Hodgkin’s Lymphoma

Alberto Anelli, Lorenzo Giovine, Elia Perli

Outline

- Introduction
- Statistical analysis
- Machine Learning Analysis

Introduction

Essentials about Hodgkin's lymphoma

The disease:

- Blood cancer that originates from lymphocytes (especially B-type) and is found in the lymphatic system
- Presence of Reed-Sternberg cells, large abnormal cells found in lymph node biopsies.
- Painless swelling of lymph nodes, fever, night sweats, unexplained weight loss, and fatigue.

Introduction Essentials about Hodgkin’s lymphoma

Epidemiology

- Incidence: 2–3 cases per 100,000 people annually worldwide.
- Affects mostly young adults (15–35 years) and older adults (>55 years) (bimodal distribution)
- Slight male predominance.
- More common in developed countries; associations with Epstein-Barr Virus (EBV) in certain regions.

Introduction Essentials about Hodgkin’s lymphoma

Why is it important?

- Hodgkin’s lymphoma is one of the most curable cancers, with a 5-year survival rate of around 86%.
- Studying Hodgkin's lymphoma could be valuable for the advancement of cancer biology, comprehending immune interactions, and the development of targeted therapies

Introduction The base study

CLINICAL CANCER RESEARCH | PRECISION MEDICINE AND IMAGING

A Gene Expression–based Model to Predict Metabolic Response After Two Courses of ABVD in Hodgkin Lymphoma Patients

Stefano Luminari[1,2], Benedetta Donati[3], Massimiliano Casali[4], Riccardo Valli[5], Raffaella Santi[6], Benedetta Puccini[7], Sofya Kovalchuk[7], Alessia Ruffini[1,8], Angelo Fama[1], Valentina Berti[9], Valentina Fragliasso[3], Magda Zanelli[5], Federica Vergoni[6], Annibale Versari[4], Luigi Rigacci[10], Francesco Merli[1], and Alessia Ciarrocchi[3]

ABSTRACT

Purpose: Early response to ABVD, assessed with interim FDG-PET (iPET), is prognostic for classical Hodgkin lymphoma (cHL) and supports the use of response adapted therapy. The aim of this study was to identify a gene-expression profile on diagnostic biopsy to predict iPET positivity (iPET$^+$).

Experimental Design: Consecutive untreated patients with stage I–IV cHL who underwent iPET after two cycles of ABVD were identified. Expression of 770 immune-related genes was analyzed by digital expression profiling (NanoString Technology). iPET was centrally reviewed according to the five-point Deauville scale (DS 1-5). An iPET$^+$ predictive model was derived by multivariate regression analysis and assessed in a validation set identified using the same inclusion criteria.

Results: A training set of 121 and a validation set of 117 patients were identified, with 23 iPET$^+$ cases in each group. Sixty-three (52.1%), 19 (15.7%), and 39 (32.2%) patients had stage I–II, III, and IV, respectively. Diagnostic biopsy of iPET$^+$ cHLs showed transcriptional profile distinct from iPET$^-$. Thirteen genes were stringently associated with iPET$^+$. This signature comprises two functionally stromal-related nodes. Lymphocytes/monocytes ratio (LMR) was also associated to iPET$^+$. In the training cohort a 5-gene/LMR integrated score predicted iPET$^+$ [AUC, 0.88; 95% confidence interval (CI), 0.80–0.96]. The score achieved a 100% sensitivity to identify DS5 cases. Model performance was confirmed in the validation set (AUC, 0.68; 95% CI, 0.52–0.84). Finally, iPET score was higher in patients with event versus those without.

Conclusions: In cHL, iPET is associated with a genetic signature and can be predicted by applying an integrated gene-based model on the diagnostic biopsy.

Introduction

and dacarbazine (ABVD) or of the more intense bleomycin, etoposide,

Background

Definition:

Metabolic activity = biochemical processes and reactions occurring within cells or tissues that are necessary for maintaining life and function

Metabolic response = measure of the variation of metabolic activity

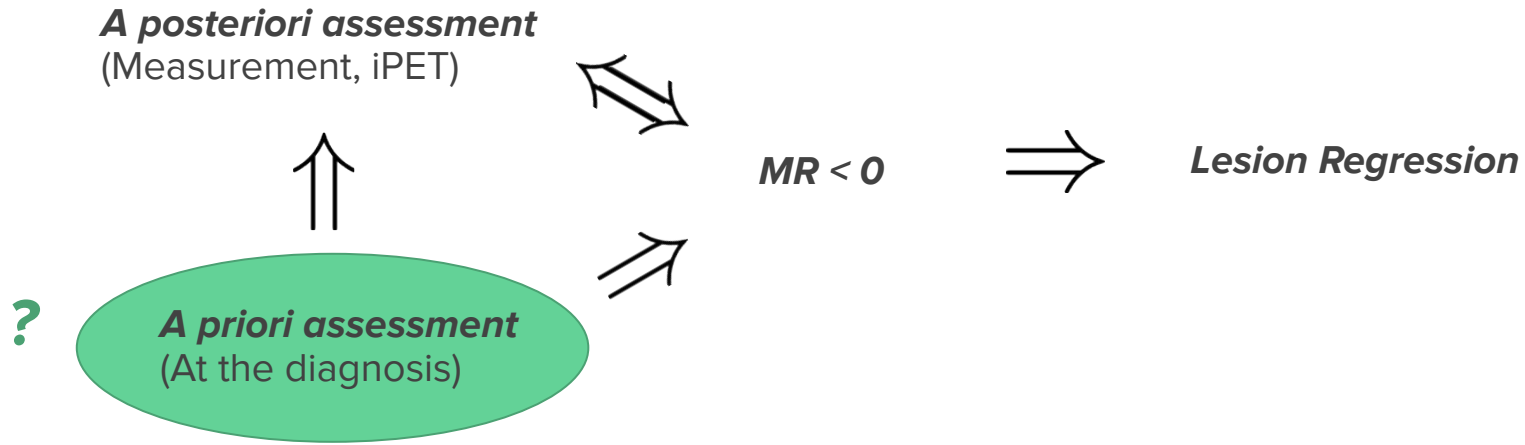
Remark (important):

Negative metabolic response after chemotherapy, in cHL, is correlated with lesion regression

Background

Remark: The direct assessment of metabolic regression can be done only a posteriori.

Goals: Find a proxy for metabolic regression using Genomic Statistical Analysis



Background

Applications:

1. Assessing the relationship between metabolic response and genetic to highlight some mechanisms that underlie cancer biology
2. Further the development of **response-adapted therapies**

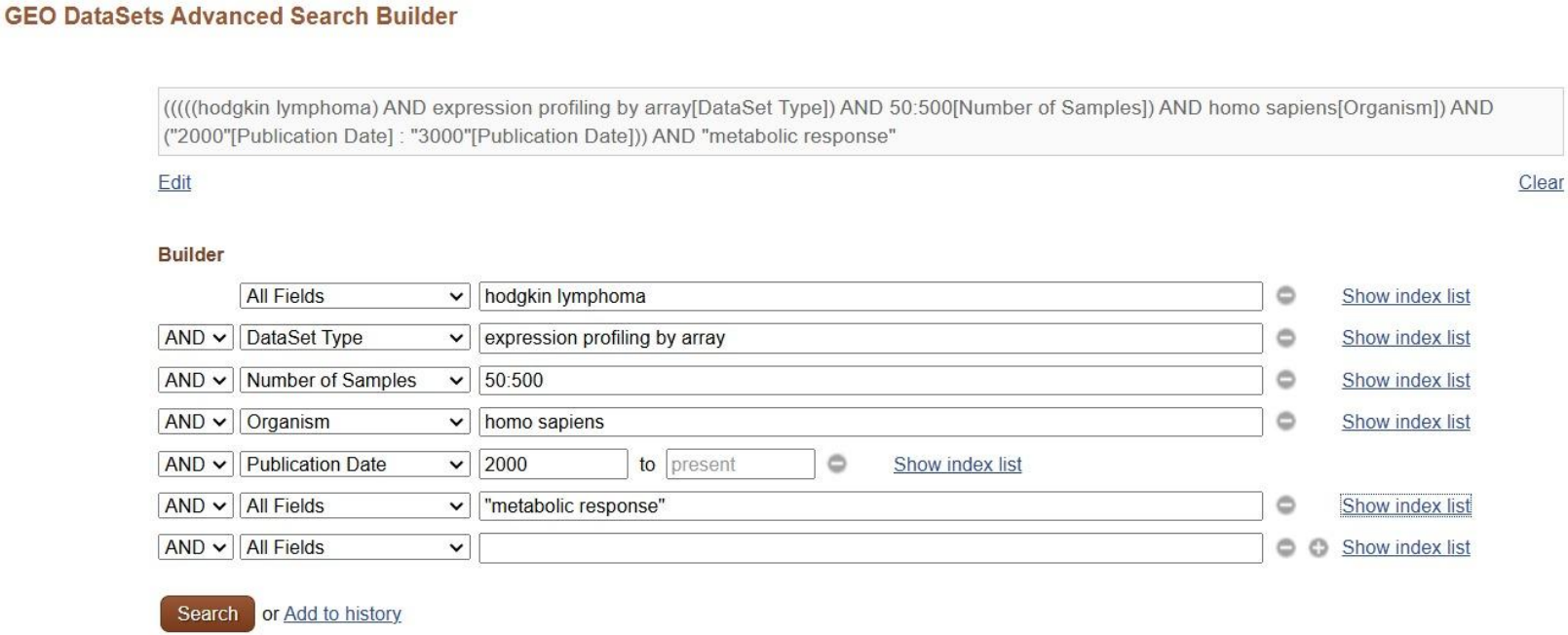
Remark:

This kind of cancer is well-suited for application **2.**, due to its low-risk profile

Methods

1. Dataset search
2. Data normalization
3. Feature selection
4. Model training
5. Model validation

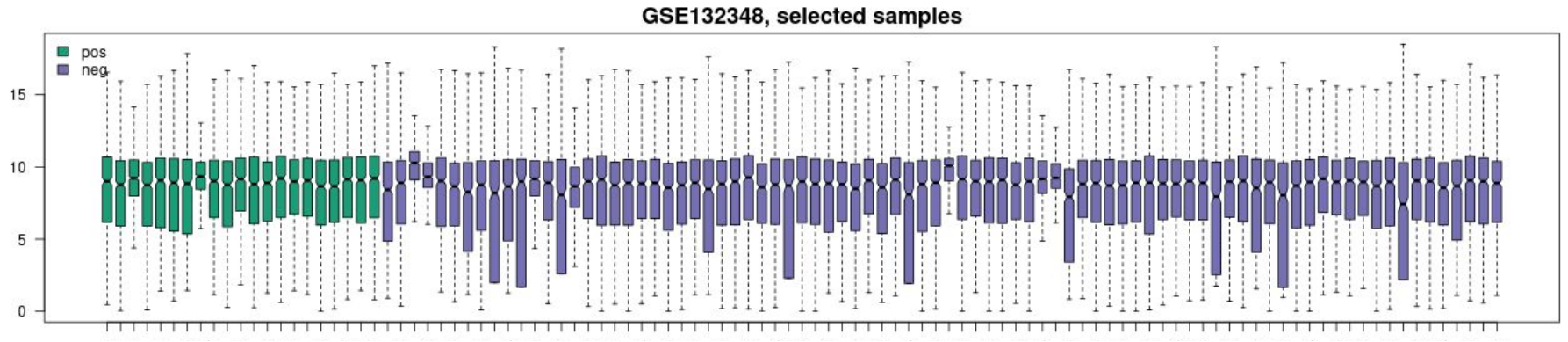
Finding the right dataset



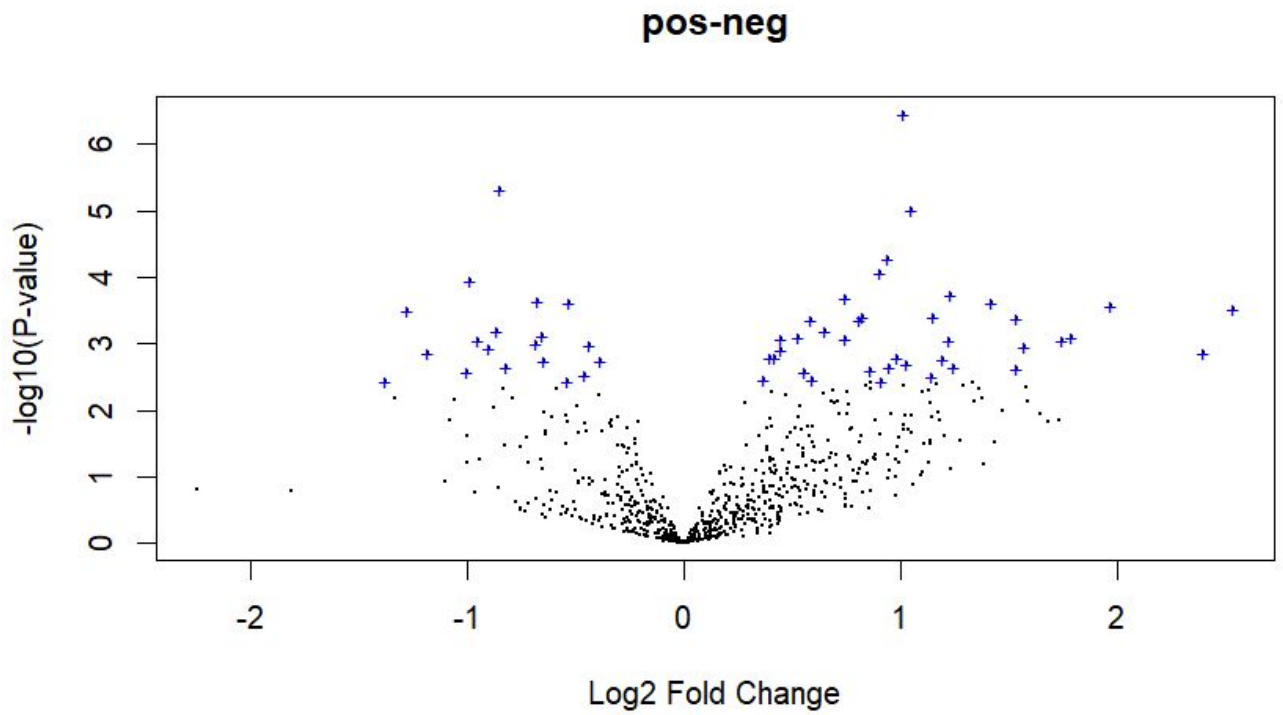
That’s the right one!

Series GSE132348	Query DataSets for GSE132348
Status	Public on Sep 26, 2019
Title	A gene expression-based model to predict metabolic response after two courses of ABVD in patients with classical Hodgkin Lymphoma
Organism	Homo sapiens
Experiment type	Expression profiling by array
Summary	RNA extracted from diagnostic tumor samples of 106 patients affected by Hodgkin Lymphoma was analyzed on the nCounter system using the PanCancer Immune Profiling Panel.
Overall design	Gene expression profile were linked to interim PET response in Hodgkin Lymphoma.
Contributor(s)	Ciarrocchi A, Luminari S, Donati B
Citation(s)	Luminari S, Donati B, Casali M, Valli R et al. A Gene Expression-based Model to Predict Metabolic Response After Two Courses of ABVD in Hodgkin Lymphoma Patients. *Clin Cancer Res* 2020 Jan 15;26(2):373-383. PMID: 31645353

Normalising data



Analysis of differentially expressed genes

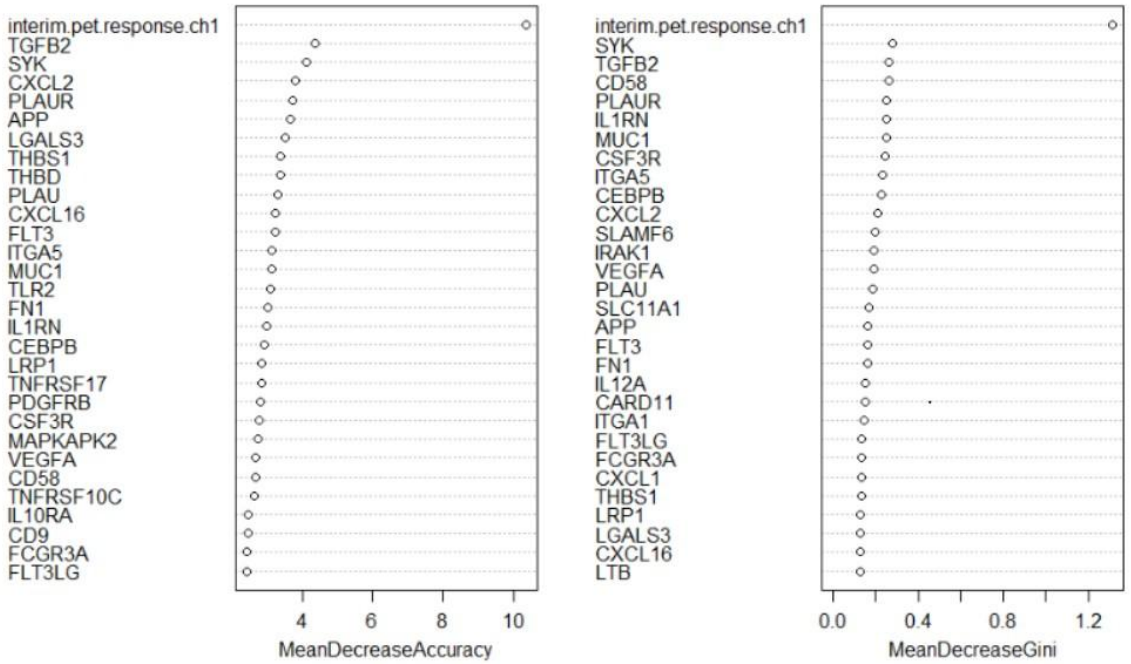


Random Forest

```
Call:
 randomForest(formula = train$group ~ ., data = train, ntree = 1000,
importance = TRUE, cutoff = c(0.5, 0.4))
               Type of random forest: classification
                     Number of trees: 1000
No. of variables tried at each split: 28

        OOB estimate of  error rate: 13.51%
Confusion matrix:
    neg pos class.error
neg  55   4  0.06779661
pos   6   9  0.40000000
```

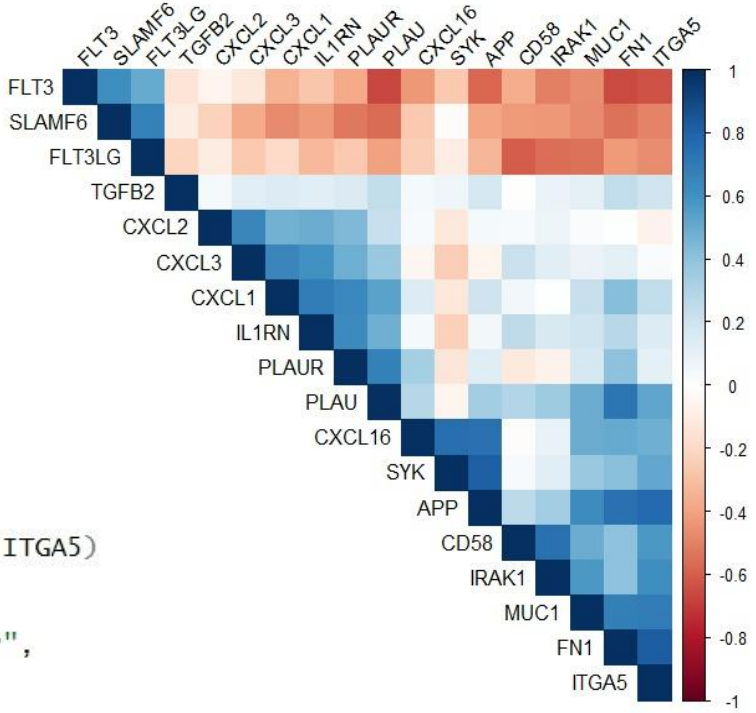
```
> #test the model
> pred <- predict(rf_new, test, type="response")
> confusionMatrix(as.factor(test$group), as.factor(pred))
Confusion Matrix and Statistics

          Reference
Prediction neg pos
       neg  25   0
       pos   3   3

               Accuracy : 0.9032
                 95% CI : (0.7425, 0.9796)
    No Information Rate : 0.9032
    P-Value [Acc > NIR] : 0.6474

                  Kappa : 0.6173

 Mcnemar's Test P-Value : 0.2482

            Sensitivity : 0.8929
            Specificity : 1.0000
         Pos Pred Value : 1.0000
         Neg Pred Value : 0.5000
             Prevalence : 0.9032
         Detection Rate : 0.8065
   Detection Prevalence : 0.8065
      Balanced Accuracy : 0.9464

       'Positive' Class : neg
```

Importanza


interim.pet.response.ch1
TGFB2
SYK
CXCL2
PLAUR
APP
LGALS3
THBS1
THBD
PLAU
CXCL16
FLT3
ITGA5
MUC1
TLR2
FN1
IL1RN
CEBPB
LRP1
TNFRSF17
PDGFRB
CSF3R
MAPKAPK2
VEGFA
CD58
TNFRSF10C
IL10RA
CD9
FCGR3A
FLT3LG
4
6
8
10
MeanDecreaseAccuracy
interim.pet.response.ch1
SYK
TGFB2
CD58
PLAUR
IL1RN
MUC1
CSF3R
ITGA5
CEBPB
CXCL2
SLAMF6
IRAK1
VEGFA
PLAU
SLC11A1
APP
FLT3
FN1
IL12A
CARD11
ITGA1
FLT3LG
FCGR3A
CXCL1
THBS1
LRP1
LGALS3
CXCL16
LTB
0.0
0.4
0.8
1.2
MeanDecreaseGini

Correlation Matrix plot

```
gset_subset <- gset_new %>% select(CXCL1, CXCL2, CXCL3,
                                   PLAUR, TGFB2, SYK,
                                   CD58, IL1RN, MUC1,
                                   PLAU, IRAK1, FLT3,
                                   CXCL16, APP, CD58,
                                   FN1, SLAMF6, FLT3LG, ITGA5)
#correlation matrix
correlation_matrix <- cor(gset_subset)
corrplot(correlation_matrix, method="color", type="upper",
         order="hclust", tl.col="black", tl.srt=45)
```



Conclusions

The **RandomForest model** gives an accurate representation of chemotherapy response on the basis of genetical parameters. It has an accuracy of about 90% in the test set.

Considering the biological question, we valued specificity over sensitivity, since it is preferable to clinically classify a treatment-responsive lymphoma as non responsive, instead of viceversa.

Thank you for your attention