Exploring the Metabolic Profile of Juvenile Myelomonocytic Leukemia Through Newborn Screening

Steven Brooks, MS[1], Lillian Rubino[2], Elliot Stieglitz, MD[3], Gang Peng, PhD[1]
[1]Department of Medical and Molecular Genetics, Indiana University School of Medicine
[2]Student Researcher, Department of Medical and Molecular Genetics, Indiana University School of Medicine
[3]Department of Pediatrics, University of California, San Francisco

Background

Juvenile Myelomonocytic Leukemia (JMML) is a rare pediatric leukemia involving myeloid cell overproduction, typically diagnosed only after clinical symptoms appear. Although **Newborn Screening (NBS)** routinely tests infants for treatable conditions using dried blood spots, JMML is not currently included. This study investigates metabolic differences between JMML patients and healthy controls to identify potential biomarkers detectable in NBS samples.

Participants

Blood sample and clinical data were obtained from 58 JMML patients and 58 healthy controls using dried blood spots collected through the California Newborn Screening program. Metabolic intensity was measured using mass spectrometry with two modes: reverse-phase negative with a PFAS-free kit (RPNPF) and HILIC chromatography with positive ionization (ZHP).

Total samples (n=116): 58 JMML (35 M/23 F) and 58 controls (34 M/24 F)

Data Preprocess

> 65 % missing → removed

20 – 65 % missing → Fisher's exact test

< 20 % missing → Filter, impute, transform & t-tests

Annotated Metabolites

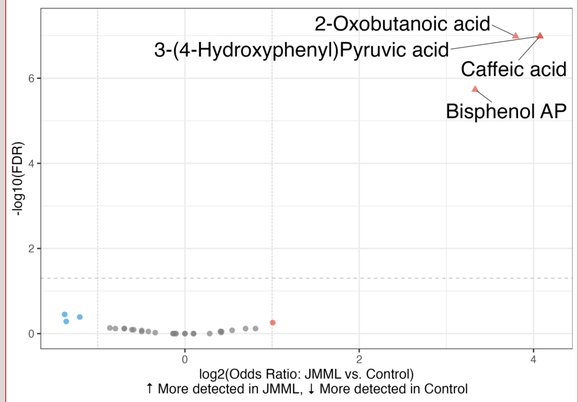
Step	ZHP	RPNPF
Initial	315	298
After ≤65% Missingness	258	213
After Pooled Noise-to-Signal ≤ 0.3	222	162
After CV (group) ≤ 0.3	178	155

Unannotated Metabolites

Step	ZHP	RPNPF
Initial	42759	15914
After ≤65% Missingness	41807	15196
After Pooled Noise-to-Signal ≤ 0.3	28869	8014
After CV (group) ≤ 0.3	10143	6693

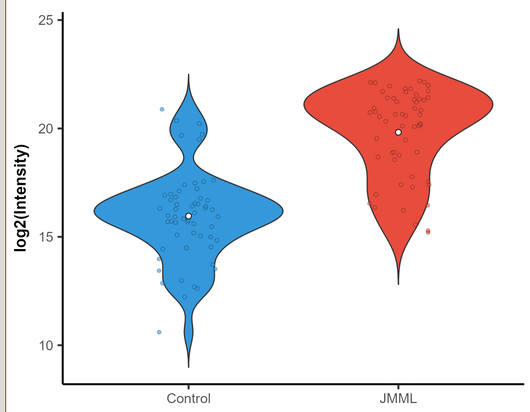
Annotated Results

JMML-based Metabolite Differences
Fisher's Exact Test, ZHP Platform



Red points indicate metabolites enriched in JMML, while blue points in controls. Among the metabolites significantly enriched in JMML cases, Bisphenol AP, an endocrine-disrupting chemical commonly found in plastics, was highly detected, along with Caffeic acid, 2-Oxobutanoic acid, and 3-(4-Hydroxyphenyl)Pyruvic acid.

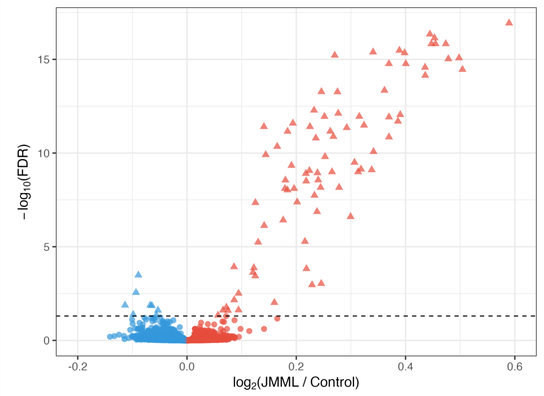
Distribution of 2-Hydroxychrysene Intensities across Diagnostic Groups



Violin plot showing the log2-transformed intensities of 2-hydroxychrysene in dried blood spot samples from JMML patients (red) and controls (blue). White dots indicate group means. JMML samples show significantly elevated levels of 2-hydroxychrysene after multiple-test correction.

Unannotated Results

JMML-based Metabolite Differences
T-Test, ZHP Platform



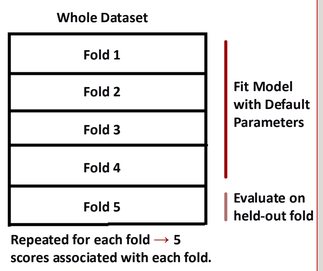
Volcano plot displaying the -$\log_{10}$ FDR versus $\log_2$ fold change in Box-Cox transformed metabolite intensities. Red points indicate metabolites significantly enriched in JMML (FDR < 0.05), while blue points are enriched in controls. Triangles represent metabolites passing the FDR threshold, and dotted lines mark $\log_2$(alpha level).

Datamining Methodology

Goal: classify a dried blood spot card as JMML positive or negative

Setup

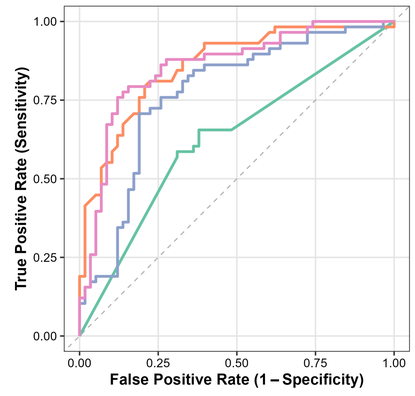
- 5-fold cross-validation
- Classes balanced in each fold
- Compare logistic regression, random forest, SVM, and XGBoost
- Default parameters
- AUROC curve used to visualize results



Features: Box-cox transformed intensities, binary detection indicators, sample metadata (race/ethnicity, sex, hemoglobin levels)

Datamining Results

Model Performance for Classifying JMML Using Metabolomic Data



Cross-validated receiver operating characteristic (ROC) curves for four models classifying JMML versus control using Box-Cox QRLIC imputed intensities and binarized-on-missingness metabolites. The solid curves show the pooled ROC for Logistic Regression (green), Support Vector Machine (orange), Random Forest (blue), and XGBoost (magenta). The dashed diagonal line represents random-chance performance (AUROC = 0.5).

Conclusion

- Several metabolites and chemicals were found different between JMML patients and healthy controls
- Maternal consumption of coffee is associated with acute lymphoblastic leukemia[1]. High Caffeic acid and BPA level in JMML patients might be related to maternal coffee intake.
- Machine learning models demonstrated good classification performance and may support early JMML detection through newborn screening using DBS.

References

1. Milne E, et al. *Cancer Causes & Control*. 2018;29(5):455–464. https://doi.org/10.1007/s10552-018-1028-y

Acknowledgments

This project is supported in part by the NIH R25 CA269039 to the IU Simon Comprehensive Cancer Center Grant for the SPARC Program and NIH 5U54CA196519.