohort; n=126) were compared to total cohort (tested cohort and non-tested cohort; n=707) in terms of clinical characteristics and outcomes of patients from UKALL 2003 and UKALL 2011. There was no statistically significant difference in sex, white blood cell count (WBC) at diagnosis, central nervous system (CNS) involvement, marrow response at day 8/28, positivity of minimal/measurable residual disease (MRD) at the end of induction (EOI), and clinical outcomes, such as event-free survival, relapse rate and overall survival. However, the tested cohort was younger (8.8 vs 10.0, *p*-value=0.005) and had a higher level of bone marrow (BM) at diagnosis (91% vs 90%, *p*-value=0.03) as well as higher complete remission rate (98% vs 94%, *p*-value=0.04). The higher complete remission rate may be explained by the younger age of this cohort, as younger patients often tolerate treatment toxicity better. Overall, the tested cohort was representative of the total cohort containing all patients from UKALL 2003 and UKALL 2011 (Table 6.1.).

UKALL 2003 & UKALL 2011	Tested cohort	Tested + non-tested cohort	*p*-value
Total number (%)	126 (15)	707 (85)	-
Sex (male, n (%))	97 (77)	530 (75)	0.63
Age (years, median (range))	8.8 (1.0-21.9)	10.0 (1.0-24.3)	**0.005**
WBC (10^9/L, median (range))	108.9 (1.0-881.0)	85.2 (0.4-1800.0)	0.16
CNS involvement (Y, n (%))	9 (7)	37 (6)	0.43
BM blast % at diagnosis (median (range))	91 (14-100)	90 (4-100)	**0.03**
BM blast % at day 8 (median (range))	14 (0-96)	10 (0-99)	0.09
BM blast % at day 28 (median (range))	1 (0-70)	1 (0-95)	0.91
MRD % at EOI (≥0.01, n (%))	65 (61)	316 (59)	0.75
Complete remission (Y, n (%))	124 (98)	664 (94)	**0.04**
EFS at 5 years (%, (95% CI))	78 (70-85)	79 (76-82)	0.85
RR at 5 years ((%, (95% CI))	18 (12-26)	16 (12-19)	0.48
OS at 5 years ((%, (95% CI))	85 (78-91)	86 (83-88)	0.89

Table 6.1. Clinical characteristics and outcomes of patients in the tested cohort and total cohort (tested + non-tested cases) from UKALL 2003 and UKALL 2011 (WBC, white blood cell count; CNS, central nervous system; BM, bone marrow; MRD, minimal/measurable residual disease; EOI, end of induction; EFS, event-free survival; RR, relapse rate; OS, overall survival; CI, confidence interval; Y, yes).

6.3.2. The Genetic Classifier (TALLSorts)

RNA-seq data from 126 T-cell ALL samples were run through a pipeline of the TALLSorts algorithm, which was a validation process of this classifier in this thesis. The distribution of predicted scores was checked over subtypes estimated by the algorithm. BCL11B, HOXA_KMT2A, NKX2, TLX1 and TLX3 subtypes were clearly separated by the classifier, whereas the rest of the predicted subtypes, such as HOXA_MLLT10, TAL/LMO and diverse subtype, were diffused (Figure 6.6.). More specifically, HOXA_MLLT10 subtype was widely spread across all the samples, and most of the second calls belonged to this subtype (Figure 6.7.).

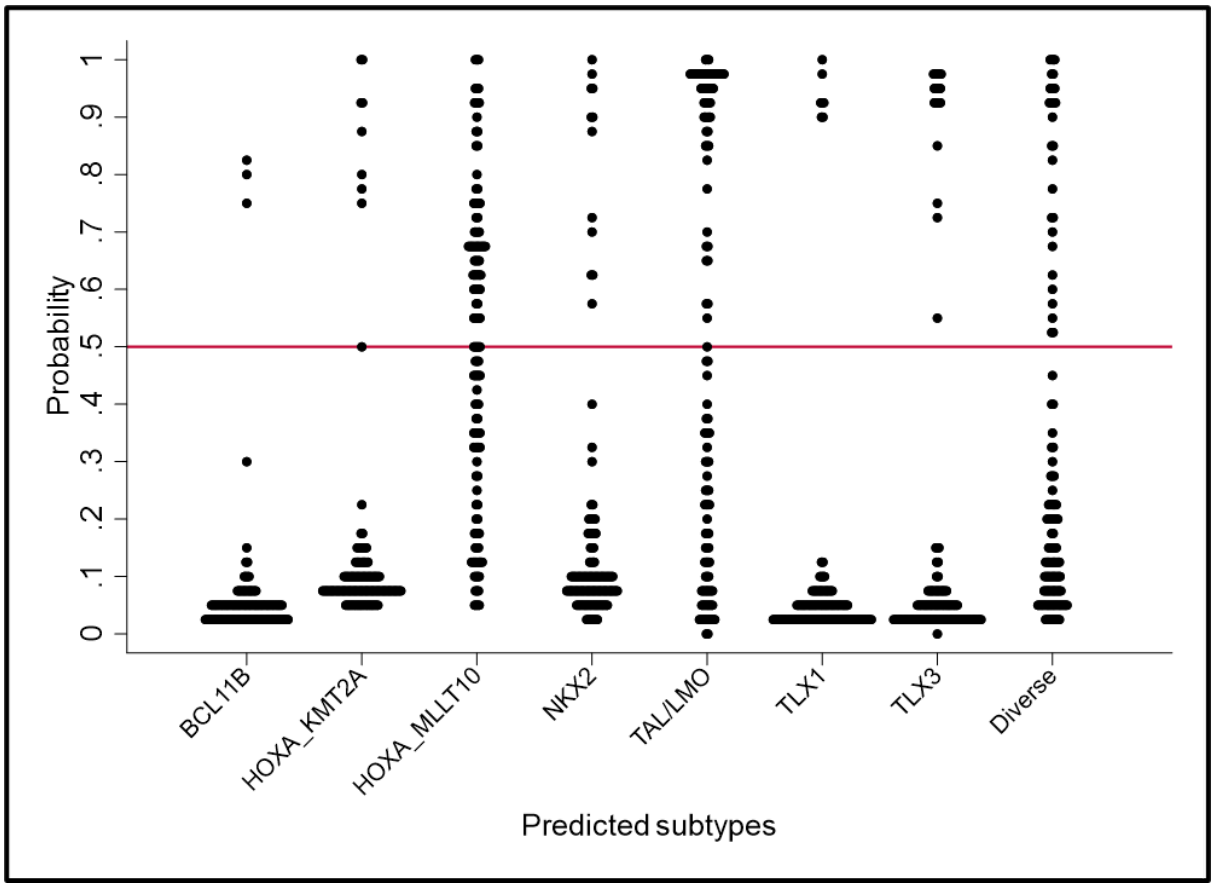

Figure 6.6. Probability (predicted) scores of the subtypes estimated by the TALLSorts (red line indicates the probability score of 0.5).

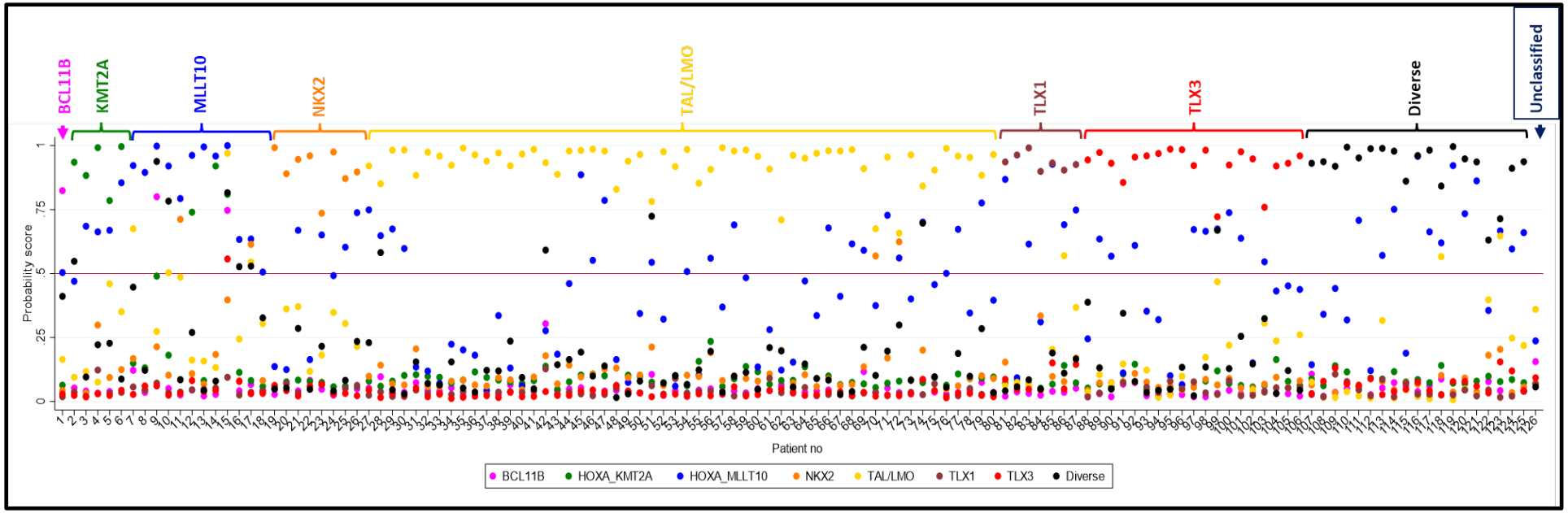

Figure 6.7. Distribution of the probability (predicted) scores of the subtypes over patients (samples).

The algorithm defined eight subgroups on the basis of the highest probability (predicted) scores. The most frequent subgroup was TAL/LMO (54, 42.9%), while BCL11B was the rarest subgroup (1, 0.8%). 1 out of 126 samples could not be classified by the algorithm as its highest score did not reach the probability of 0.50 (Table 6.2.).

Predicted Subtypes (n=126)	n (%)
BCL11B	1 (0.8)
HOXA_KMT2A	5 (4.0)
HOXA_MLLT10	12 (9.5)
NKX2	8 (6.4)
TAL/LMO	54 (42.9)
TLX1	7 (5.6)
TLX3	19 (15.1)
Diverse	19 (15.1)
Unclassified	1 (0.8)

Table 6.2. The number and proportion of the predicted subtypes by the TALLSorts.

Characterised genetic subgroup consisted of positive cases with known gene fusions detected by the laboratory methods (n=62, 49%), whereas those who were not fully tested for the known gene fusions or those without the known gene fusions were classified as uncharacterised genetic subgroup (n=64, 51%). These two subgroups were further investigated separately below.

6.3.2.1. The Characterised Genetic Subgroup

Genetic data, containing gene fusions identified by the laboratory methods, were available in 62 out of 126 samples. In this subgroup, characterised genetic subgroup, the highest proportion belonged to the predicted TAL/LMO subtype, while the predicted diverse subtype had the lowest proportion (Figure 6.8.).

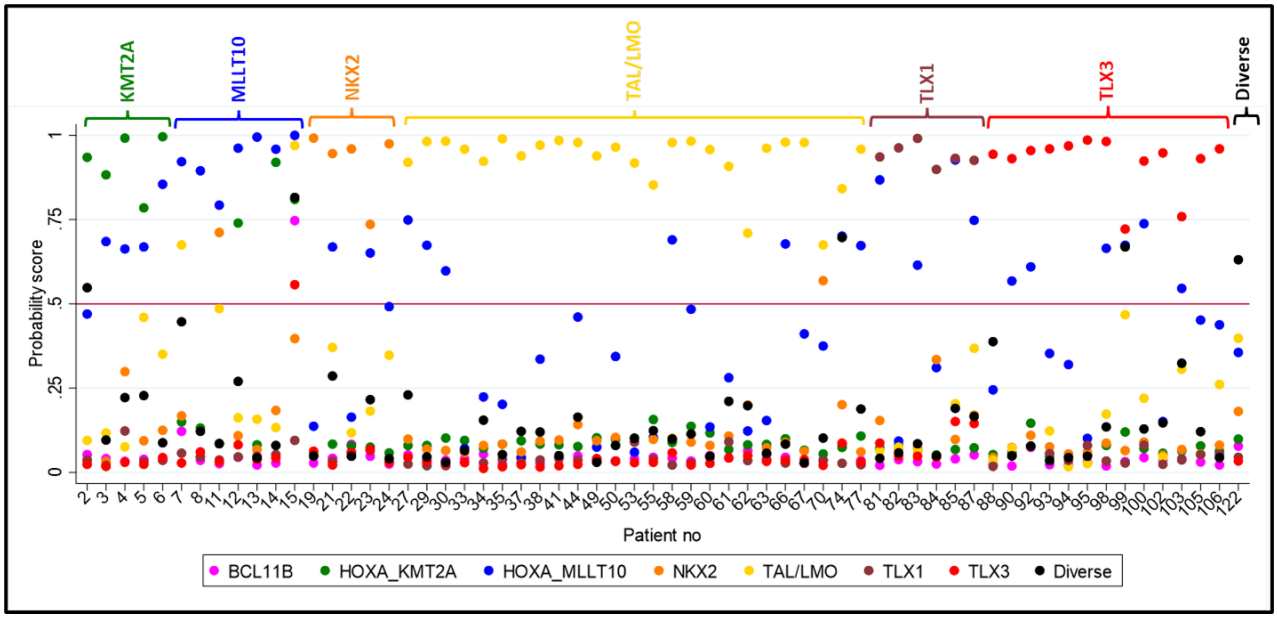

Figure 6.8. Distribution of probability (predicted) scores of the subtypes in the characterised genetic subgroup. The x-axis represents each patient (sample), and the y-axis indicates the probability (predicted) scores of the subtypes.

DNA subtypes (standard of care testing) and predicted subtypes by TALLSorts were compared using a confusion matrix. It revealed that 54 out of 62 samples were correctly predicted by the classifier, indicating that probability of predicting correct subtypes was 87%. Moreover, all cases with TLX1 subtype (n=6) and TLX3 subtype (n=13) were correctly estimated by the algorithm, and those with TAL/LMO subtype were also predicted with accuracy of 89% (24/27). On the other hand, the accuracy was lower for some subtypes, such as HOXA KMT2A (63%) and NKX2 (71%) (Figure 6.9., Table 6.3.).

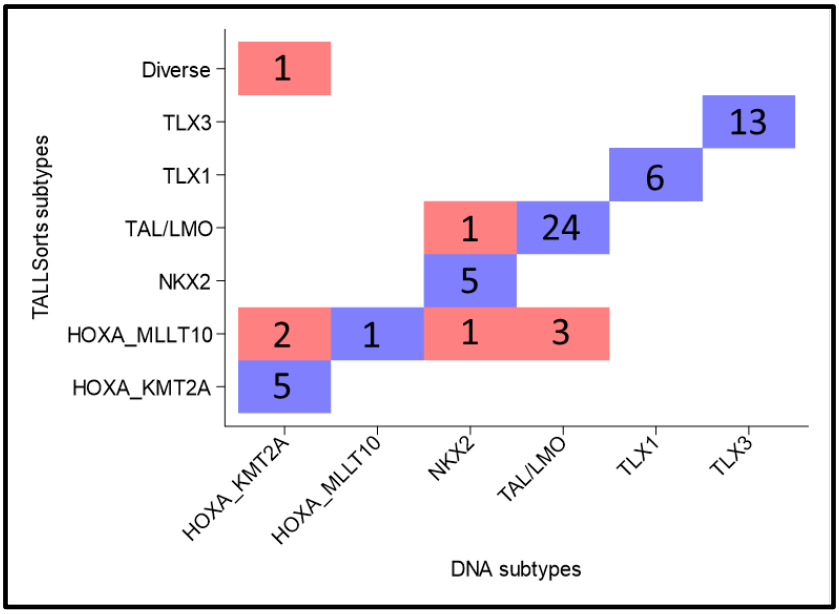

Figure 6.9. Confusion matrix comparing the DNA subtypes and predicted subtypes in the characterised genetic subgroup (the number of patients/samples was shown in the box).

Subtypes	Total	Matched subtype	Unmatched subtypes
HOXA_KMT2A	8 (100%)	5 (63%)	3 (37%)
HOXA_MLLT10	1 (100%)	1 (100%)	0 (0%)
NKX2	7 (100%)	5 (71%)	2 (29%)
TAL/LMO	27 (100%)	24 (89%)	3 (11%)
TLX1	6 (100%)	6 (100%)	0 (0%)
TLX3	13 (100%)	13 (100%)	0 (0%)

Table 6.3. The accuracy of TALLSorts predicting correct subtypes (DNA subtypes) detected by the laboratory methods.

Of 62 cases, eight calls did not match DNA subtypes in the characterised genetic subgroup (Figure 6.9.). The probability scores of these eight subtypes were visually checked in Figure 6.10. 5 out of 8 samples (patients 7, 11, 12, 14, and 15) had at least two positive calls (probability score >0.5) and were predicted as HOXA_MLLT10 subtype. More interestingly, these five samples matched their DNA subtypes if the second highest scores were considered as the final predicted subtypes. On the other hand, 4 out of 5 cases had poor-quality samples (only one sample (sample 11) with good quality).

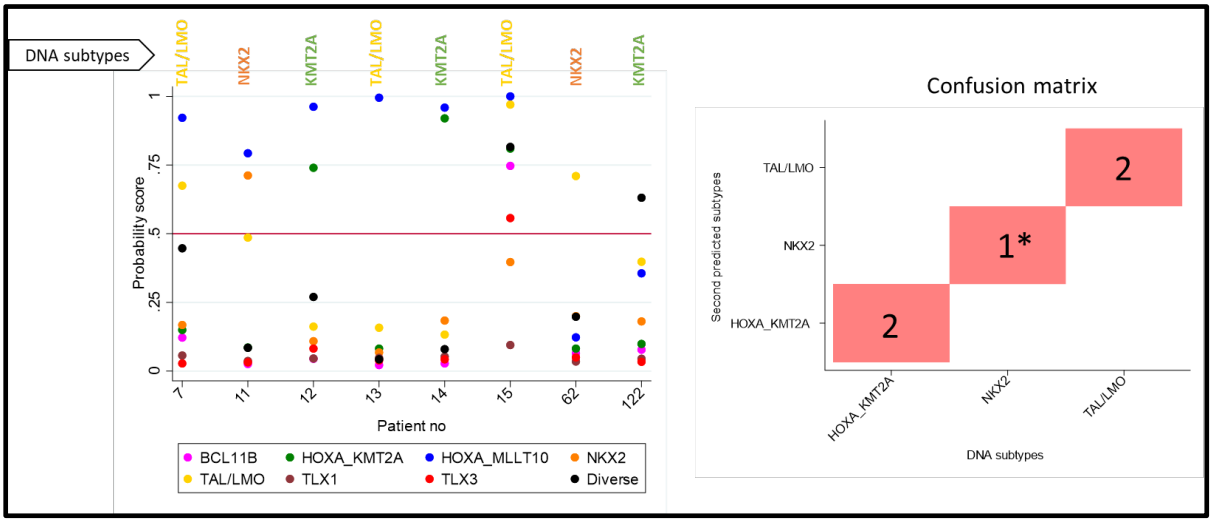

Figure 6.10. Distribution of probability (predicted) scores and confusion matrix for unmatched subtypes in the characterised genetic subgroup. Confusion matrix shows the predicted subtypes with the second highest calls over the cut-off of 0.5 (*, good quality).

To check the impact of sample/data quality, distribution of the samples with good or poor quality was investigated in the characterised genetic subgroup. 69% of the samples with poor quality were correctly estimated by the algorithm, and the predictions of the samples with good quality matched their DNA subtypes with a proportion of 93% (Table 6.4).

Characterised genetic subgroup (n=62)	good quality (n, %)	poor quality (n, %)	*p*-value
matched subtypes	43 (93)	11 (69)	0.011
unmatched subtypes	3 (7)	5 (31)	

Table 6.4. Distribution of sample quality in the characterised genetic subgroup.

6.3.2.2. The Uncharacterised Genetic Subgroup

The number of samples assigned to uncharacterised genetic subgroup was 64 (51%). Nine of them did not have any known gene fusions; however, the algorithm classified them as HOXA_MLLT10 subtype (n=1), TAL/LMO subtype (n=3), TLX1 subtype (n=1), and TLX3 subtype (n=4) (Figure 6.11.). There was only one sample (sample 9) with poor quality among these cases, others were in good quality.

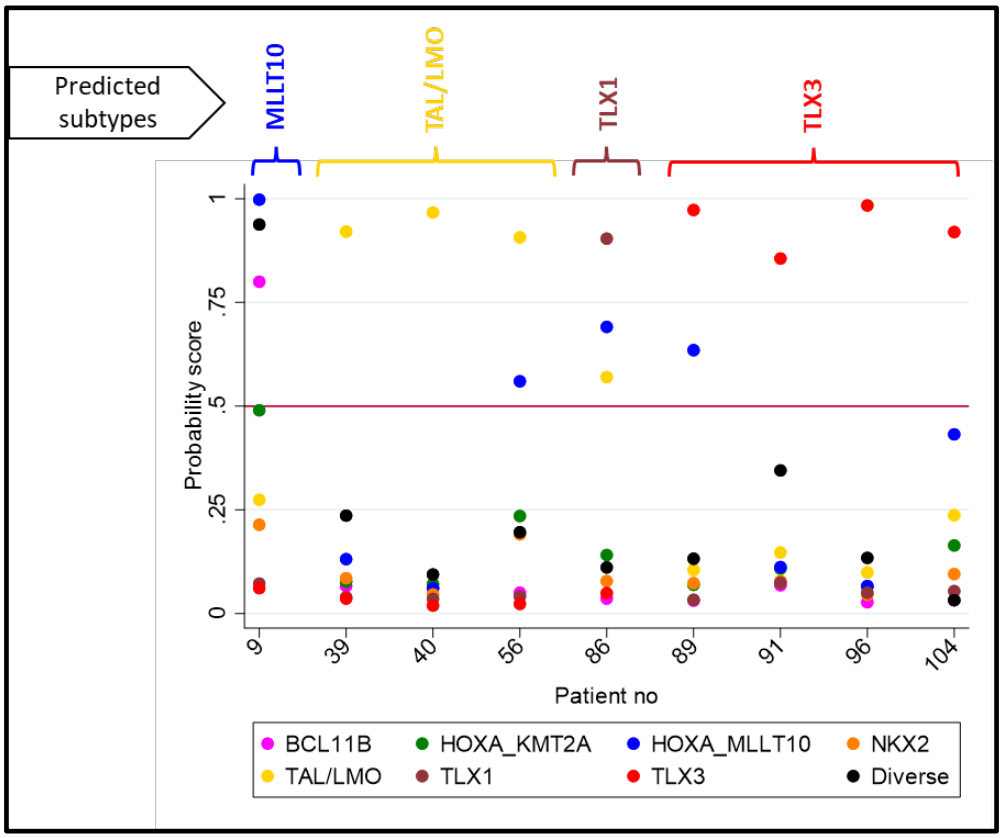

Figure 6.11. The probability (predicted) scores of the samples without the known gene fusions.

6.3.2.3. Distribution of the Predicted Subtypes within the Risk Groups

The predicted subtypes by the TALLSorts were checked within the standard-risk and high-risk groups identified by the D8BM% model. A higher proportion of TLX3 and NKX2 subtypes was observed in the standard-risk group, while the high-risk group had more diverse subtype (Figure 6.12.). However, only diverse subtype was statistically varied between the risk groups (7% vs 36%, *p*-value <0.001) (Table 6.5.).

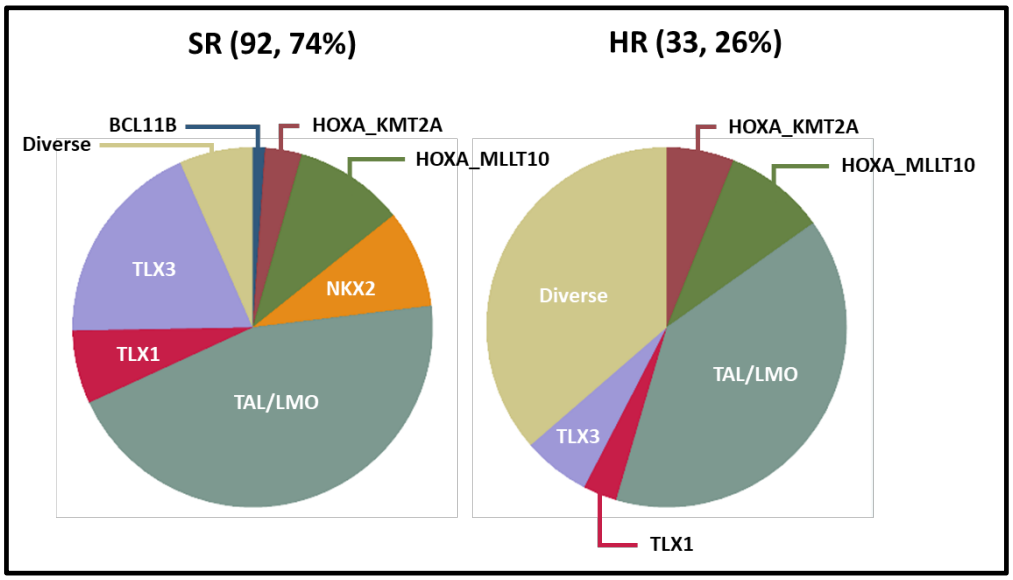

Figure 6.12. Proportion of the predicted subtypes in the risk groups defined by the D8BM% model. (SR, standard-risk; HR, high-risk).

UKALL 2003 & UKALL 2011	SR (92, 74%)	HR (33, 26%)	*p*-value
BCL11B subtype (Y, n (%))	1 (1)	0 (0)	0.73
HOXA_KMT2A subtype (Y, n (%))	3 (3)	2 (6)	0.49
HOXA_MLLT10 subtype (Y, n (%))	9 (10)	3 (9)	1.0
NKX2 subtype (Y, n (%))	8 (9)	0 (0)	0.11
TAL/LMO subtype (Y, n (%))	41 (45)	13 (39)	0.57
TLX1 subtype (Y, n (%))	6 (7)	1 (3)	0.67
TLX3 subtype (Y, n (%))	17 (19)	2 (6)	0.09
Diverse subtype (Y, n (%))	6 (7)	12 (36)	**<0.001**

Table 6.5. Distribution of the predicted subtypes in the risk groups defined by the D8BM% model (SR, standard-risk; HR, high-risk; Y, yes).

6.3.3. Identification of Differentially Expressed Genes between the Risk Groups

In total, 1810 genes were dysregulated between the standard-risk group and high-risk group identified by the D8BM% model (adjusted *p*-value <0.05 & absolute log2 fold change >0), and there were no outliers or low-count genes (Figure 6.13.).

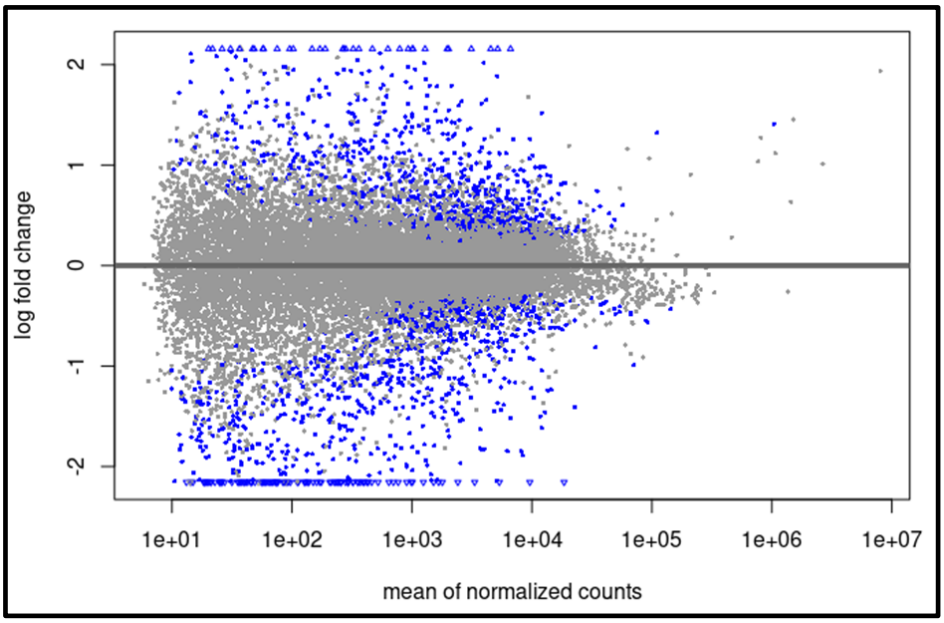

Figure 6.13. Dysregulated genes (blue) with adjusted *p*-value <0.05 and absolute log2 fold change >0 between the risk groups.

To identify important genes, 156 out of 1810 genes were selected based on adjusted *p*-value of <0.05 & absolute log2 fold change of ≥2. Among them, there were 39 up-regulated and 117 down-regulated genes in the high-risk group (Figure 6.14.). In addition to dysregulation of the genes in the risk group, gene set enrichment analysis revealed that 156 genes have different biological roles in heart development, skeletal system development and immune response (Figure 6.15.).

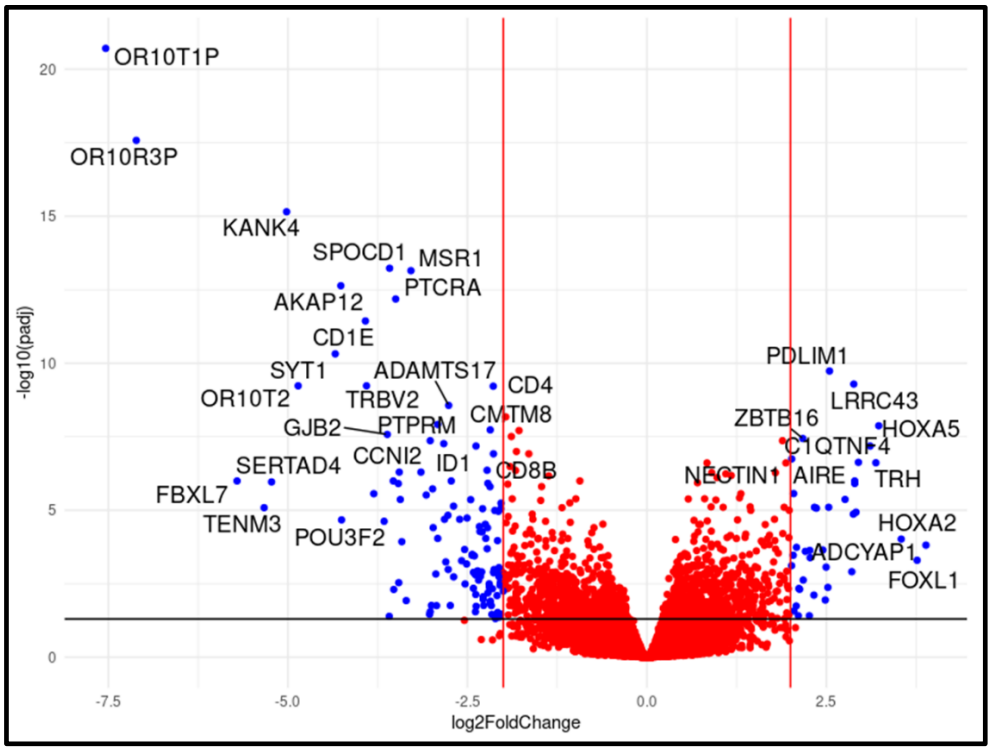

Figure 6.14. Volcano plot for differentially expressed genes with their gene symbols in the high-risk group. Blue colour demonstrates 156 genes with adjusted *p*-value <0.05 and absolute log2 fold change ≥2, and red colour indicates nonsignificant genes.

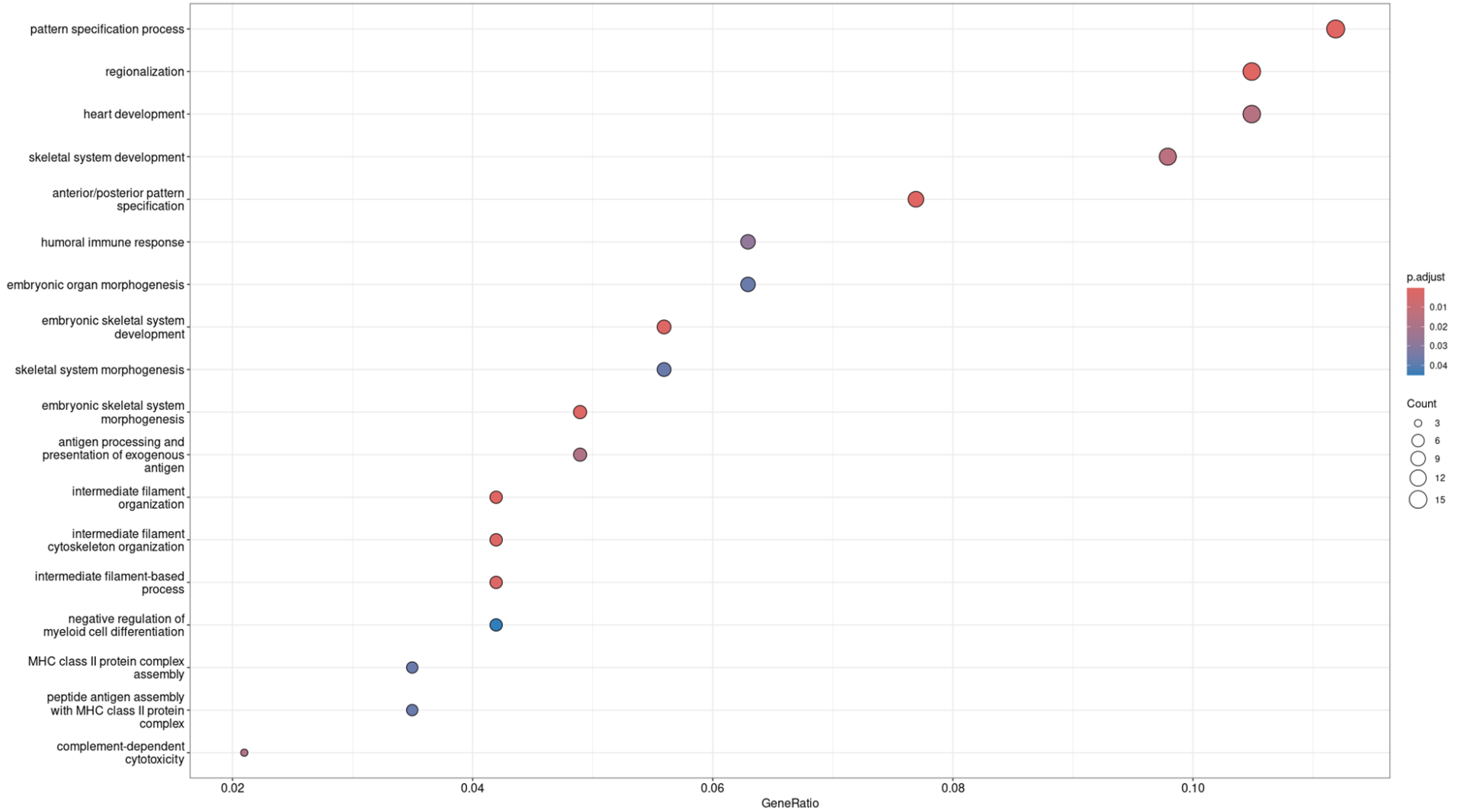

Figure 6.15. Gene set enrichment analysis for the 156 genes (*p*-adjust, adjusted *p*-value).

6.3.4. The Genes Overlapping between Two Different Strategies

The TALLSorts algorithm uses 309 genes to stratify patients into eight different T-cell ALL subtypes. When comparing the genes from TALLSorts (n=309) and differential expression analysis (n=156), there were 12 overlapping genes (*ADAMTS19, ANKRD1, FBXL7, FRZB, GJB2, HOXA3, HOXA4, HOXA6, HOXA7, MEIS1, SKIDA1, TRH*), which had various biological roles in the cells (Figure 6.16., Table 6.6.). Of 12 genes, four genes were highly dysregulated in the high-risk group (up-regulation of *TRH* and down-regulations of *ANKRD1, FBXL7, and GJB2*) (Table 6.6., Figure 6.17.).

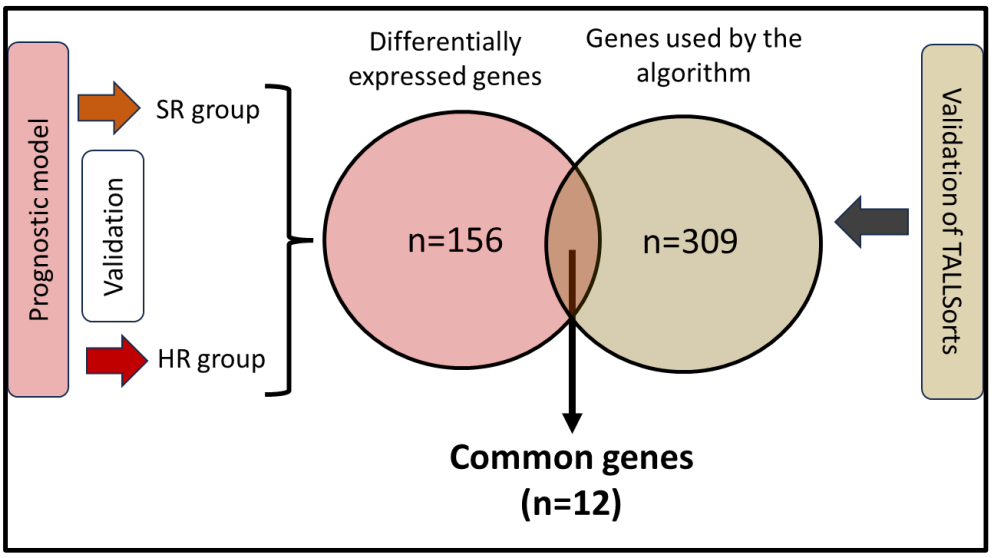

Figure 6.16. Methodology for identification of the overlapping genes between the DESeq2 model and TALLSorts algorithm (SR, standard-risk; HR, high-risk).

Gene symbol	Gene name	Biological function	log_2Fold Change	padj
FBXL7	F-box and leucine rich repeat protein 7	Ubiquitination, proteasomal degregation	-5.70816	$1.01x10^{-6}$
GJB2	gap junction protein beta 2	Gap junction channel activity	-3.61571	$2.64x10^{-8}$
ANKRD1	ankyrin repeat domain 1	Transcription factor binding, histone deacetylase binding	-3.52595	0.004966
FRZB	frizzled related protein	G protein-coupled receptor activity, Wnt-protein binding	-2.80092	$5.74x10^{-4}$
ADAMTS19	ADAM metallopeptidase with thrombospondin type 1 motif 19	Peptidase activity, metalloendopeptidase activity	-2.05231	0.042857
MEIS1	Meis homeobox 1	Sequence-specific DNA binding, chromatin binding	2.01448	$7.70x10^{-4}$
HOXA7	homeobox A7	DNA-binding transcription factor activity, transcription factor binding	2.116294	0.004594
HOXA4	homeobox A4	DNA-binding transcription factor activity, sequence-specific DNA binding	2.496061	$8.61x10^{-4}$
SKIDA1	SKI/DACH domain containing 1	NA	2.531838	$7.93x10^{-6}$
HOXA6	homeobox A6	DNA-binding transcription factor activity, sequence-specific DNA binding	2.874964	$1.36x10^{-5}$
HOXA3	homeobox A3	DNA-binding transcription factor activity, HMG box domain binding	2.912166	$1.17x10^{-5}$
TRH	thyrotropin releasing hormone	Neuropeptide hormone activity, thyrotropin-releasing hormone activity	3.189169	$2.43x10^{-7}$

Table 6.6. Summary information about 12 overlapping genes between the genetic classifier and differential expression analysis (padj, adjusted *p*-value) (information source for gene name and biological function was obtained from www.genecards.org).

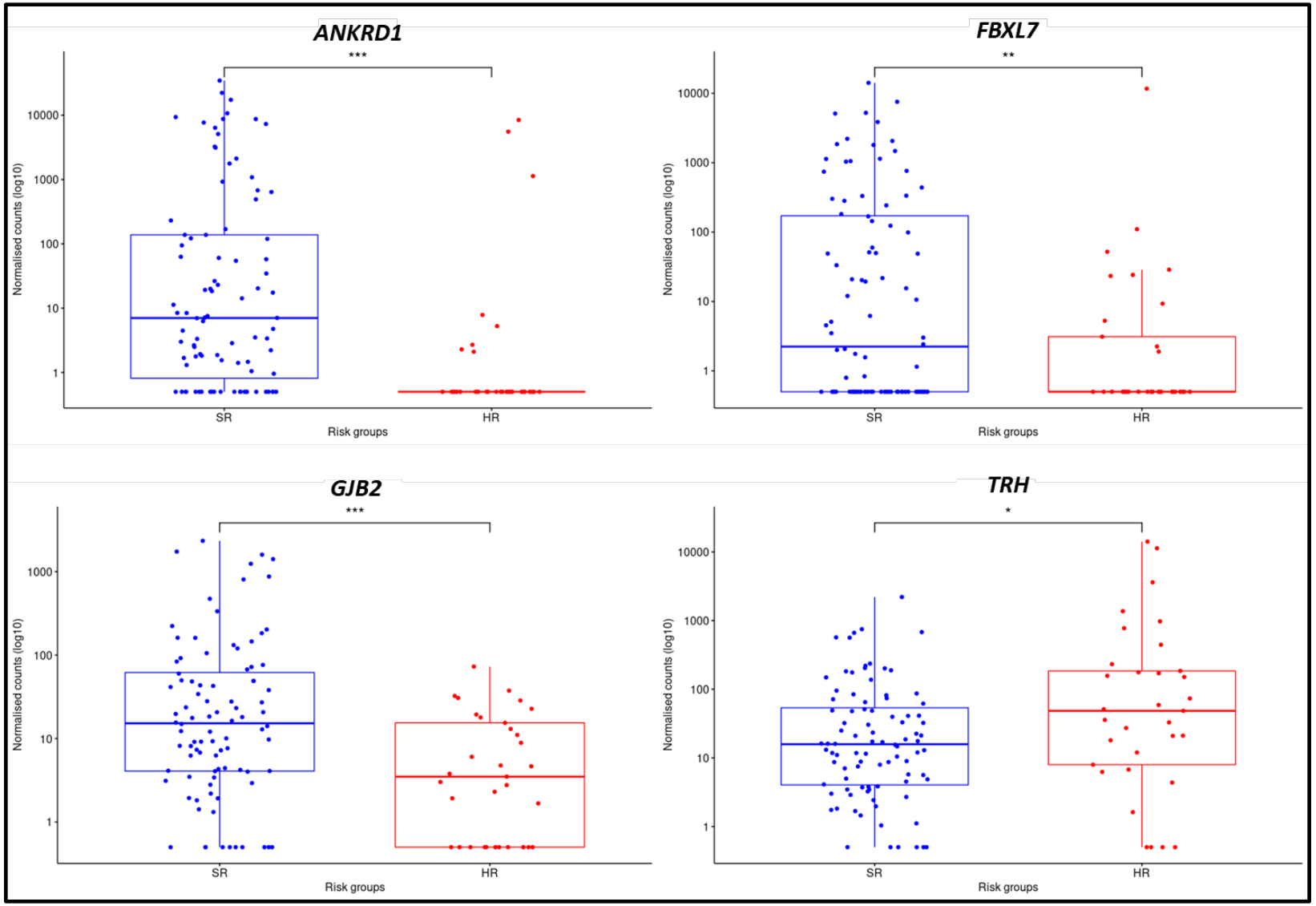

Figure 6.17. Normalised expression levels of the four highly dysregulated genes in the risk groups (SR, standard-risk; HR, high-risk).

6.3.5. Unsupervised Clustering Analysis

6.3.5.1. Identification of Optimal Cluster Number

Unsupervised clustering analysis was performed with the ConsensusClusterPlus to discover genetic subgroups. The results shown in Figure 6.18. were examined to identify optimal cluster number (number of the subgroups). The Cumulative Distribution Function (CDF) plot indicates the cumulative proportion of paired samples versus consensus values across cluster groups (k). The plot suggested that a higher and flatter curve (S-shape curve) was in cluster 5 (k=5), meaning that an optimal and stable cluster could be with five subgroups (Figure 6.18.(A)). Delta area curve (relative change in area under the CDF curve) is a quantitative measure that evaluates the difference in the area under the CDF curve as the number of clusters (k) increase. Only small drop in the change demonstrates no improvement when adding more clusters. A cluster with five subgroups (k=5) would still be the optimal cluster number based on the relative change as it was within the area that had small changes (Figure 6.18.(B)). A consensus matrix shows consensus values in blue (strong correlation) and white (weak correlation) across the samples. The ideal matrix represents the “cleanest” clusters (always either together or separate). Therefore, a cluster with five subgroups (k=5) was still supported by this matrix as an optimal k (Figure 6.18.(C)). More evidence comes from a tracking plot (cluster assignment with different colours) across the samples. Most of the samples remained within the same clusters after the cluster 5 (k=5), which indicated stable and reliable clustering (Figure 6.18.(D)).

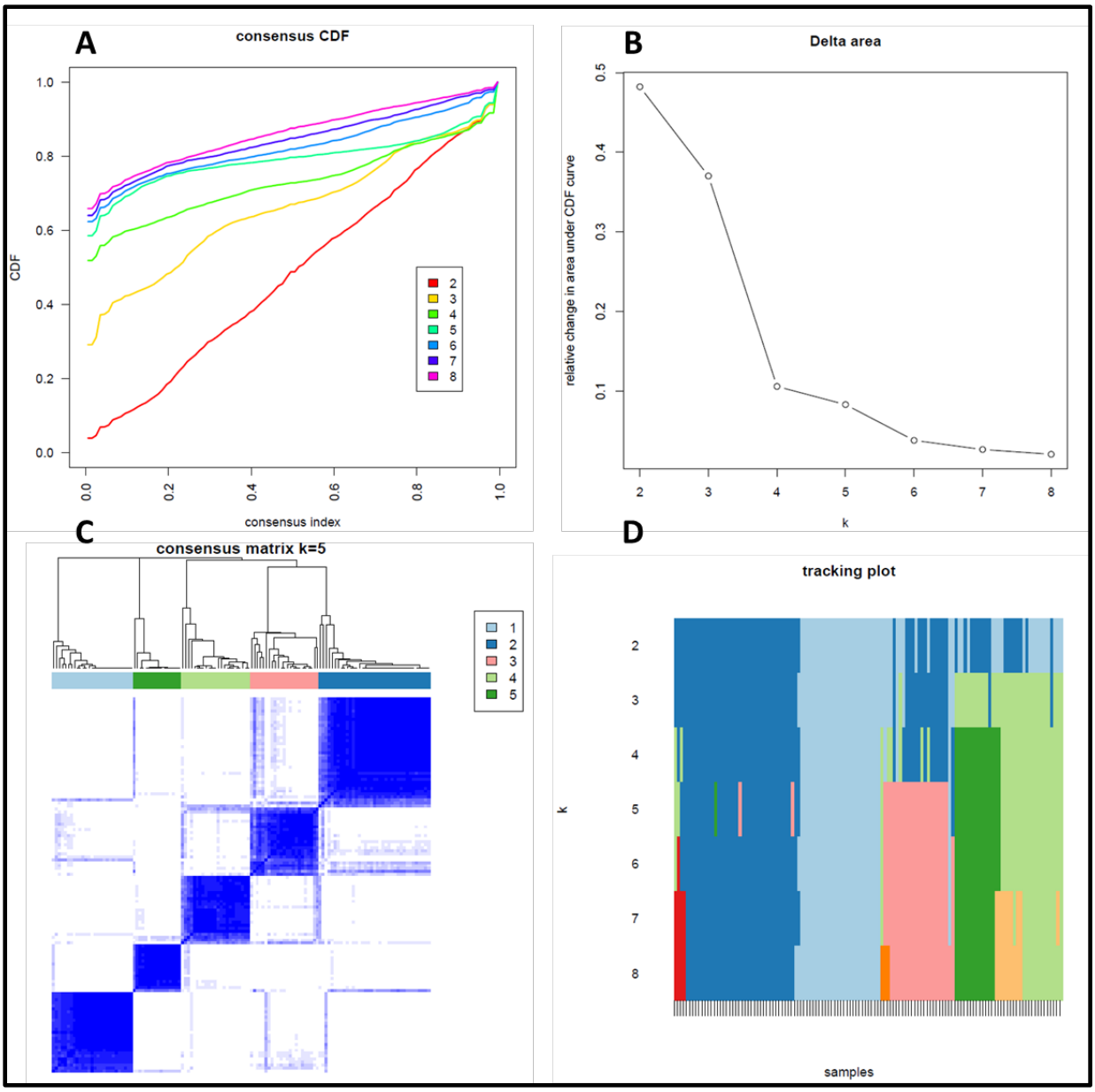

Figure 6.18. Identification of optimal cluster number using the ConsensusClusterPlus. The Cumulative Distribution Function (CDF) graph represents the cumulative proportion of paired samples and consensus values in different clusters (A), delta area curve is a quantitative measure of CDF curves (B), consensus matrix shows consensus values (C), and tracking plot demonstrates samples in different clusters (D). The colours labelled with numbers in the CDF and tracking plots represent different clusters, while in the consensus matrix, they indicate different subgroups in the cluster number (k)=5.

6.3.5.2. Distribution of Clinical Characteristics, Risk Groups and Predicted Subtypes in the Clusters

The unsupervised clustering analysis revealed that 126 samples/patients could be categorised into the following five clusters: Cluster 1 (n=27, 22%), Cluster 2 (n=37, 29%), Cluster 3 (n=23, 18%), Cluster 4 (n=23, 18%), and Cluster 5 (n=16, 13%). Cluster 1 and Cluster 5 had higher levels of WBC at diagnosis, BM blast percentages at day 8 and day 28, and more cases with positive MRD (≥0.01%), in comparison to Cluster 2 and Cluster 4. Although Cluster 1 and Cluster 5 had differences in these variables, they did not have a significant disadvantage in the rates of event-free survival, relapse and overall survival, compared to the other clusters (Table 6.7.).

Variables	Values	Cluster 1 (27, 22%)	Cluster 2 (37, 29%)	Cluster 3 (23, 18%)	Cluster 4 (23, 18%)	Cluster 5 (16, 13%)	*p*-value
Sex	male, n (%)	23 (85)	27 (73)	19 (83)	15 (65)	13 (81)	0.44
Age (years)	median (range)	8.4 (2.0-16.7)	6.0 (1.0-17.0)	9.3 (1.6-21.9)	8.2 (3.2-18.0)	11.1 (4.3-20.1)	0.08
WBCx10^9/L	median (range)	256.0 (27.8-881.0)	50.7 (3.3-777.0)	119.5 (8.6-636.0)	68.2 (8.0-436.4)	109.4 (1.0-488.0)	**0.007**
BM blast % at diagnosis	median (range)	92 (66-100)	89 (14-100)	92 (80-98)	95 (66-99)	90 (85-99)	0.54
Regimen C	n (%)	19 (70)	13 (35)	10 (43)	9 (39)	13 (81)	**0.004**
BM blast % at day 8	median (range)	50 (5-95)	10 (1-96)	6 (0-38)	13 (0-75)	65 (1-95)	**0.0001**
BM blast % at day 28	median (range)	2 (0-70)	1 (0-10)	2 (0-15)	0 (0-4.5)	2 (0-5)	**0.02**
EOI-MRD (%)	≥0.01, n (%)	23 (96)	13 (46)	12 (55)	8 (35)	9 (90)	**<0.001**
Event-free survival rate	% (95%CI)	78 (57-89)	72 (54-84)	78 (55-90)	91 (69-98)	75 (46-90)	0.72
Relapse rate	% (95%CI)	19 (8-40)	26 (14-44)	14 (5-37)	9 (2-30)	20 (7-50)	0.80
Overall survival rate	% (95%CI)	81 (61-92)	83 (66-92)	83 (60-93)	100 (-)	81 (52-94)	0.30

Table 6.7. Clinical characteristics and outcomes of patients in the clusters defined by the consensus clustering analysis (WBC, white blood cell count; BM, bone marrow, EOI-MRD, minimal/measurable residual disease at the end of induction).

In addition to clinical characteristics of the clusters, more patients in the high-risk group were observed in Cluster 1 and Cluster 5, whereas Cluster 2, Cluster 3, and Cluster 4 had more patients in the standard-risk group (Figure 6.19.). In terms of the subtypes predicted by TALLSorts, most cases with TAL/LMO subtype were found in Cluster 1 (21, 78%) and Cluster 3 (17, 77%), and Cluster 4 and Cluster 5 had more proportion of TLX3 subtype (14, 61%) and diverse subtype (13, 81%), respectively (Figure 6.19., Table 6.8.).

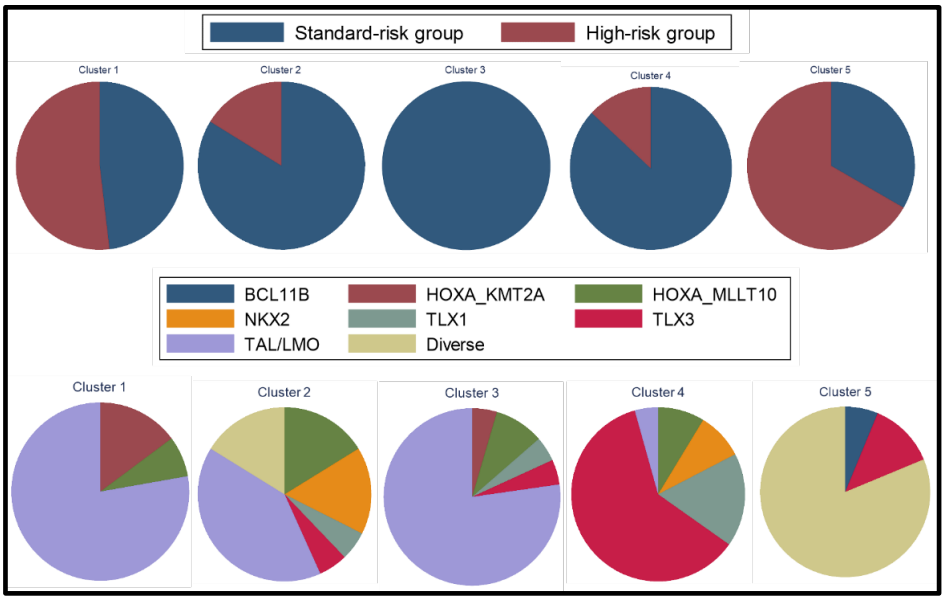

Figure 6.19. Proportion of the risk groups defined by the D8BM% model and the subtypes predicted by TALLSorts in five clusters. The predicted subtypes were based on the highest probability scores, as the TALLSorts algorithm assigned, not corrected calls.

UKALL 2003 & UKALL 2011	Cluster 1 (27, 22%)	Cluster 2 (37, 29%)	Cluster 3 (23, 18%)	Cluster 4 (23, 18%)	Cluster 5 (16, 13%)
BCL11B subtype (Y, n (%))	0 (0)	0 (0)	0 (0)	0 (0)	1 (6)
HOXA_KMT2A subtype (Y, n (%))	4 (15)	0 (0)	1 (5)	0 (0)	0 (0)
HOXA_MLLT10 subtype (Y, n (%))	2 (7)	6 (16)	2 (9)	2 (9)	0 (0)
NKX2 subtype (Y, n (%))	0 (0)	6 (16)	0 (0)	2 (9)	0 (0)
TAL/LMO subtype (Y, n (%))	21 (78)	15 (41)	17 (77)	1 (4)	0 (0)
TLX1 subtype (Y, n (%))	0 (0)	2 (5)	1 (5)	4 (17)	0 (0)
TLX3 subtype (Y, n (%))	0 (0)	2 (5)	1 (5)	14 (61)	2 (13)
Diverse subtype (Y, n (%))	0 (0)	6 (16)	0 (0)	0 (0)	13 (81)

Table 6.8. Distribution of the predicted subtypes within the clusters defined by the unsupervised analysis. The original predicted subtypes were used, as the algorithm classified according to the highest probability scores (Y, yes).

6.4. Discussion

In this chapter, bioinformatic tools were used to explore RNA sequencing data from T-cell ALL patients (n=126) for the evaluation of the genetic classifier, identification of differentially expressed genes between the risk groups defined by D8BM%, and discovery of genetic subgroups (clusters).

Overall, the accuracy of TALLSorts to correctly identify subgroups was 87% (Figure 6.9.), but it could be as high as 95% if the HOXA_MLLT10 signatures were sorted (Figure 6.10.). This rate would also be 93% ((43/46)*100)) if the samples with poor quality (16 out of 62 samples in the characterised genetic subgroup) were excluded in the analysis. Therefore, high accuracies of the algorithm in different conditions indicate that RNA sequencing data generated from the American cohort are correlated with the data generated from the UK cohort. In other words, T-cell ALL patients living in different geographical places (with a different lifestyle and environmental factors) can share similar genetic expression profiles, so this algorithm can detect T-cell ALL subtypes across all populations, despite racial or ethnic disparities. Although the genetic classifier correctly identified 100% of the cases with TLX1 and TLX3 subtypes (Table 6.3.), HOXA_MLLT10 subtype was extremely diffused over the samples (Figure 6.6., Figure 6.7.). Interestingly, the second highest probability scores of HOXA_MLLT10 subtype matched DNA subtypes (Figure 6.10.). Despite the fact that these 4 out of 5 HOXA_MLLT10 subtypes were predicted using RNA-seq data with poor quality because of DNA contamination or low amount of RNA concentration, these findings indicated that genes defined by the algorithm for HOXA_MLLT10 subtype could not be unique for identification of *MLLT10* fusions and could involve in common genetic mechanisms in the other subtypes. On the other hand, the TALLSorts algorithm correctly identified 11 (69%) samples with poor sequencing quality in total (Table 6.4.); therefore, it can be useful to identify T-cell ALL subtypes in RNA-seq data with poor-quality sequencing.

The DESeq2 analysis revealed that 156 significant genes (*p*-adjusted <0.05 and absolute log2 fold change ≥2) were differentially expressed between the standard-risk group and high-risk group defined by the D8BM% model (Figure 6.14.). It was also observed that 12 out 156 genes (*ADAMTS19, ANKRD1, FBXL7, FRZB, GJB2, HOXA3, HOXA4, HOXA6, HOXA7, MEIS1, SKIDA1, TRH*) overlapped with a subset of the genes used by the TALLSorts classifier (Figure 6.16., Table 6.6.). Although *HOXA* genes are already known to be involved in T-cell ALL biology (Soulier *et al.*, 2005), to our knowledge, there is no information about the roles of other genes, which overlap between the two different strategies, in the literature.

However, some of these genes have been found to be associated with prognosis in other cancer types. For example, Moro and colleagues reported that *FBXL7* had an anti-metastatic role, and low expression of this gene was an indicator of poor prognosis in pancreas and prostate cancers (Moro *et al.*, 2020). Another research group demonstrated that the expression of *THR* was an independent prognostic factor in acute myeloid leukaemia (AML), and AML patients who had higher expression level of the gene tended to be more sensitive to chemotherapeutic agents (Gao *et al.*, 2022).

The unsupervised analysis identified five different genetic clusters (subgroups) based on gene expression profiles (Figure 6.18.). Despite the fact that Cluster 1 and Cluster 5 had high risk features (poor marrow and MRD responses, and high proportion of the high-risk patients), there was no statistically significant differences among all clusters in event-free survival, relapse rate and overall survival (Figure 6.19., Table 6.7.). The findings support that T-cell ALL is a genetically heterogenous cancer type, and the genetic subtypes have not been widely accepted as prognostic factors (Durinck *et al.*, 2015). Although a recent report has suggested a new Next-Generation Sequencing (NGS)-based genetic classifier with 14 coding genes (Simonin *et al.*, 2024), it needs to be validated in other T-cell ALL cohorts. Currently, researchers have started investigating non-coding genome and enhancer interactions to provide new insights on T-cell ALL genetics/biology (Liu *et al.*, 2017b; Pölönen *et al.*, 2024).

Availability of RNA-seq data (n=126) and samples with poor sequencing quality (n=28, 22%) were the main limitations for the bioinformatic analyses. Also, the number of samples that were not fully tested for genetic abnormalities was another restriction, which limited the evaluation of the genetic classifier in the UKALL cohorts.

In conclusion, the TALLSorts can be a useful tool to detect most of T-cell ALL subtypes, specifically TLX1 and TLX3 subtypes. Other than coding genes or the known subtypes, non-coding regions and enhancer activities on the genome can provide more comprehensive information related to prognosis and treatment stratification for T-cell ALL patients.

Chapter 7. General Discussion

7.1. Summary of Findings and Discussion

This thesis aims to evaluate and validate potential risk and prognostic factors, improving current MRD-based stratification and treatment decisions in T-cell ALL. Chapter 3 addresses the first main objective by investigating demographic and clinical factors, such as age, sex, white blood cell count (WBC) at diagnosis, National Cancer Institute-Rome (NCI-Rome) classification, central nervous system (CNS) involvement, organomegalies, marrow and MRD responses, and genetic abnormalities. In the UKALL trials, from UKALL VIII to UKALL 2011, only treatment response variables (early marrow response and MRD response) had meaningful prognostic values among these factors, which are also supported by the previous studies in the literature. For example, there is no prognostic implication of sex difference in T-cell ALL patients treated with contemporary treatments (Gupta *et al.*, 2022; Burns *et al.*, 2021). Although age and WBC at diagnosis, therefore NCI-Rome classification, are independent prognostic factors in B-cell ALL (Schultz *et al.*, 2006), the prognostic power of these factors is debatable in T-cell ALL (Burns *et al.*, 2021; Patrick *et al.*, 2014b; Petit *et al.*, 2018; Melchior *et al.*, 2012). The main genetic subtypes are still controversial, as discussed in Chapter 1 (section 1.3.3.). Although there is some discussion about absolute threshold and effective time point of MRD, the level of 0.01% is traditionally accepted for MRD classification (Saygin *et al.*, 2022), and prognosis of T-cell ALL patients has dramatically improved thanks to the MRD-based stratifications with intensified protocols (Raetz and Teachey, 2016; Hunger *et al.*, 2012). In addition to MRD response, early treatment response variables, such as marrow response or blood response, are prognostic indicators in T-cell ALL (Griffin *et al.*, 2000; Melchior *et al.*, 2012; Möricke *et al.*, 2013).

Chapter 4 and Chapter 5 discuss the clinical usefulness of developing and validating baseline (static) and dynamic prognostic models, evaluating the second main objective of this thesis. The baseline model (D8BM% model), consisting of only bone marrow blast percentage at day 8, classified T-cell ALL patients into the standard-risk and high-risk groups in three consecutive UKALL trials, and it improved MRD stratification at the end of induction. The threshold of the D8BM% model corresponded to the raw marrow value of approximately 41%, indicating a different threshold from the cut-off used in traditional early marrow classification (25%). In other words, an arbitrary cut-off was not used in the model that was developed and validated for T-cell ALL patients specifically. Although morphologic assessment of the bone marrow samples is the main limitation to measuring level of early marrow responses, the findings demonstrate that early marrow response can be effectively

used for the risk stratification at very early stage of the treatment. Machine learning algorithms can enhance the efficiency of bone marrow assessment and can make the assessment more accurate and objective (Huang *et al.*, 2020; Wang *et al.*, 2022; Chandradevan *et al.*, 2020); therefore, low-cost marrow measurements can allow early intervention during the therapy in the low- and middle-income countries, which have limited access to the contemporary laboratory tools (Moreira *et al.*, 2023; Hayashi, Makimoto and Yuza, 2024). On the other hand, for the first time, dynamic models (NLME and AUC models) were developed for the evaluation of the treatment response (marrow and MRD values) trends at multiple time points in T-cell ALL and successfully identified the risk groups in UKALL 2003 and UKALL 2011. Validation of these models is another piece of evidence supporting that early treatment response variables are crucial to stratifying T-cell ALL patients. EOI-EOC MRD model, consisting of only MRD values at three time points, also separated patients with high-risk features into two risk groups. When comparing dynamic models (NLME model, AUC model, EOI-EOC MRD model) and the models with single-time point measurement (D8BM% model and MRD stratifications at the end of induction/consolidation), there was no meaningful advantage of using dynamic models over static models for the discrimination of relapse and non-relapse cases and improvement of current MRD stratification. A similar result was found in an article investigating dynamics of MRD response in B-cell ALL and concluding that a combination result of MRD response variables at EOI and EOC did not provide considerable prognostic information (Popov *et al.*, 2023b). Taking all findings together, dynamic interactions of the treatment response variables did not have additional prognostic value in T-cell ALL risk stratification, so measurements at single-time point were sufficient to stratify T-cell ALL patients for further treatment options.

The third main objective is examined in Chapter 6 where the bioinformatic tools have been used for the genetics of T-cell ALL. The genetic classifier using RNA-seq data, TALLSorts, detected the majority of the subtypes in T-cell ALL, and the DESeq2 algorithm identified 156 differentially expressed genes between the risk groups defined by the D8BM% model. Furthermore, there were 12 overlapping genes between these two detection strategies. To our knowledge, the roles of most overlapping genes in T-cell ALL are uncertain although some of them are associated with unfavourable prognosis in other cancers (Moro *et al.*, 2020; Gao *et al.*, 2022). Additionally, the unsupervised clustering analysis on the basis of the gene expression profile revealed five different genetic subgroups; however, these subgroups did not differ from each other in terms of clinical endpoints. This can explain why recent studies

have focused on non-coding genome and enhancer interactions for the deregulation of oncogene activations in T-cell ALL (Liu *et al.*, 2017b; Pölönen *et al.*, 2024).

Overall, early treatment response measured at a single-time point before the induction phase is useful and efficient to stratify T-cell ALL patients for the prevention of side effects of treatment toxicity; therefore, it should be incorporated into current risk stratification for further treatment allocation in this disease.

7.2. Strengths and Limitations

T-cell ALL is a rare disease among other cancer types, so it can be difficult to have enough number of patients for epidemiological investigations. Comprehensive and consecutive UKALL trials carried out since 1970s were used in this thesis; therefore, sample size of T-cell ALL population was enough to do all the analyses. In addition, the lowest median follow-up time was five years in the latest trial (UKALL 2011), indicating sufficient time period to perform the survival analyses. The baseline (static) and dynamic models were successfully developed and validated using internal and external evaluations ("horizontal validation") in the five trials that had different randomisations and stratification strategies.

On the other side, there were some limitations in this thesis. Marrow and MRD responses were not evaluated in early trials in model development process because there was no available treatment response data. Also, missing data were imputed by the KNN imputation method or with random numbers based on previous findings before prognostic models were developed. Another limitation was about the detection of the genetic abnormalities, which were not fully tested in the trials; therefore, patients were categorised into four broader genetic subgroups to increase sample size of each genetic subtype. Although available samples from patients on UKALL 2003 and UKALL 2011 were sent to Illumina centre for sequencing, the number of cases with sequencing data was limited for bioinformatic analyses.

7.3. Options for Expanding This Work

The D8BM% model was built using marrow response values that were measured at day 8 by the subjective method (light microscopy). Therefore, evaluation of this model in other clinical trials, other than UKALL trials, is crucial to check its predictive ability for clinical usefulness. Additionally, machine learning algorithms can be further investigated to improve accuracy of the conventional marrow assessment for clinical utility. Also, an investigation of early MRD response before the end of induction are required for the evaluation of its additional prognostic impact in T-cell ALL risk stratification.

The NLME and AUC models, which were based on marrow and MRD values, were assessed by different laboratory techniques (light microscopy and PCR). The dynamic models should be, therefore, studied in other T-cell ALL cohorts, having response data measured by an established method for accurate assessment of clinical usefulness. For example, flow cytometry (FC) is used to detect leukemia-associated immunophenotypes for early treatment response in ALL patients (Basso *et al.*, 2009; Coustan-Smith *et al.*, 2002; Campbell *et al.*, 2023), so FC can be a standard method to repeatedly measure blasts in the bone marrow at very early stages of the treatment for the examination of early treatment response dynamics.

EOI-EOC MRD model was built using MRD values at different time points for only high-risk patients in UKALL 2011. This dynamic model can be evaluated in a cohort with large sample size of T-cell ALL patients in all risk groups, so potential advantage of MRD dynamics can be also examined in a patient group with low-risk features.

Bioinformatic analyses, using data generated from 126 samples, identified different gene expression profiles between both the risk groups defined by the D8BM% model and the clusters identified by the ConsensusClusterPlus. A larger sample size of sequencing data can be useful for more conclusive results in the bioinformatic analyses.

To conclude, T-cell ALL is a genetically heterogeneous and aggressive haematological malignancy, and MRD measured at the end of induction/consolidation is the strong prognostic factor, which is widely used for the risk stratification in T-cell ALL. In this thesis, only early treatment response has been found to be associated with both prognosis and improvement of MRD stratification at the end of induction. Hence, early treatment response should be integrated into treatment allocation strategy in T-cell ALL as treatment protocol can be revised at very early stage of the therapy to reduce the side effects of chemotherapy.

Chapter 8. References

Abd ElHafeez, S., D'Arrigo, G., Leonardis, D., Fusaro, M., Tripepi, G. and Roumeliotis, S. (2021) 'Methods to Analyze Time-to-Event Data: The Cox Regression Analysis', *Oxid Med Cell Longev,* 2021, pp. 1302811.

Alaggio, R., Amador, C., Anagnostopoulos, I., Attygalle, A. D., Araujo, I. B. O., Berti, E., Bhagat, G., Borges, A. M., Boyer, D., Calaminici, M., Chadburn, A., Chan, J. K. C., Cheuk, W., Chng, W. J., Choi, J. K., Chuang, S. S., Coupland, S. E., Czader, M., Dave, S. S., de Jong, D., Du, M. Q., Elenitoba-Johnson, K. S., Ferry, J., Geyer, J., Gratzinger, D., Guitart, J., Gujral, S., Harris, M., Harrison, C. J., Hartmann, S., Hochhaus, A., Jansen, P. M., Karube, K., Kempf, W., Khoury, J., Kimura, H., Klapper, W., Kovach, A. E., Kumar, S., Lazar, A. J., Lazzi, S., Leoncini, L., Leung, N., Leventaki, V., Li, X. Q., Lim, M. S., Liu, W. P., Louissaint, A., Jr., Marcogliese, A., Medeiros, L. J., Michal, M., Miranda, R. N., Mitteldorf, C., Montes-Moreno, S., Morice, W., Nardi, V., Naresh, K. N., Natkunam, Y., Ng, S. B., Oschlies, I., Ott, G., Parrens, M., Pulitzer, M., Rajkumar, S. V., Rawstron, A. C., Rech, K., Rosenwald, A., Said, J., Sarkozy, C., Sayed, S., Saygin, C., Schuh, A., Sewell, W., Siebert, R., Sohani, A. R., Tooze, R., Traverse-Glehen, A., Vega, F., Vergier, B., Wechalekar, A. D., Wood, B., Xerri, L. and Xiao, W. (2022) 'The 5th edition of the World Health Organization Classification of Haematolymphoid Tumours: Lymphoid Neoplasms', *Leukemia,* 36(7), pp. 1720-1748.

Allen, A., Sireci, A., Colovai, A., Pinkney, K., Sulis, M., Bhagat, G. and Alobeid, B. (2013) 'Early T-cell precursor leukemia/lymphoma in adults and children', *Leuk Res,* 37(9), pp. 1027-34.

Arber, D. A., Orazi, A., Hasserjian, R., Thiele, J., Borowitz, M. J., Le Beau, M. M., Bloomfield, C. D., Cazzola, M. and Vardiman, J. W. (2016) 'The 2016 revision to the World Health Organization classification of myeloid neoplasms and acute leukemia', *Blood,* 127(20), pp. 2391-2405.

Arber, D. A., Orazi, A., Hasserjian, R. P., Borowitz, M. J., Calvo, K. R., Kvasnicka, H.-M., Wang, S. A., Bagg, A., Barbui, T., Branford, S., Bueso-Ramos, C. E., Cortes, J. E., Dal Cin, P., DiNardo, C. D., Dombret, H., Duncavage, E. J., Ebert, B. L., Estey, E. H., Facchetti, F., Foucar, K., Gangat, N., Gianelli, U., Godley, L. A., Gökbuget, N., Gotlib, J., Hellström-Lindberg, E., Hobbs, G. S., Hoffman, R., Jabbour, E. J., Kiladjian, J.-J., Larson, R. A., Le Beau, M. M., Loh, M. L.-C., Löwenberg, B., Macintyre, E., Malcovati, L., Mullighan, C. G., Niemeyer, C., Odenike, O. M., Ogawa, S., Orfao, A., Papaemmanuil, E., Passamonti, F., Porkka, K., Pui, C.-H., Radich, J. P., Reiter, A., Rozman, M., Rudelius, M., Savona, M. R., Schiffer, C. A., Schmitt-Graeff, A., Shimamura, A., Sierra, J., Stock, W. A., Stone, R. M., Tallman, M. S., Thiele, J., Tien, H.-F., Tzankov, A., Vannucchi, A. M., Vyas, P., Wei, A. H., Weinberg, O. K., Wierzbowska, A., Cazzola, M., Döhner, H. and Tefferi, A. (2022) 'International Consensus Classification of Myeloid Neoplasms and Acute Leukemias: integrating morphologic, clinical, and genomic data', *Blood,* 140(11), pp. 1200-1228.

Attarbaschi, A., Pisecker, M., Inthal, A., Mann, G., Janousek, D., Dworzak, M., Pötschger, U., Ullmann, R., Schrappe, M., Gadner, H., Haas, O. A., Panzer-Grümayer, R., Strehl, S. and Group, o. b. o. t. A. B. F. M. S. (2010) 'Prognostic relevance of TLX3 (HOX11L2) expression in childhood T-cell acute lymphoblastic leukaemia treated with Berlin–Frankfurt–Münster (BFM) protocols containing early and late re-intensification elements', *British Journal of Haematology,* 148(2), pp. 293-300.

Awan, S. E., Bennamoun, M., Sohel, F., Sanfilippo, F. and Dwivedi, G. (2022) 'A reinforcement learning-based approach for imputing missing data', *Neural Computing and Applications,* 34(12), pp. 9701-9716.

Ballerini, P., Blaise, A., Busson-Le Coniat, M., Su, X. Y., Zucman-Rossi, J., Adam, M., van den Akker, J., Perot, C., Pellegrino, B., Landman-Parker, J., Douay, L., Berger, R. and Bernard, O. A. (2002) 'HOX11L2 expression defines a clinical subtype of pediatric T-ALL associated with poor prognosis', *Blood,* 100(3), pp. 991-997.

Bandapalli, O. R., Schuessele, S., Kunz, J. B., Rausch, T., Stütz, A. M., Tal, N., Geron, I., Gershman, N., Izraeli, S., Eilers, J., Vaezipour, N., Kirschner-Schwabe, R., Hof, J., von Stackelberg, A., Schrappe, M., Stanulla, M., Zimmermann, M., Koehler, R., Avigad, S., Handgretinger, R., Frismantas, V., Bourquin, J. P., Bornhauser, B., Korbel, J. O., Muckenthaler, M. U. and Kulozik, A. E. (2014) 'The activating STAT5B N642H mutation is a common abnormality in pediatric T-cell acute lymphoblastic leukemia and confers a higher risk of relapse', *Haematologica,* 99(10), pp. e188-92.

Bandapalli, O. R., Zimmermann, M., Kox, C., Stanulla, M., Schrappe, M., Ludwig, W. D., Koehler, R., Muckenthaler, M. U. and Kulozik, A. E. (2013) 'NOTCH1 activation clinically antagonizes the unfavorable effect of PTEN inactivation in BFM-treated children with precursor T-cell acute lymphoblastic leukemia', *Haematologica,* 98(6), pp. 928-36.

Basso, G., Veltroni, M., Valsecchi, M. G., Dworzak, M. N., Ratei, R., Silvestri, D., Benetello, A., Buldini, B., Maglia, O., Masera, G., Conter, V., Arico, M., Biondi, A. and Gaipa, G. (2009) 'Risk of Relapse of Childhood Acute Lymphoblastic Leukemia Is Predicted By Flow Cytometric Measurement of Residual Disease on Day 15 Bone Marrow', *Journal of Clinical Oncology,* 27(31), pp. 5168-5174.

Belson, M., Kingsley, B. and Holmes, A. (2007) 'Risk Factors for Acute Leukemia in Children: A Review', *Environmental Health Perspectives,* 115(1), pp. 138-45.

Bene, M. C., Castoldi, G., Knapp, W., Ludwig, W. D., Matutes, E., Orfao, A. and van't Veer, M. B. (1995) 'Proposals for the immunological classification of acute leukemias. European Group for the Immunological Characterization of Leukemias (EGIL)', *Leukemia,* 9(10), pp. 1783-6.

Bennett, J. M., Catovsky, D., Daniel, M. T., Flandrin, G., Galton, D. A., Gralnick, H. R. and Sultan, C. (1981) 'The morphological classification of acute lymphoblastic leukaemia: concordance among observers and clinical correlations', *Br J Haematol,* 47(4), pp. 553-61.

Berry, D. A., Zhou, S., Higley, H., Mukundan, L., Fu, S., Reaman, G. H., Wood, B. L., Kelloff, G. J., Jessup, J. M. and Radich, J. P. (2017) 'Association of Minimal Residual Disease With Clinical Outcome in Pediatric and Adult Acute Lymphoblastic Leukemia: A Meta-analysis', *JAMA Oncology,* 3(7), pp. e170580-e170580.

Bhatla, T., Hogan, L., MD, Teachey, D. T., Bautista, F., Moppett, J. P., Velasco, P., Micalizzi, C., Rossig, C., Shukla, N. N., Gilad, G., Locatelli, F., BARUCHEL, A., Zwaan, M., Bezler, N. S., Rubio-San-Simón, A., Taussig, D., Raetz, E. A., Mao, Z. J., Wood, B., Alvarez Arias, D., Dr, Krevvata, M., Nnane, I., Bandyopadhyay, N., Lopez Solano, L., Dennis, R. M., Carson, R. and Vora, A. (2024) 'Daratumumab in Pediatric Relapsed/Refractory Acute Lymphoblastic Leukemia or Lymphoblastic Lymphoma: DELPHINUS Study', *Blood*.

Buchmann, S., Schrappe, M., Baruchel, A., Biondi, A., Borowitz, M., Campbell, M., Cario, G., Cazzaniga, G., Escherich, G., Harrison, C. J., Heyman, M., Hunger, S. P., Kiss, C., Liu, H.-C., Locatelli, F., Loh, M. L., Manabe, A., Mann, G., Pieters, R., Pui, C.-H., Rives, S., Schmiegelow, K., Silverman, L. B., Stary, J., Vora, A., Brown, P. and Consortium, o. b. o. t. P.-d.-L. (2022) 'Remission, treatment failure, and relapse in pediatric ALL: an international consensus of the Ponte-di-Legno Consortium', *Blood,* 139(12), pp. 1785-1793.

Burns, M. A., Place, A. E., Stevenson, K. E., Gutiérrez, A., Forrest, S., Pikman, Y., Vrooman, L. M., Harris, M. H., Hunt, S. K., O'Brien, J. E., Asselin, B. L., Athale, U. H., Clavell, L. A., Cole, P. D., Gennarini, L. M., Kahn, J. M., Kelly, K. M., Laverdiere, C., Leclerc, J. M., Michon, B., Schorin, M. A., Sulis, M. L., Welch, J. J. G., Neuberg, D. S., Sallan, S. E. and Silverman, L. B. (2021) 'Identification of prognostic factors in childhood T-cell acute lymphoblastic leukemia: Results from DFCI ALL Consortium Protocols 05-001 and 11-001', *Pediatr Blood Cancer,* 68(1), pp. e28719.

Campbell, M., Kiss, C., Zimmermann, M., Riccheri, C., Kowalczyk, J., Felice, M. S., Kuzmanovic, M., Kovacs, G., Kosmidis, H., Gonzalez, A., Bilic, E., Castillo, L., Kolenova, A., Jazbec, J., Popa, A., Konstantinov, D., Kappelmayer, J., Szczepanski, T., Dworzak, M., Buldini, B., Gaipa, G., Marinov, N., Rossi, J., Nagy, A., Gaspar, I., Stary, J. and Schrappe, M. (2023) 'Childhood Acute Lymphoblastic Leukemia: Results of the Randomized Acute Lymphoblastic Leukemia Intercontinental-Berlin-Frankfurt-Münster 2009 Trial', *Journal of Clinical Oncology,* 41(19), pp. 3499-3511.

Cancer-Research-UK (2021) *Acute lymphoblastic leukaemia (ALL) incidence statistics*. Available at: https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/leukaemia-all/incidence#heading-Zero (Accessed: 17/10/2023.

Cassidy, A., Myles, J. P., van Tongeren, M., Page, R. D., Liloglou, T., Duffy, S. W. and Field, J. K. (2008) 'The LLP risk model: an individual risk prediction model for lung cancer', *British Journal of Cancer,* 98(2), pp. 270-276.

Cazzaniga, G. and Biondi, A. (2005) 'Molecular monitoring of childhood acute lymphoblastic leukemia using antigen receptor gene rearrangements and quantitative polymerase chain reaction technology', *Haematologica,* 90(3), pp. 382-390.

Cazzaniga, G., Gaipa, G., Rossi, V. and Biondi, A. (2003) 'Minimal residual disease as a surrogate marker for risk assignment to ALL patients', *Rev Clin Exp Hematol,* 7(3), pp. 292-323.

Chandradevan, R., Aljudi, A. A., Drumheller, B. R., Kunananthaseelan, N., Amgad, M., Gutman, D. A., Cooper, L. A. D. and Jaye, D. L. (2020) 'Machine-based detection and classification for bone marrow aspirate differential counts: initial development focusing on nonneoplastic cells', *Laboratory Investigation,* 100(1), pp. 98-109.

Chang, J. H.-C., Poppe, M. M., Hua, C.-H., Marcus, K. J. and Esiashvili, N. (2021) 'Acute lymphoblastic leukemia', *Pediatric Blood & Cancer,* 68(S2), pp. e28371.

Chessells, J. M., Bailey, C. C. and Richards, S. M. (1995) 'Intensification of treatment and survival in all children with lymphoblastic leukaemia: results of UK Medical Research Council trial UKALL X', *The Lancet,* 345(8943), pp. 143-148.

Chessells, J. M., Veys, P., Kempski, H., Henley, P., Leiper, A., Webb, D. and Hann, I. M. (2003) 'Long-term follow-up of relapsed childhood acute lymphoblastic leukaemia', *British Journal of Haematology,* 123(3), pp. 396-405.

Chiaretti, S., Zini, G. and Bassan, R. (2014) 'Diagnosis and subclassification of acute lymphoblastic leukemia', *Mediterr J Hematol Infect Dis,* 6(1), pp. e2014073.

Chonghaile, T. N., Roderick, J. E., Glenfield, C., Ryan, J., Sallan, S. E., Silverman, L. B., Loh, M. L., Hunger, S. P., Wood, B., DeAngelo, D. J., Stone, R., Harris, M., Gutierrez, A., Kelliher, M. A. and Letai, A. (2014) 'Maturation stage of T-cell acute lymphoblastic leukemia determines BCL-2 versus BCL-XL dependence and sensitivity to ABT-199', *Cancer Discov,* 4(9), pp. 1074-87.

Chowdhury, M. Z. I. and Turin, T. C. (2020) 'Variable selection strategies and its importance in clinical prediction modelling', *Fam Med Community Health,* 8(1), pp. e000262.

Clappier, E., Collette, S., Grardel, N., Girard, S., Suarez, L., Brunie, G., Kaltenbach, S., Yakouben, K., Mazingue, F., Robert, A., Boutard, P., Plantaz, D., Rohrlich, P., van Vlierberghe, P., Preudhomme, C., Otten, J., Speleman, F., Dastugue, N., Suciu, S., Benoit, Y., Bertrand, Y., Cavé, H. and on behalf of the, E.-C. (2010) 'NOTCH1 and FBXW7 mutations have a favorable impact on early response to treatment, but not on outcome, in children with T-cell acute lymphoblastic leukemia (T-ALL) treated on EORTC trials 58881 and 58951', *Leukemia,* 24(12), pp. 2023-2031.

Cordo, V., van der Zwet, J. C. G., Canté-Barrett, K., Pieters, R. and Meijerink, J. P. P. (2021) 'T-cell Acute Lymphoblastic Leukemia: A Roadmap to Targeted Therapies', *Blood Cancer Discov,* 2(1), pp. 19-31.

Coustan-Smith, E., Mullighan, C. G., Onciu, M., Behm, F. G., Raimondi, S. C., Pei, D., Cheng, C., Su, X., Rubnitz, J. E., Basso, G., Biondi, A., Pui, C. H., Downing, J. R. and Campana, D. (2009) 'Early T-cell precursor leukaemia: a subtype of very high-risk acute lymphoblastic leukaemia', *Lancet Oncol,* 10(2), pp. 147-56.

Coustan-Smith, E., Sancho, J., Behm, F. G., Hancock, M. L., Razzouk, B. I., Ribeiro, R. C., Rivera, G. K., Rubnitz, J. E., Sandlund, J. T., Pui, C.-H. and Campana, D. (2002) 'Prognostic importance of measuring early clearance of leukemic cells by flow cytometry in childhood acute lymphoblastic leukemia', *Blood,* 100(1), pp. 52-58.

Craddock, M., Crockett, C., McWilliam, A., Price, G., Sperrin, M., van der Veer, S. N. and Faivre-Finn, C. (2022) 'Evaluation of Prognostic and Predictive Models in the Oncology Clinic', *Clinical Oncology,* 34(2), pp. 102-113.

Davidian, M. and Giltinan, D. M. (2003) 'Nonlinear Models for Repeated Measurement Data: An Overview and Update', *Journal of Agricultural, Biological, and Environmental Statistics,* 8(4), pp. 387-419.

de Glas, N. A., Bastiaannet, E., Engels, C. C., de Craen, A. J. M., Putter, H., van de Velde, C. J. H., Hurria, A., Liefers, G. J. and Portielje, J. E. A. (2016) 'Validity of the online PREDICT tool in older patients with breast cancer: a population-based study', *British Journal of Cancer,* 114(4), pp. 395-400.

DelRocco, N. J., Loh, M. L., Borowitz, M. J., Gupta, S., Rabin, K. R., Zweidler-McKay, P., Maloney, K. W., Mattano, L. A., Larsen, E., Angiolillo, A., Schore, R. J., Burke, M. J., Salzer, W. L., Wood, B. L., Carroll, A. J., Heerema, N. A., Reshmi, S. C., Gastier-Foster, J. M., Harvey, R., Chen, I. M., Roberts, K. G., Mullighan, C. G., Willman, C., Winick, N., Carroll, W. L., Rau, R. E., Teachey, D. T., Hunger, S. P., Raetz, E. A., Devidas, M. and Kairalla, J. A. (2024) 'Enhanced Risk Stratification for Children and Young Adults with B-Cell Acute Lymphoblastic Leukemia: A Children's Oncology Group Report', *Leukemia,* 38(4), pp. 720-728.

Down, S. K., Lucas, O., Benson, J. R. and Wishart, G. C. (2014) 'Effect of PREDICT on chemotherapy/trastuzumab recommendations in HER2-positive patients with early-stage breast cancer', *Oncol Lett,* 8(6), pp. 2757-2761.

Dunsmore, K. P., Winter, S. S., Devidas, M., Wood, B. L., Esiashvili, N., Chen, Z., Eisenberg, N., Briegel, N., Hayashi, R. J., Gastier-Foster, J. M., Carroll, A. J., Heerema, N. A., Asselin, B. L., Rabin, K. R., Zweidler-Mckay, P. A., Raetz, E. A., Loh, M. L., Schultz, K. R., Winick, N. J., Carroll, W. L. and Hunger, S. P. (2020) 'Children's Oncology Group AALL0434: A Phase III Randomized Clinical Trial Testing Nelarabine in Newly Diagnosed T-Cell Acute Lymphoblastic Leukemia', *J Clin Oncol,* 38(28), pp. 3282-3293.

Durinck, K., Goossens, S., Peirs, S., Wallaert, A., Van Loocke, W., Matthijssens, F., Pieters, T., Milani, G., Lammens, T., Rondou, P., Van Roy, N., De Moerloose, B., Benoit, Y., Haigh, J., Speleman, F., Poppe, B. and Van Vlierberghe, P. (2015) 'Novel biological insights in T-cell acute lymphoblastic leukemia', *Experimental Hematology,* 43(8), pp. 625-639.

Eden, O. B., Lilleyman, J. S., Richards, S., Shaw, M. P. and Peto, J. (1991) 'Results of Medical Research Council Childhood Leukaemia Trial UKALL VIII (Report to the Medical Research Council on behalf of the Working Party on Leukaemia in Childhood)', *British Journal of Haematology,* 78(2), pp. 187-196.

Enshaei, A., O'Connor, D., Bartram, J., Hancock, J., Harrison, C. J., Hough, R., Samarasinghe, S., den Boer, M. L., Boer, J. M., de Groot-Kruseman, H. A., Marquart, H. V., Noren-Nystrom, U., Schmiegelow, K., Schwab, C., Horstmann, M. A., Escherich, G., Heyman, M., Pieters, R., Vora, A., Moppett, J. and Moorman, A. V. (2020) 'A validated novel continuous prognostic index to deliver stratified medicine in pediatric acute lymphoblastic leukemia', *Blood,* 135(17), pp. 1438-1446.

Ferrando, A. (2010) 'NOTCH mutations as prognostic markers in T-ALL', *Leukemia,* 24(12), pp. 2003-2004.

Ferrando, A. A., Neuberg, D. S., Staunton, J., Loh, M. L., Huard, C., Raimondi, S. C., Behm, F. G., Pui, C.-H., Downing, J. R., Gilliland, D. G., Lander, E. S., Golub, T. R. and Look, A. T. (2002) 'Gene expression signatures define novel oncogenic pathways in T cell acute lymphoblastic leukemia', *Cancer Cell,* 1(1), pp. 75-87.

Field, J. K., Vulkan, D., Davies, M. P. A., Duffy, S. W. and Gabe, R. (2021) 'Liverpool Lung Project lung cancer risk stratification model: calibration and prospective validation', *Thorax,* 76(2), pp. 161-168.

Ford, A. M., Colman, S. and Greaves, M. (2023) 'Covert pre-leukaemic clones in healthy co-twins of patients with childhood acute lymphoblastic leukaemia', *Leukemia,* 37(1), pp. 47-52.

Gao, Y., Zhou, J. F., Mao, J. Y., Jiang, L. and Li, X. P. (2022) 'Identification of the Thyrotropin-Releasing Hormone (TRH) as a Novel Biomarker in the Prognosis for Acute Myeloid Leukemia', *Biomolecules,* 12(10).

Gaynon, P. S., Desai, A. A., Bostrom, B. C., Hutchinson, R. J., Lange, B. J., Nachman, J. B., Reaman, G. H., Sather, H. N., Steinherz, P. G., Trigg, M. E., Tubergen, D. G. and Uckun, F. M. (1997) 'Early response to therapy and outcome in childhood acute lymphoblastic leukemia: a review', *Cancer,* 80(9), pp. 1717-26.

GeneCards (2023a) *NKX2-1 Gene - NK2 Homeobox 1*. Available at: https://www.genecards.org/cgi-bin/carddisp.pl?gene=NKX2-1&keywords=nkx21 (Accessed: 15th September 2023.

GeneCards (2023b) *TLX1 Gene - T Cell Leukemia Homeobox 1*. Available at: https://www.genecards.org/cgi-bin/carddisp.pl?gene=TLX1&keywords=tlx1 (Accessed: 15th September 2023.

GeneCards (2023c) *TLX3 Gene - T Cell Leukemia Homeobox 3*. Available at: https://www.genecards.org/cgi-bin/carddisp.pl?gene=TLX3&keywords=tlx3 (Accessed: 15th September 2023.

Genescà, E. and la Starza, R. (2022) 'Early T-Cell Precursor ALL and Beyond: Immature and Ambiguous Lineage T-ALL Subsets', *Cancers,* 14(8), pp. 1873.

Girardi, T., Vicente, C., Cools, J. and De Keersmaecker, K. (2017) 'The genetics and molecular biology of T-ALL', *Blood,* 129(9), pp. 1113-1123.

Goh, W. W. B., Wang, W. and Wong, L. (2017) 'Why Batch Effects Matter in Omics Data, and How to Avoid Them', *Trends Biotechnol,* 35(6), pp. 498-507.

Gossai, N. P., Devidas, M., Chen, Z., Wood, B. L., Zweidler-McKay, P. A., Rabin, K. R., Loh, M. L., Raetz, E. A., Winick, N. J., Burke, M. J., Carroll, A. J., Esiashvili, N., Heerema, N. A., Carroll, W. L., Hunger, S. P., Dunsmore, K. P., Winter, S. S. and Teachey, D. T. (2023) 'Central nervous system status is prognostic in T-cell acute lymphoblastic leukemia: a Children's Oncology Group report', *Blood,* 141(15), pp. 1802-1811.

Gottardo, N. G., Jacoby, P. A., Sather, H. N., Reaman, G. H., Baker, D. L. and Kees, U. R. (2005) 'Significance of HOX11L2/TLX3 expression in children with T-cell acute lymphoblastic leukemia treated on Children's Cancer Group protocols', *Leukemia,* 19(9), pp. 1705-1708.

Goulden, N. (2013) *UKALL 2011 Trial - United Kingdom National Randomised Trial For Children and Young Adults with Acute Lymphoblastic Leukaemia and Lymphoma 2011* Available at: https://www.northerncanceralliance.nhs.uk/wp-content/uploads/2019/01/UKALL2011-Protocol-v3.0-01-Oct-2013.pdf (Accessed: 18th September 2023.

Goulden, N. J., Kirkwood, A. A., Moppett, J., Samarasinghe, S., Lawson, S., Rowntree, C., Hough, R., Athwal, R., Kearns, P. R. and Vora, A. (2017) 'UKALL 2011: Randomised Trial Investigating a Short Induction Dexamethasone Schedule for Children and Young Adults with Acute Lymphoblastic Leukaemia', *Blood,* 130(Supplement 1), pp. 141-141.

Griffin, T. C., Shuster, J. J., Buchanan, G. R., Murphy, S. B., Camitta, B. M. and Amylon, M. D. (2000) 'Slow disappearance of peripheral blood blasts is an adverse prognostic factor in childhood T cell acute lymphoblastic leukemia: a Pediatric Oncology Group study', *Leukemia,* 14(5), pp. 792-5.

Grotel, M., Meijerink, J., Beverloo, H., Langerak, A., Buys-Gladdines, J., Schneider, P., Poulsen, T., Boer, M., Horstmann, M., Kamps, W., Veerman, A., Wering, E., Noesel, M. and Pieters, R. (2006) 'The outcome of molecular-cytogenetic subgroups in pediatric T-cell acute lymphoblastic leukemia: A retrospective study of patients treated according to DCOG or COALL protocols', *Haematologica,* 91, pp. 1212-21.

Gu, A., Schmidt, B., Lonsdale, A., Jalaldeen, R., Kosasih, H. J., Brown, L. M., Sadras, T., Ekert, P. G. and Oshlack, A. (2023) 'TALLSorts: a T-cell acute lymphoblastic leukemia subtype classifier using RNA-seq expression data', *Blood Advances,* 7(24), pp. 7402-7406.

Gunsilius, E., Gastl, G. and Petzer, A. L. (2001) 'Hematopoietic stem cells', *Biomedicine & Pharmacotherapy,* 55(4), pp. 186-194.

Gupta, S., Teachey, D. T., Chen, Z., Rabin, K. R., Dunsmore, K. P., Larsen, E. C., Maloney, K. W., Mattano Jr, L. A., Winter, S. S., Carroll, A. J., Heerema, N. A., Borowitz, M. J., Wood, B. L., Carroll, W. L., Raetz, E. A., Winick, N. J., Loh, M. L., Hunger, S. P. and Devidas, M. (2022) 'Sex-based disparities in outcome in pediatric acute lymphoblastic leukemia: a Children's Oncology Group report', *Cancer,* 128(9), pp. 1863-1870.

Guru Murthy, G. S., Pondaiah, S. K., Abedin, S. and Atallah, E. (2019) 'Incidence and survival of T-cell acute lymphoblastic leukemia in the United States', *Leuk Lymphoma,* 60(5), pp. 1171-1178.

Gutierrez, A., Sanda, T., Grebliunaite, R., Carracedo, A., Salmena, L., Ahn, Y., Dahlberg, S., Neuberg, D., Moreau, L. A., Winter, S. S., Larson, R., Zhang, J., Protopopov, A., Chin, L., Pandolfi, P. P., Silverman, L. B., Hunger, S. P., Sallan, S. E. and Look, A. T. (2009) 'High frequency of PTEN, PI3K, and AKT abnormalities in T-cell acute lymphoblastic leukemia', *Blood,* 114(3), pp. 647-650.

Halabi, S., Li, C. and Luo, S. (2019) 'Developing and Validating Risk Assessment Models of Clinical Outcomes in Modern Oncology', *JCO Precis Oncol,* 3.

Halabi, S. and Owzar, K. (2010) 'The Importance of Identifying and Validating Prognostic Factors in Oncology', *Seminars in Oncology,* 37(2), pp. e9-e18.

Hann, I., Vora, A., Richards, S., Hill, F., Gibson, B., Lilleyman, J., Kinsey, S., Mitchell, C., Eden, O. B. and on behalf of the, U. K. M. R. C. s. W. P. o. C. L. (2000) 'Benefit of intensified treatment for all children with acute lymphoblastic leukaemia: results from MRC UKALL XI and MRC ALL97 randomised trials', *Leukemia,* 14(3), pp. 356-363.

Hann, I. M., EVANS, D. I. K., PALMER, M. K., JONES, P. J. M. and HAWORTH, C. (1979) 'The prognostic significance of morphological features in childhood acute lymphoblastic leukaemia', *Clinical & Laboratory Haematology,* 1(3), pp. 215-226.

Harris, J. K. (2021) 'Primer on binary logistic regression', *Fam Med Community Health,* 9(Suppl 1).

Harris, N. L., Jaffe, E. S., Diebold, J., Flandrin, G., Muller-Hermelink, H. K., Vardiman, J., Lister, T. A. and Bloomfield, C. D. (1999) 'World Health Organization classification of neoplastic diseases of the hematopoietic and lymphoid tissues: report of the Clinical Advisory Committee meeting-Airlie House, Virginia, November 1997', *J Clin Oncol,* 17(12), pp. 3835-49.

Harrison, C. J. and Johansson, B. (2015) 'Acute lymphoblastic leukemia', *Cancer Cytogenetics*, pp. 198-251.

Hayashi, H., Makimoto, A. and Yuza, Y. (2024) 'Treatment of Pediatric Acute Lymphoblastic Leukemia: A Historical Perspective', *Cancers,* 16(4), pp. 723.

Hill, F. G. H., Richards, S., Gibson, B., Hann, I., Lilleyman, J., Kinsey, S., Mitchell, C., Harrison, C. J., Eden, O. B. and Leukaemia, t. U. M. R. C. W. P. o. C. (2004) 'Successful treatment without cranial radiotherapy of children receiving intensified chemotherapy for acute lymphoblastic leukaemia: results of the risk-stratified randomized central nervous system treatment trial MRC UKALL XI (ISRC TN 16757172)', *British Journal of Haematology,* 124(1), pp. 33-46.

Hjalgrim, L. L., Rostgaard, K., Schmiegelow, K., Söderhäll, S., Kolmannskog, S., Vettenranta, K., Kristinsson, J., Clausen, N., Melbye, M., Hjalgrim, H. and Gustafsson, G. r. (2003) 'Age- and Sex-Specific Incidence of Childhood Leukemia by Immunophenotype in the Nordic Countries', *JNCI: Journal of the National Cancer Institute,* 95(20), pp. 1539-1544.

Homminga, I., Pieters, R., Langerak, A. W., de Rooi, J. J., Stubbs, A., Verstegen, M., Vuerhard, M., Buijs-Gladdines, J., Kooi, C., Klous, P., van Vlierberghe, P., Ferrando, A. A., Cayuela, J. M., Verhaaf, B., Beverloo, H. B., Horstmann, M., de Haas, V., Wiekmeijer, A. S., Pike-Overzet, K., Staal, F. J., de Laat, W., Soulier, J., Sigaux, F. and Meijerink, J. P. (2011) 'Integrated transcript and genome analyses reveal NKX2-1 and MEF2C as potential oncogenes in T cell acute lymphoblastic leukemia', *Cancer Cell,* 19(4), pp. 484-97.

Homminga, I., Vuerhard, M. J., Langerak, A. W., Buijs-Gladdines, J., Pieters, R. and Meijerink, J. P. (2012) 'Characterization of a pediatric T-cell acute lymphoblastic leukemia patient with simultaneous LYL1 and LMO2 rearrangements', *Haematologica,* 97(2), pp. 258-61.

Hu, S., Li, S., Ning, W., Huang, X., Liu, X., Deng, Y., Franceschi, D., Ogbuehi, A., Lethaus, B., Savkovic, V., Li, H., Gaus, S., Zimmerer, R., Ziebolz, D., Schmalz, G. and Huang, S. (2022) 'Identifying crosstalk genetic biomarkers linking a neurodegenerative disease, Parkinson's disease, and periodontitis using integrated bioinformatics analyses', *Frontiers in Aging Neuroscience,* 14.

Huang, F., Guang, P., Li, F., Liu, X., Zhang, W. and Huang, W. (2020) 'AML, ALL, and CML classification and diagnosis based on bone marrow cell morphology combined with convolutional neural network: A STARD compliant diagnosis research', *Medicine,* 99(45), pp. e23154.

Hunger, S. P., Lu, X., Devidas, M., Camitta, B. M., Gaynon, P. S., Winick, N. J., Reaman, G. H. and Carroll, W. L. (2012) 'Improved survival for children and adolescents with acute lymphoblastic leukemia between 1990 and 2005: a report from the children's oncology group', *J Clin Oncol,* 30(14), pp. 1663-9.

Ibrahim, J. G., Chu, H. and Chen, M.-H. (2012) 'Missing Data in Clinical Studies: Issues and Methods', *Journal of Clinical Oncology,* 30(26), pp. 3297-3303.

Inukai, T., Kiyokawa, N., Campana, D., Coustan-Smith, E., Kikuchi, A., Kobayashi, M., Takahashi, H., Koh, K., Manabe, A., Kumagai, M., Ikuta, K., Hayashi, Y., Tsuchida, M., Sugita, K. and Ohara, A. (2012) 'Clinical significance of early T-cell precursor acute lymphoblastic leukaemia: results of the Tokyo Children's Cancer Study Group Study L99-15', *British Journal of Haematology,* 156(3), pp. 358-365.

Jain, M. (2011) 'Next-generation sequencing technologies for gene expression profiling in plants', *Briefings in Functional Genomics,* 11(1), pp. 63-70.

Jenkins, D. A., Sperrin, M., Martin, G. P. and Peek, N. (2018) 'Dynamic models to predict health outcomes: current status and methodological challenges', *Diagnostic and Prognostic Research,* 2(1), pp. 23.

Jenkinson, S., Kirkwood, A. A., Goulden, N., Vora, A., Linch, D. C. and Gale, R. E. (2016) 'Impact of PTEN abnormalities on outcome in pediatric patients with T-cell acute lymphoblastic leukemia treated on the MRC UKALL2003 trial', *Leukemia,* 30(1), pp. 39-47.

Jenkinson, S., Koo, K., Mansour, M. R., Goulden, N., Vora, A., Mitchell, C., Wade, R., Richards, S., Hancock, J., Moorman, A. V., Linch, D. C. and Gale, R. E. (2013) 'Impact of NOTCH1/FBXW7 mutations on outcome in pediatric T-cell acute lymphoblastic leukemia patients treated on the MRC UKALL 2003 trial', *Leukemia,* 27(1), pp. 41-47.

Karrman, K. and Johansson, B. (2017) 'Pediatric T-cell acute lymphoblastic leukemia', *Genes, Chromosomes and Cancer,* 56(2), pp. 89-116.

Kashef, A., Khatibi, T. and Mehrvar, A. (2020) 'Prediction of Cranial Radiotherapy Treatment in Pediatric Acute Lymphoblastic Leukemia Patients Using Machine Learning: A Case Study at MAHAK Hospital', *Asian Pac J Cancer Prev,* 21(11), pp. 3211-3219.

Kato, M. and Manabe, A. (2018) 'Treatment and biology of pediatric acute lymphoblastic leukemia', *Pediatr Int,* 60(1), pp. 4-12.

Katz, A. J., Chia, V. M., Schoonen, W. M. and Kelsh, M. A. (2015) 'Acute lymphoblastic leukemia: an assessment of international incidence, survival, and disease burden', *Cancer Causes & Control,* 26(11), pp. 1627-1642.

Kees, U. R., Heerema, N. A., Kumar, R., Watt, P. M., Baker, D. L., La, M. K., Uckun, F. M. and Sather, H. N. (2003) 'Expression of HOX11 in childhood T-lineage acute lymphoblastic leukaemia can occur in the absence of cytogenetic aberration at 10q24: a study from the Children's Cancer Group (CCG)', *Leukemia,* 17(5), pp. 887-893.

Kirkwood, A. A., Goulden, N., Moppett, J., Samarasinghe, S., Mee, J., Hough, R., Kearns, P. R., Lawson, S., Rowntree, C. J. and Vora, A. (2022) 'High Dose Methotrexate Does Not Reduce the Risk of CNS Relapse in Children and Young Adults with Acute Lymphoblastic Leukaemia and Lymphoblastic Lymphoma. Results of the Randomised Phase III Study UKALL 2011', *Blood,* 140(Supplement 1), pp. 516-518.

Krieger, D., Moericke, A., Oschlies, I., Zimmermann, M., Schrappe, M., Reiter, A. and Burkhardt, B. (2010) 'Frequency and clinical relevance of DNA microsatellite alterations of the CDKN2A/B, ATM and p53 gene loci: a comparison between pediatric precursor T-cell lymphoblastic lymphoma and T-cell lymphoblastic leukemia', *Haematologica,* 95(1), pp. 158-62.

Kumar, B. V., Connors, T. J. and Farber, D. L. (2018) 'Human T Cell Development, Localization, and Function throughout Life', *Immunity,* 48(2), pp. 202-213.

Lauten, M., Möricke, A., Beier, R., Zimmermann, M., Stanulla, M., Meissner, B., Odenwald, E., Attarbaschi, A., Niemeyer, C., Niggli, F., Riehm, H. and Schrappe, M. (2012) 'Prediction of outcome by early bone marrow response in childhood acute lymphoblastic leukemia treated in the ALL-BFM 95 trial: differential effects in precursor B-cell and T-cell leukemia', *Haematologica,* 97(7), pp. 1048-56.

Ledent, V. and Vervoort, M. (2001) 'The basic helix-loop-helix protein family: comparative genomics and phylogenetic analysis', *Genome Res,* 11(5), pp. 754-70.

Lee, J. W. and Cho, B. (2017) 'Prognostic factors and treatment of pediatric acute lymphoblastic leukemia', *Korean J Pediatr,* 60(5), pp. 129-137.

Liu, X., Zou, Y., Chen, X., Wang, S., Guo, Y., Yang, W., Zhang, L., Chen, Y., Zhang, Y. and Zhu, X. (2020) 'Minimal residual disease surveillance at day 90 predicts long-term survival in pediatric patients with T-cell acute lymphoblastic leukemia', *Leukemia & Lymphoma,* 61(14), pp. 3460-3467.

Liu, Y., Easton, J., Shao, Y., Maciaszek, J., Wang, Z., Wilkinson, M. R., McCastlain, K., Edmonson, M., Pounds, S. B., Shi, L., Zhou, X., Ma, X., Sioson, E., Li, Y., Rusch, M., Gupta, P., Pei, D., Cheng, C., Smith, M. A., Auvil, J. G., Gerhard, D. S., Relling, M. V., Winick, N. J., Carroll, A. J., Heerema, N. A., Raetz, E., Devidas, M., Willman, C. L., Harvey, R. C., Carroll, W. L., Dunsmore, K. P., Winter, S. S., Wood, B. L., Sorrentino, B. P., Downing, J. R., Loh, M. L., Hunger, S. P., Zhang, J. and Mullighan, C. G. (2017a) 'The genomic landscape of pediatric and young adult T-lineage acute lymphoblastic leukemia', *Nat Genet,* 49(8), pp. 1211-1218.

Liu, Y., Easton, J., Shao, Y., Maciaszek, J., Wang, Z., Wilkinson, M. R., McCastlain, K., Edmonson, M., Pounds, S. B., Shi, L., Zhou, X., Ma, X., Sioson, E., Li, Y., Rusch, M., Gupta, P., Pei, D., Cheng, C., Smith, M. A., Auvil, J. G., Gerhard, D. S., Relling, M. V., Winick, N. J., Carroll, A. J., Heerema, N. A., Raetz, E., Devidas, M., Willman, C. L., Harvey, R. C., Carroll, W. L., Dunsmore, K. P., Winter, S. S., Wood, B. L., Sorrentino, B. P., Downing, J. R., Loh, M. L., Hunger, S. P., Zhang, J. and Mullighan, C. G. (2017b) 'The genomic landscape of pediatric and young adult T-lineage acute lymphoblastic leukemia', *Nature Genetics,* 49(8), pp. 1211-1218.

Love, M. I., Anders, S. and Huber, W. (2024) *Analyzing RNA-seq data with DESeq2*: Bioconductor. Available at: https://www.bioconductor.org/packages/release/bioc/vignettes/DESeq2/inst/doc/DESeq2.html (Accessed: 28/05/2024 2024).

Love, M. I., Huber, W. and Anders, S. (2014) 'Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2', *Genome Biol,* 15(12), pp. 550.

Malard, F. and Mohty, M. (2020) 'Acute lymphoblastic leukaemia', *Lancet,* 395(10230), pp. 1146-1162.

Marcotte, E. L., Spector, L. G., Mendes-de-Almeida, D. P. and Nelson, H. H. (2021) 'The Prenatal Origin of Childhood Leukemia: Potential Applications for Epidemiology and Newborn Screening', *Front Pediatr,* 9, pp. 639479.

Melchior, L., Anja, M., Rita, B., Martin, Z., Martin, S., Barbara, M., Edelgard, O., Andishe, A., Charlotte, N., Felix, N., Hansjörg, R. and Martin, S. (2012) 'Prediction of outcome by early bone marrow response in childhood acute lymphoblastic leukemia treated in the ALL-BFM 95 trial: differential effects in precursor B-cell and T-cell leukemia', *Haematologica,* 97(7), pp. 1048-1056.

Mendes, R. D., Sarmento, L. M., Canté-Barrett, K., Zuurbier, L., Buijs-Gladdines, J. G. C. A. M., Póvoa, V., Smits, W. K., Abecasis, M., Yunes, J. A., Sonneveld, E., Horstmann, M. A., Pieters, R., Barata, J. T. and Meijerink, J. P. P. (2014) 'PTEN microdeletions in T-cell acute lymphoblastic leukemia are caused by illegitimate RAG-mediated recombination events', *Blood,* 124(4), pp. 567-578.

Miller, D. R., Leikin, S., Albo, V., Sather, H., Karon, M. and Hammond, D. (1983) 'Prognostic factors and therapy in acute lymphoblastic leukemia of childhood: CCG-141: A report from childrens cancer study group', *Cancer,* 51(6), pp. 1041-1049.

Miller, D. R., Leikin, S., Albo, V., Vitale, L., Sather, H., Coccia, P., Nesbit, M., Karon, M. and Hammond, D. (1980) 'Use of prognostic factors in improving the design and efficiency of clinical trials in childhood leukemia: Children's Cancer Study Group Report', *Cancer treatment reports,* 64(2-3), pp. 381-392.

Mirji, G., Bhat, J., Kode, J., Banavali, S., Sengar, M., Khadke, P., Sait, O. and Chiplunkar, S. (2016) 'Risk stratification of T-cell Acute Lymphoblastic Leukemia patients based on gene expression, mutations and copy number variation', *Leukemia Research,* 45, pp. 33-39.

Mitchell, C., Payne, J., Wade, R., Vora, A., Kinsey, S., Richards, S. and Eden, T. (2009) 'The impact of risk stratification by early bone-marrow response in childhood lymphoblastic leukaemia: results from the United Kingdom Medical Research Council trial ALL97 and ALL97/99', *Br J Haematol,* 146(4), pp. 424-36.

Mitchell, C., Richards, S., Harrison, C. J. and Eden, T. (2010) 'Long-term follow-up of the United Kingdom medical research council protocols for childhood acute lymphoblastic leukaemia, 1980-2001', *Leukemia,* 24(2), pp. 406-18.

Mitchell, C. D., Richards, S. M., Kinsey, S. E., Lilleyman, J., Vora, A., Eden, T. O. and Party, M. R. C. C. L. W. (2005) 'Benefit of dexamethasone compared with prednisolone for childhood acute lymphoblastic leukaemia: results of the UK Medical Research Council ALL97 randomized trial', *British journal of haematology,* 129(6), pp. 734-745.

Molina, O., Bataller, A., Thampi, N., Ribera, J., Granada, I., Velasco, P., Fuster, J. L. and Menéndez, P. (2022) 'Near-Haploidy and Low-Hypodiploidy in B-Cell Acute Lymphoblastic Leukemia: When Less Is Too Much', *Cancers,* 14(1), pp. 32.

Monti, S., Tamayo, P., Mesirov, J. and Golub, T. (2003) 'Consensus Clustering: A Resampling-Based Method for Class Discovery and Visualization of Gene Expression Microarray Data', *Machine Learning,* 52(1), pp. 91-118.

Moorman, A. V. (2016) 'New and emerging prognostic and predictive genetic biomarkers in B-cell precursor acute lymphoblastic leukemia', *Haematologica,* 101(4), pp. 407-416.

Moorman, A. V. 'Advances in Risk Stratification in Pediatric Acute Lymphoblastic Leukemia', *EHA*.

Moreira, D. C., González-Ramella, O., Echavarría Valenzuela, M., Carrillo, A. K., Faughnan, L., Job, G., Chen, Y., Villegas, C., Ellis Irigoyen, A., Barra Urbays, R., Ramírez Martinez, M., Altamirano Alvarez, E., León Espitia, J. A., López Facundo, N. A., Colunga Pedraza, J. E., Reyes Gutierrez, F. M., Aguilar Román, A. B., Tamez Gómez, E. L., Portillo Zavala, C. S., Negroe Ocampo, N. D. C., Pulido Sanchez, S. G., Cortés Alva, D., Casillas Toral, P., Salas Villa, K., Mendoza Sánchez, P. J., Pérez Alvarado, C., Tamayo Pedraza, G., González Zamorano, M., Ávila Alba, J. M. R., Becerril Becerril, J., Ramírez Durán, H., Sandoval Cabrera, A., Pineda Gordillo, A., de la Rosa Alonso, D. I., Mejía Marín, L. J., Benítez Can, L., Gutiérrez Martinez, I., Jiménez Osorio, M. I., Echeandia, N., Casillas, E., Guerrero-Gomez, K., Devidas, M. and Friedrich, P. (2023) 'Evaluation of factors leading to poor outcomes for pediatric acute lymphoblastic leukemia in Mexico: a multi-institutional report of 2,116 patients', *Front Oncol,* 13, pp. 1255555.

Möricke, A., Lauten, M., Beier, R., Odenwald, E., Stanulla, M., Zimmermann, M., Attarbaschi, A., Niggli, F. and Schrappe, M. (2013) 'Prediction of outcome by early response in childhood acute lymphoblastic leukemia', *Klin Padiatr,* 225 Suppl 1, pp. S50-6.

Moro, L., Simoneschi, D., Kurz, E., Arbini, A. A., Jang, S., Guaragnella, N., Giannattasio, S., Wang, W., Chen, Y.-A., Pires, G., Dang, A., Hernandez, E., Kapur, P., Mishra, A., Tsirigos, A., Miller, G., Hsieh, J.-T. and Pagano, M. (2020) 'Epigenetic silencing of the ubiquitin ligase subunit FBXL7 impairs c-SRC degradation and promotes epithelial-to-mesenchymal transition and metastasis', *Nature Cell Biology,* 22(9), pp. 1130-1142.

Mroczek, A., Zawitkowska, J., Kowalczyk, J. and Lejman, M. (2021) 'Comprehensive Overview of Gene Rearrangements in Childhood T-Cell Acute Lymphoblastic Leukaemia', *Int J Mol Sci,* 22(2).

Mullighan, C. G., Zhang, J., Harvey, R. C., Collins-Underwood, J. R., Schulman, B. A., Phillips, L. A., Tasian, S. K., Loh, M. L., Su, X., Liu, W., Devidas, M., Atlas, S. R., Chen, I.-M., Clifford, R. J., Gerhard, D. S., Carroll, W. L., Reaman, G. H., Smith, M., Downing, J. R., Hunger, S. P. and Willman, C. L. (2009) 'JAK mutations in high-risk childhood acute lymphoblastic leukemia', *Proceedings of the National Academy of Sciences,* 106(23), pp. 9414-9418.

Murti, D. M. P., Pujianto, U., Wibawa, A. P. and Akbar, M. I. 'K-Nearest Neighbor (K-NN) based Missing Data Imputation'. *2019 5th International Conference on Science in Information Technology (ICSITech)*, 23-24 Oct. 2019, 83-88.

Natarajan, V., Bandapalli, O. R., Rajkumar, T., Sagar, T. G. and Karunakaran, N. (2015) 'NOTCH1 and FBXW7 Mutations Favor Better Outcome in Pediatric South Indian T-Cell Acute Lymphoblastic Leukemia', *Journal of Pediatric Hematology/Oncology,* 37(1).

NHS (2022) *Standard protocol prepared for the Targeted Lung Health Checks Programme*. Available at: https://www.england.nhs.uk/wp-content/uploads/2019/02/B1646-standard-protocol-targeted-lung-health-checks-programme-v2.pdf (Accessed: 10/10/2024 2024).

Noronha, E. P., Marques, L. V. C., Andrade, F. G., Thuler, L. C. S., Terra-Granado, E., Pombo-de-Oliveira, M. S., , t. B. C. S. G. o. A. L., da Paz Zampier, C., da Conceição Barbosa, T., Chagas Neto, P., Dallapicola Brisson, G., dos Santos Bueno, F. V., Cezar Sardou, I., Gonçalves Aguiar, B., Silva Dias, A. C., Pedral Sampaio, G., Antônio Gomes Oliveira, R., de Oliveira, C. T., Casagranda, C., Ramos Vera, G., Ribeiro Neves, G., Maria Quezado Magalhães, I., Carlos Cordoba, J., Teixeira Costa, J., de Brito, P. C., Ferreira Marques, R., Pereira, R., Guedes, R. and Epelman, S. (2021) 'Corrigendum: The Profile of Immunophenotype and Genotype Aberrations in Subsets of Pediatric T-Cell Acute Lymphoblastic Leukemia', *Frontiers in Oncology,* 10.

O'Connor, D., Demeulemeester, J., Conde, L., Kirkwood, A., Fung, K., Papaleonidopoulou, F., Bloye, G., Farah, N., Rahman, S., Hancock, J., Bateman, C., Inglott, S., Mee, J., Herrero, J., Van Loo, P., Moorman, A. V., Vora, A. and Mansour, M. R. (2023) 'The Clinicogenomic Landscape of Induction Failure in Childhood and Young Adult T-Cell Acute Lymphoblastic Leukemia', *J Clin Oncol,* 41(19), pp. 3545-3556.

Paganin, M., Grillo, M. F., Silvestri, D., Scapinello, G., Buldini, B., Cazzaniga, G., Biondi, A., Valsecchi, M. G., Conter, V., te Kronnie, G. and Basso, G. (2018) 'The presence of mutated and deleted PTEN is associated with an increased risk of relapse in childhood T cell acute lymphoblastic leukaemia treated with AIEOP-BFM ALL protocols', *British Journal of Haematology,* 182(5), pp. 705-711.

Pan, L., Liu, G., Lin, F., Zhong, S., Xia, H., Sun, X. and Liang, H. (2017) 'Machine learning applications for prediction of relapse in childhood acute lymphoblastic leukemia', *Sci Rep,* 7(1), pp. 7402.

Paola, B., Judith, L.-P., Jean Michel, C., Vahid, A., Myriam, L., Virginie, G., Yves, P., Gérard, M., Thierry, L., Claudine, S., Sylvie, F., Anne, H., François, S., Marie Françoise, A., Luc, D., Guy, L. and André, B. (2008) 'Impact of genotype on survival of children with T-cell acute lymphoblastic leukemia treated according to the French protocol FRALLE-93: the effect of TLX3/HOX11L2 gene expression on outcome', *Haematologica,* 93(11), pp. 1658-1665.

Park, M. J., Taki, T., Oda, M., Watanabe, T., Yumura-Yagi, K., Kobayashi, R., Suzuki, N., Hara, J., Horibe, K. and Hayashi, Y. (2009) 'FBXW7 and NOTCH1 mutations in childhood T cell acute lymphoblastic leukaemia and T cell non-Hodgkin lymphoma', *Br J Haematol,* 145(2), pp. 198-206.

Patrick, K., Wade, R., Goulden, N., Mitchell, C., Moorman, A. V., Rowntree, C., Jenkinson, S., Hough, R. and Vora, A. (2014a) 'Outcome for children and young people with Early T-cell precursor acute lymphoblastic leukaemia treated on a contemporary protocol, UKALL 2003', *Br J Haematol,* 166(3), pp. 421-4.

Patrick, K., Wade, R., Goulden, N., Mitchell, C., Rowntree, C., Hancock, J., Hough, R. E. and Vora, A. J. (2014b) 'Improved Outcome for Children and Young People with T-Acute Lymphoblastic Leukaemia: Results of the UKALL 2003 Trial', *Blood,* 124(21), pp. 3702-3702.

Paulsson, K. and Johansson, B. (2009) 'High hyperdiploid childhood acute lymphoblastic leukemia', *Genes, Chromosomes and Cancer,* 48(8), pp. 637-660.

Petit, A., Trinquand, A., Chevret, S., Ballerini, P., Cayuela, J.-M., Grardel, N., Touzart, A., Brethon, B., Lapillonne, H., Schmitt, C., Thouvenin, S., Michel, G., Preudhomme, C., Soulier, J., Landman-Parker, J., Leverger, G., Macintyre, E., Baruchel, A., Asnafi, V. and Group, o. b. o. t. F. A. L. L. S. (2018) 'Oncogenetic mutations combined with MRD improve outcome prediction in pediatric T-cell acute lymphoblastic leukemia', *Blood,* 131(3), pp. 289-300.

Pieters, R., de Groot-Kruseman, H., Van der Velden, V., Fiocco, M., van den Berg, H., de Bont, E., Egeler, R. M., Hoogerbrugge, P., Kaspers, G., Van der Schoot, E., De Haas, V. and Van Dongen, J. (2016) 'Successful Therapy Reduction and Intensification for Childhood Acute Lymphoblastic Leukemia Based on Minimal Residual Disease Monitoring: Study ALL10 From the Dutch Childhood Oncology Group', *J Clin Oncol,* 34(22), pp. 2591-601.

Pocock, R., Farah, N., Richardson, S. E. and Mansour, M. R. (2021) 'Current and emerging therapeutic approaches for T-cell acute lymphoblastic leukaemia', *British Journal of Haematology,* 194(1), pp. 28-43.

Pölönen, P., Di Giacomo, D., Seffernick, A. E., Elsayed, A., Kimura, S., Benini, F., Montefiori, L. E., Wood, B. L., Xu, J., Chen, C., Cheng, Z., Newman, H., Myers, J., Iacobucci, I., Li, E., Sussman, J., Hedges, D., Hui, Y., Diorio, C., Uppuluri, L., Frank, D., Fan, Y., Chang, Y., Meshinchi, S., Ries, R., Shraim, R., Li, A., Bernt, K. M., Devidas, M., Winter, S. S., Dunsmore, K. P., Inaba, H., Carroll, W. L., Ramirez, N. C., Phillips, A. H., Kriwacki, R. W., Yang, J. J., Vincent, T. L., Zhao, Y., Ghate, P. S., Wang, J., Reilly, C., Zhou, X., Sanders, M. A., Takita, J., Kato, M., Takasugi, N., Chang, B. H., Press, R. D., Loh, M., Rampersaud, E., Raetz, E., Hunger, S. P., Tan, K., Chang, T.-C., Wu, G., Pounds, S. B., Mullighan, C. G. and Teachey, D. T. (2024) 'The genomic basis of childhood T-lineage acute lymphoblastic leukaemia', *Nature,* 632(8027), pp. 1082-1091.

Popov, A., Henze, G., Roumiantseva, J., Budanov, O., Belevtsev, M., Verzhbitskaya, T., Boyakova, E., Movchan, L., Tsaur, G., Fadeeva, M., Lagoyko, S., Zharikova, L., Miakova, N., Litvinov, D., Khlebnikova, O., Streneva, O., Stolyarova, E., Ponomareva, N., Novichkova, G., Fechina, L., Aleinikova, O. and Karachunskiy, A. (2023a) 'A simple procedure to identify children with B-lineage acute lymphoblastic leukemia who can be successfully treated with low or moderate intensity: Sequential versus single-point minimal residual disease measurement', *Pediatric Blood & Cancer,* 70(6), pp. e30295.

Popov, A., Henze, G., Roumiantseva, J., Budanov, O., Verzhbitskaya, T., Boyakova, E., Tsaur, G., Fadeeva, M., Lagoyko, S., Zharikova, L., Miakova, N., Litvinov, D., Khlebnikova, O., Streneva, O., Ponomareva, N., Novichkova, G., Fechina, L. and Karachunskiy, A. (2023b) 'Flow cytometric MRD at the end of consolidation in childhood B-lineage acute lymphoblastic leukemia has significant prognostic value but limited clinical implications: Results of study ALL-MB 2008', *Leukemia Research,* 125, pp. 106998.

PREDICT (2024) *PREDICT tool*. Available at: https://breast.predict.cam/tool (Accessed: 10/10/2024 2024).

Pui, C. H. and Evans, W. E. (2013) 'A 50-year journey to cure childhood acute lymphoblastic leukemia', *Semin Hematol,* 50(3), pp. 185-96.

Raetz, E. A. and Teachey, D. T. (2016) 'T-cell acute lymphoblastic leukemia', *Hematology,* 2016(1), pp. 580-588.

Raji, O. Y., Duffy, S. W., Agbaje, O. F., Baker, S. G., Christiani, D. C., Cassidy, A. and Field, J. K. (2012) 'Predictive accuracy of the Liverpool Lung Project risk model for stratifying patients for computed tomography screening for lung cancer: a case-control and cohort validation study', *Ann Intern Med,* 157(4), pp. 242-50.

Redaelli, A., Laskin, B. L., STEPHENS, J. M., BOTTEMAN, M. F. and PASHOS, C. L. (2005) 'A systematic literature review of the clinical and epidemiological burden of acute lymphoblastic leukaemia (ALL)', *European Journal of Cancer Care,* 14(1), pp. 53-62.

Reismüller, B., Attarbaschi, A., Peters, C., Dworzak, M. N., Pötschger, U., Urban, C., Fink, F.-M., Meister, B., Schmitt, K., Dieckmann, K., Henze, G., Haas, O. A., Gadner, H., Mann, G. and Group, o. b. o. t. A. B.-F.-M. S. (2009) 'Long-term outcome of initially homogenously treated and relapsed childhood acute lymphoblastic leukaemia in Austria – A population-based report of the Austrian Berlin-Frankfurt-Münster (BFM) Study Group', *British Journal of Haematology,* 144(4), pp. 559-570.

Ren, S.-h., Qin, Y.-f., Qin, H., Wang, H.-d., Li, G.-m., Zhu, Y.-l., Sun, C.-l., Shao, B., Zhang, J.-y., Hao, J.-p. and Wang, H. (2022) 'N6-Methyladenine-Related Signature for Immune Microenvironment and Response to Immunotherapy in Hepatocellular Carcinoma', *International Journal of General Medicine,* 15, pp. 3525-3540.

Rheingold, S. R., Ji, L., Xu, X., Devidas, M., Brown, P. A., Gore, L., Winick, N. J., Carroll, W. L., Hunger, S., Raetz, E. A., Loh, M. L. and Bhojwani, D. (2019) 'Prognostic factors for survival after relapsed acute lymphoblastic leukemia (ALL): A Children's Oncology Group (COG) study', *Journal of Clinical Oncology,* 37(15_suppl), pp. 10008-10008.

Richards, M. P. (2005) 'Techniques for Gene Expression Profiling', in Walker, J.M. and Rapley, R. (eds.) *Medical Biomethods Handbook*. Totowa, NJ: Humana Press, pp. 507-518.
Richter-Pechańska, P., Kunz, J. B., Hof, J., Zimmermann, M., Rausch, T., Bandapalli, O. R., Orlova, E., Scapinello, G., Sagi, J. C., Stanulla, M., Schrappe, M., Cario, G., Kirschner-Schwabe, R., Eckert, C., Benes, V., Korbel, J. O., Muckenthaler, M. U. and Kulozik, A. E. (2017) 'Identification of a genetically defined ultra-high-risk group in relapsed pediatric T-lymphoblastic leukemia', *Blood Cancer Journal,* 7(2), pp. e523-e523.

Rieger, M. A. and Schroeder, T. (2012) 'Hematopoiesis', *Cold Spring Harb Perspect Biol,* 4(12).

Roberts, K. G., Gu, Z., Payne-Turner, D., McCastlain, K., Harvey, R. C., Chen, I. M., Pei, D., Iacobucci, I., Valentine, M., Pounds, S. B., Shi, L., Li, Y., Zhang, J., Cheng, C., Rambaldi, A., Tosi, M., Spinelli, O., Radich, J. P., Minden, M. D., Rowe, J. M., Luger, S., Litzow, M. R., Tallman, M. S., Wiernik, P. H., Bhatia, R., Aldoss, I., Kohlschmidt, J., Mrózek, K., Marcucci, G., Bloomfield, C. D., Stock, W., Kornblau, S., Kantarjian, H. M., Konopleva, M., Paietta, E., Willman, C. L. and Mullighan, C. G. (2017) 'High Frequency and Poor Outcome of Philadelphia Chromosome-Like Acute Lymphoblastic Leukemia in Adults', *J Clin Oncol,* 35(4), pp. 394-401.

Royston, P., Moons, K. G., Altman, D. G. and Vergouwe, Y. (2009) 'Prognosis and prognostic research: Developing a prognostic model', *Bmj,* 338, pp. b604.

Samarasinghe, S., Vora, A., Goulden, N. J., Antony, G. and Moorman, A. V. (2021) 'Ten Year Outcomes of UKALL 2003: A Randomised Clinical Trial of Adjusting Treatment Intensity Based on Minimal Residual Disease', *Blood,* 138(Supplement 1), pp. 364-364.

Sang, M., Ma, L., Sang, M., Zhou, X., Gao, W. and Geng, C. (2014) 'LIM-domain-only proteins: multifunctional nuclear transcription coregulators that interacts with diverse proteins', *Molecular Biology Reports,* 41(2), pp. 1067-1073.

Saygin, C., Cannova, J., Stock, W. and Muffly, L. (2022) 'Measurable residual disease in acute lymphoblastic leukemia: methods and clinical context in adult patients', *Haematologica,* 107(12), pp. 2783-2793.

Schneider, N. R., Carroll, A. J., Shuster, J. J., Pullen, D. J., Link, M. P., Borowitz, M. J., Camitta, B. M., Katz, J. A. and Amylon, M. D. (2000) 'New recurring cytogenetic abnormalities and association of blast cell karyotypes with prognosis in childhood T-cell acute lymphoblastic leukemia: a Pediatric Oncology Group report of 343 cases', *Blood,* 96(7), pp. 2543-2549.

Schrappe, M., Valsecchi, M. G., Bartram, C. R., Schrauder, A., Panzer-Grümayer, R., Möricke, A., Parasole, R., Zimmermann, M., Dworzak, M., Buldini, B., Reiter, A., Basso, G., Klingebiel, T., Messina, C., Ratei, R., Cazzaniga, G., Koehler, R., Locatelli, F., Schäfer, B. W., Aricò, M., Welte, K., van Dongen, J. J. M., Gadner, H., Biondi, A. and Conter, V. (2011) 'Late MRD response determines relapse risk overall and in subsets of childhood T-cell ALL: results of the AIEOP-BFM-ALL 2000 study', *Blood,* 118(8), pp. 2077-2084.

Schultz, K. R., Pullen, D. J., Sather, H. N., Shuster, J. J., Devidas, M., Borowitz, M. J., Carroll, A. J., Heerema, N. A., Rubnitz, J. E., Loh, M. L., Raetz, E. A., Winick, N. J., Hunger, S. P., Carroll, W. L., Gaynon, P. S. and Camitta, B. M. (2006) 'Risk- and response-based classification of childhood B-precursor acute lymphoblastic leukemia: a combined analysis of prognostic markers from the Pediatric Oncology Group (POG) and Children's Cancer Group (CCG)', *Blood,* 109(3), pp. 926-935.

Shuster, J. J., Falletta, J. M., Pullen, D. J., Crist, W. M., Humphrey, G. B., Dowell, B. L., Wharam, M. D. and Borowitz, M. (1990) 'Prognostic Factors in Childhood T-Cell Acute Lymphoblastic Leukemia: A Pediatric Oncology Group Study', *Blood,* 75(1), pp. 166-173.

Simonin, M., Vasseur, L., Lengliné, E., Lhermitte, L., Cabannes-Hamy, A., Balsat, M., Schmidt, A., Dourthe, M.-E., Touzart, A., Graux, C., Grardel, N., Cayuela, J.-M., Arnoux, I., Gandemer, V., Huguet, F., Ducassou, S., Lhéritier, V., Chalandon, Y., Ifrah, N., Dombret, H., Macintyre, E., Petit, A., Rousselot, P., Lambert, J., Baruchel, A., Boissel, N. and Asnafi, V. (2024) 'NGS-based stratification refines the risk stratification in T-ALL and identifies a very-high-risk subgroup of patients', *Blood,* 144(15), pp. 1570-1580.

Sin, C. F. and Man, P. M. (2021) 'Early T-Cell Precursor Acute Lymphoblastic Leukemia: Diagnosis, Updates in Molecular Pathogenesis, Management, and Novel Therapies', *Front Oncol,* 11, pp. 750789.

Smith, M., Arthur, D., Camitta, B., Carroll, A. J., Crist, W., Gaynon, P., Gelber, R., Heerema, N., Korn, E. L., Link, M., Murphy, S., Pui, C. H., Pullen, J., Reamon, G., Sallan, S. E., Sather, H., Shuster, J., Simon, R., Trigg, M., Tubergen, D., Uckun, F. and Ungerleider, R. (1996) 'Uniform approach to risk classification and treatment assignment for children with acute lymphoblastic leukemia', *Journal of Clinical Oncology,* 14(1), pp. 18-24.

Song, Y. Y. and Lu, Y. (2015) 'Decision tree methods: applications for classification and prediction', *Shanghai Arch Psychiatry,* 27(2), pp. 130-5.

Soulier, J., Clappier, E., Cayuela, J.-M., Regnault, A., García-Peydró, M., Dombret, H., Baruchel, A., Toribio, M.-L. and Sigaux, F. o. (2005) 'HOXA genes are included in genetic and biologic networks defining human acute T-cell leukemia (T-ALL)', *Blood,* 106(1), pp. 274-286.

Szczepański, T., Flohr, T., van der Velden, V. H. J., Bartram, C. R. and van Dongen, J. J. M. (2002) 'Molecular monitoring of residual disease using antigen receptor genes in childhood acute lymphoblastic leukaemia', *Best Practice & Research Clinical Haematology,* 15(1), pp. 37-57.

Taj, M. M., Moorman, A. V., Hamadeh, L., Petit, A., Asnafi, V., Alby-Laurent, F., Vora, A., Mansour, M. R., Gale, R., Chevret, S., Moppett, J., Baruchel, A. and Macintyre, E. (2022) 'Prognostic value of Oncogenetic mutations in pediatric T Acute Lymphoblastic Leukemia: a comparison of UKALL2003 and FRALLE2000T protocols', *Leukemia,* 36(1), pp. 263-266.

Tasian, S. K., Loh, M. L. and Hunger, S. P. (2015) 'Childhood acute lymphoblastic leukemia: Integrating genomics into therapy', *Cancer,* 121(20), pp. 3577-3590.

Teachey, D. T., Devidas, M., Wood, B. L., Chen, Z., Hayashi, R. J., Hermiston, M. L., Annett, R. D., Archer, J. H., Asselin, B. L., August, K. J., Cho, S. Y., Dunsmore, K. P., Fisher, B. T., Freedman, J. L., Galardy, P. J., Harker-Murray, P., Horton, T. M., Jaju, A. I., Lam, A., Messinger, Y. H., Miles, R. R., Okada, M., Patel, S. I., Schafer, E. S., Schechter, T., Singh, N., Steele, A. C., Sulis, M. L., Vargas, S. L., Winter, S. S., Wood, C., Zweidler-McKay, P., Bollard, C. M., Loh, M. L., Hunger, S. P. and Raetz, E. A. (2022) 'Children's Oncology Group Trial AALL1231: A Phase III Clinical Trial Testing Bortezomib in Newly Diagnosed T-Cell Acute Lymphoblastic Leukemia and Lymphoma', *J Clin Oncol,* 40(19), pp. 2106-2118.

Teachey, D. T. and Pui, C. H. (2019) 'Comparative features and outcomes between paediatric T-cell and B-cell acute lymphoblastic leukaemia', *Lancet Oncol,* 20(3), pp. e142-e154.

Terwilliger, T. and Abdul-Hay, M. (2017) 'Acute lymphoblastic leukemia: a comprehensive review and 2017 update', *Blood Cancer Journal,* 7(6), pp. e577-e577.

Thongprayoon, C., Dumancas, C. Y., Nissaisorakarn, V., Keddis, M. T., Kattah, A. G., Pattharanitima, P., Petnak, T., Vallabhajosyula, S., Garovic, V. D., Mao, M. A., Dillon, J. J., Erickson, S. B. and Cheungpasitporn, W. (2021) 'Machine Learning Consensus Clustering Approach for Hospitalized Patients with Phosphate Derangements', *Journal of Clinical Medicine,* 10(19), pp. 4441.

Todd, I. (2010) 'Cells of the Immune System', *eLS.*

Toft, N., Birgens, H., Abrahamsson, J., Griškevičius, L., Hallböök, H., Heyman, M., Klausen, T. W., Jónsson Ó, G., Palk, K., Pruunsild, K., Quist-Paulsen, P., Vaitkeviciene, G., Vettenranta, K., Åsberg, A., Frandsen, T. L., Marquart, H. V., Madsen, H. O., Norén-Nyström, U. and Schmiegelow, K. (2018) 'Results of NOPHO ALL2008 treatment for patients aged 1-45 years with acute lymphoblastic leukemia', *Leukemia,* 32(3), pp. 606-615.

Trinquand, A., Tanguy-Schmidt, A., Ben Abdelali, R., Lambert, J., Beldjord, K., Lengliné, E., De Gunzburg, N., Payet-Bornet, D., Lhermitte, L., Mossafa, H., Lhéritier, V., Bond, J., Huguet, F., Buzyn, A., Leguay, T., Cahn, J. Y., Thomas, X., Chalandon, Y., Delannoy, A., Bonmati, C., Maury, S., Nadel, B., Macintyre, E., Ifrah, N., Dombret, H. and Asnafi, V. (2013) 'Toward a NOTCH1/FBXW7/RAS/PTEN-based oncogenetic risk classification of adult T-cell acute lymphoblastic leukemia: a Group for Research in Adult Acute Lymphoblastic Leukemia study', *J Clin Oncol,* 31(34), pp. 4333-42.

Vaitkevičienė, G., Forestier, E., Hellebostad, M., Heyman, M., Jonsson, O. G., Lähteenmäki, P. M., Rosthoej, S., Söderhäll, S., Schmiegelow, K., Haematology, O. b. o. t. N. S. o. P. and Oncology (2011) 'High white blood cell count at diagnosis of childhood acute lymphoblastic leukaemia: biological background and prognostic impact. Results from the NOPHO ALL-92 and ALL-2000 studies', *European Journal of Haematology,* 86(1), pp. 38-46.

Van Vlierberghe, P., Pieters, R., Beverloo, H. B. and Meijerink, J. P. P. (2008) 'Molecular-genetic insights in paediatric T-cell acute lymphoblastic leukaemia', *British Journal of Haematology,* 143(2), pp. 153-168.

Vardiman, J. W., Thiele, J., Arber, D. A., Brunning, R. D., Borowitz, M. J., Porwit, A., Harris, N. L., Le Beau, M. M., Hellström-Lindberg, E., Tefferi, A. and Bloomfield, C. D. (2009) 'The 2008 revision of the World Health Organization (WHO) classification of myeloid neoplasms and acute leukemia: rationale and important changes', *Blood,* 114(5), pp. 937-951.

Vicente, C., Schwab, C., Broux, M., Geerdens, E., Degryse, S., Demeyer, S., Lahortiga, I., Elliott, A., Chilton, L., La Starza, R., Mecucci, C., Vandenberghe, P., Goulden, N., Vora, A., Moorman, A. V., Soulier, J., Harrison, C. J., Clappier, E. and Cools, J. (2015) 'Targeted sequencing identifies associations between IL7R-JAK mutations and epigenetic modulators in T-cell acute lymphoblastic leukemia', *Haematologica,* 100(10), pp. 1301-10.

Vinti, L., Strocchio, L., Buldini, B., Silvestri, D., Conter, V., Merli, P., Rizzari, C., Parasole, R., Paganin, M., Basso, G., Zecca, M. and Locatelli, F. (2019) 'Long-Term Outcome of Relapsed Acute T-Lymphoblastic Leukemia (T-ALL) in Children and Adolescents', *Blood,* 134, pp. 2580.

Vora, A. (2017) *Childhood Acute Lymphoblastic Leukemia.*

Vora, A., Goulden, N., Mitchell, C., Hancock, J., Hough, R., Rowntree, C., Moorman, A. V. and Wade, R. (2014) 'Augmented post-remission therapy for a minimal residual disease-defined high-risk subgroup of children and young people with clinical standard-risk and intermediate-risk acute lymphoblastic leukaemia (UKALL 2003): a randomised controlled trial', *Lancet Oncol,* 15(8), pp. 809-18.

Vora, A., Goulden, N., Wade, R., Mitchell, C., Hancock, J., Hough, R., Rowntree, C. and Richards, S. (2013) 'Treatment reduction for children and young adults with low-risk acute lymphoblastic leukaemia defined by minimal residual disease (UKALL 2003): a randomised controlled trial', *The Lancet Oncology,* 14(3), pp. 199-209.

Vora, A., Mitchell, C. D., Lennard, L., Eden, T. O., Kinsey, S. E., Lilleyman, J. and Richards, S. M. (2006) 'Toxicity and efficacy of 6-thioguanine versus 6-mercaptopurine in childhood lymphoblastic leukaemia: a randomised trial', *The Lancet,* 368(9544), pp. 1339-1348.

Wada, M., Bartram, C. R., Nakamura, H., Hachiya, M., Chen, D.-L., Borenstein, J., Miller, C. W., Ludwig, L., Hansen-Hagge, T. E., Ludwig, W.-D., Reiter, A., Mizoguchi, H. and Koeffler, H. P. (1993) 'Analysis of p53 Mutations in a Large Series of Lymphoid Hematologic Malignancies of Childhood', *Blood,* 82(10), pp. 3163-3169.

Wang, C.-W., Huang, S.-C., Lee, Y.-C., Shen, Y.-J., Meng, S.-I. and Gaol, J. L. (2022) 'Deep learning for bone marrow cell detection and classification on whole-slide images', *Medical Image Analysis,* 75, pp. 102270.

Wang, Y., Zeng, H.-m. and Zhang, L.-p. (2018) 'ETV6/RUNX1-positive childhood acute lymphoblastic leukemia in China: excellent prognosis with improved BFM protocol', *Italian Journal of Pediatrics,* 44(1), pp. 94.

Wei, W., Chen, X., Zou, Y., Chang, L., An, W., Wan, Y., Liu, T., Yang, W., Chen, Y., Guo, Y. and Zhu, X. (2015) 'Prediction of outcomes by early treatment responses in childhood T-cell acute lymphoblastic leukemia: a retrospective study in China', *BMC Pediatrics,* 15(1), pp. 80.

Wen, J., Zhou, M., Shen, Y., long, Y., Guo, Y., Song, L. and Xiao, J. (2022) 'Poor treatment responses were related to poor outcomes in pediatric B cell acute lymphoblastic leukemia with KMT2A rearrangements', *BMC Cancer,* 22(1), pp. 859.

Willemse, M. J., Seriu, T., Hettinger, K., d'Aniello, E., Hop, W. C. J., Panzer-Grümayer, E. R., Biondi, A., Schrappe, M., Kamps, W. A., Masera, G., Gadner, H., Riehm, H., Bartram, C. R. and van Dongen, J. J. M. (2002) 'Detection of minimal residual disease identifies differences in treatment response between T-ALL and precursor B-ALL', *Blood,* 99(12), pp. 4386-4393.

Wishart, G. C., Azzato, E. M., Greenberg, D. C., Rashbass, J., Kearins, O., Lawrence, G., Caldas, C. and Pharoah, P. D. P. (2010) 'PREDICT: a new UK prognostic model that predicts survival following surgery for invasive breast cancer', *Breast Cancer Research,* 12(1), pp. R1.

Wishart, G. C., Rakha, E., Green, A., Ellis, I., Ali, H. R., Provenzano, E., Blows, F. M., Caldas, C. and Pharoah, P. D. (2014) 'Inclusion of KI67 significantly improves performance of the PREDICT prognostication and prediction model for early breast cancer', *BMC Cancer,* 14, pp. 908.

Wong, H. S., Subramaniam, S., Alias, Z., Taib, N. A., Ho, G. F., Ng, C. H., Yip, C. H., Verkooijen, H. M., Hartman, M. and Bhoo-Pathy, N. (2015) 'The predictive accuracy of PREDICT: a personalized decision-making tool for Southeast Asian women with breast cancer', *Medicine (Baltimore),* 94(8), pp. e593.

Wood, B. L., Winter, S. S., Dunsmore, K. P., Devidas, M., Chen, S., Asselin, B., Esiashvili, N., Loh, M. L., Winick, N. J., Carroll, W. L., Raetz, E. A. and Hunger, S. P. (2014) 'T-Lymphoblastic Leukemia (T-ALL) Shows Excellent Outcome, Lack of Significance of the Early Thymic Precursor (ETP) Immunophenotype, and Validation of the Prognostic Value of End-Induction Minimal Residual Disease (MRD) in Children's Oncology Group (COG) Study AALL0434', *Blood,* 124(21), pp. 1.

Xia, Y., Brown, L., Yang, C. Y., Tsan, J. T., Siciliano, M. J., Espinosa, R., 3rd, Le Beau, M. M. and Baer, R. J. (1991) 'TAL2, a helix-loop-helix gene activated by the (7;9)(q34;q32) translocation in human T-cell leukemia', *Proc Natl Acad Sci U S A,* 88(24), pp. 11416-20.

Yuan, Y., Li, J., Xue, T.-L., Hu, H.-R., Lin, W., Liu, S.-G., Zhang, R.-D., Zheng, H.-Y. and Gao, C. (2022) 'Prognostic significance of NOTCH1/FBXW7 mutations in pediatric T cell acute lymphoblastic leukemia: a study of minimal residual disease risk-directed CCLG-ALL 2008 treatment protocol', *Leukemia & Lymphoma,* 63(7), pp. 1624-1633.

Zhang, J., Ding, L., Holmfeldt, L., Wu, G., Heatley, S. L., Payne-Turner, D., Easton, J., Chen, X., Wang, J., Rusch, M., Lu, C., Chen, S. C., Wei, L., Collins-Underwood, J. R., Ma, J., Roberts, K. G., Pounds, S. B., Ulyanov, A., Becksfort, J., Gupta, P., Huether, R., Kriwacki, R. W., Parker, M., McGoldrick, D. J., Zhao, D., Alford, D., Espy, S., Bobba, K. C., Song, G., Pei, D., Cheng, C., Roberts, S., Barbato, M. I., Campana, D., Coustan-Smith, E., Shurtleff, S. A., Raimondi, S. C., Kleppe, M., Cools, J., Shimano, K. A., Hermiston, M. L., Doulatov, S., Eppert, K., Laurenti, E., Notta, F., Dick, J. E., Basso, G., Hunger, S. P., Loh, M. L., Devidas, M., Wood, B., Winter, S., Dunsmore, K. P., Fulton, R. S., Fulton, L. L., Hong, X., Harris, C. C., Dooling, D. J., Ochoa, K., Johnson, K. J., Obenauer, J. C., Evans, W. E., Pui, C. H., Naeve, C. W., Ley, T. J., Mardis, E. R., Wilson, R. K., Downing, J. R. and Mullighan, C. G. (2012) 'The genetic basis of early T-cell precursor acute lymphoblastic leukaemia', *Nature,* 481(7380), pp. 157-63.

Zuurbier, L., Homminga, I., Calvert, V., Winkel, M., Buijs-Gladdines, J., Kooi, C., Smits, W., Sonneveld, E., Veerman, A., Kamps, W., Horstmann, M., Petricoin, E., Pieters, R. and Meijerink, J. (2010) 'NOTCH1 and/or FBXW7 mutations predict for initial good prednisone response but not for improved outcome in pediatric T-cell acute lymphoblastic leukemia patients treated on DCOG or COALL protocols', *Leukemia : official journal of the Leukemia Society of America, Leukemia Research Fund, U.K,* 24, pp. 2014-22.