SUPPLEMENTARY APPENDIX

Bone marrow histomorphological criteria can accurately diagnose hemophagocytic lymphohistiocytosis

Eric Gars,[1] Natasha Purington,[1] Gregory Scott,[1] Karen Chisholm,[2] Dita Gratzinger,[1] Beth A. Martin[1*] Robert S. Ohgami[1*]

[1]Stanford University, Stanford, CA and [2]Seattle Children's Hospital and University of Washington, Seattle, WA, USA

BAM and RSO contributed equally to this work.

 doi:10.3324/haematol.2017.186627Received: December 17, 2017.

Accepted: June 13, 2018.

Pre-published: June 14, 2018.

Correspondence: ericgars@stanford.edu/rohgami@stanford.edu

Statistical Analysis

Descriptive statistics such as means, medians, standard deviations and interquartile ranges were generated to evaluate distributions of key variables by HLH diagnosis. Correlations among age, HLH diagnostic criteria, and lineages were assessed using Spearman's rank. The Student's T-test was used to test for significant differences by HLH diagnosis in patient demographic data. The Kruskal-Wallis rank sum test was used to determine if there were significant differences in hemophagocytosis lineage by HLH diagnosis.

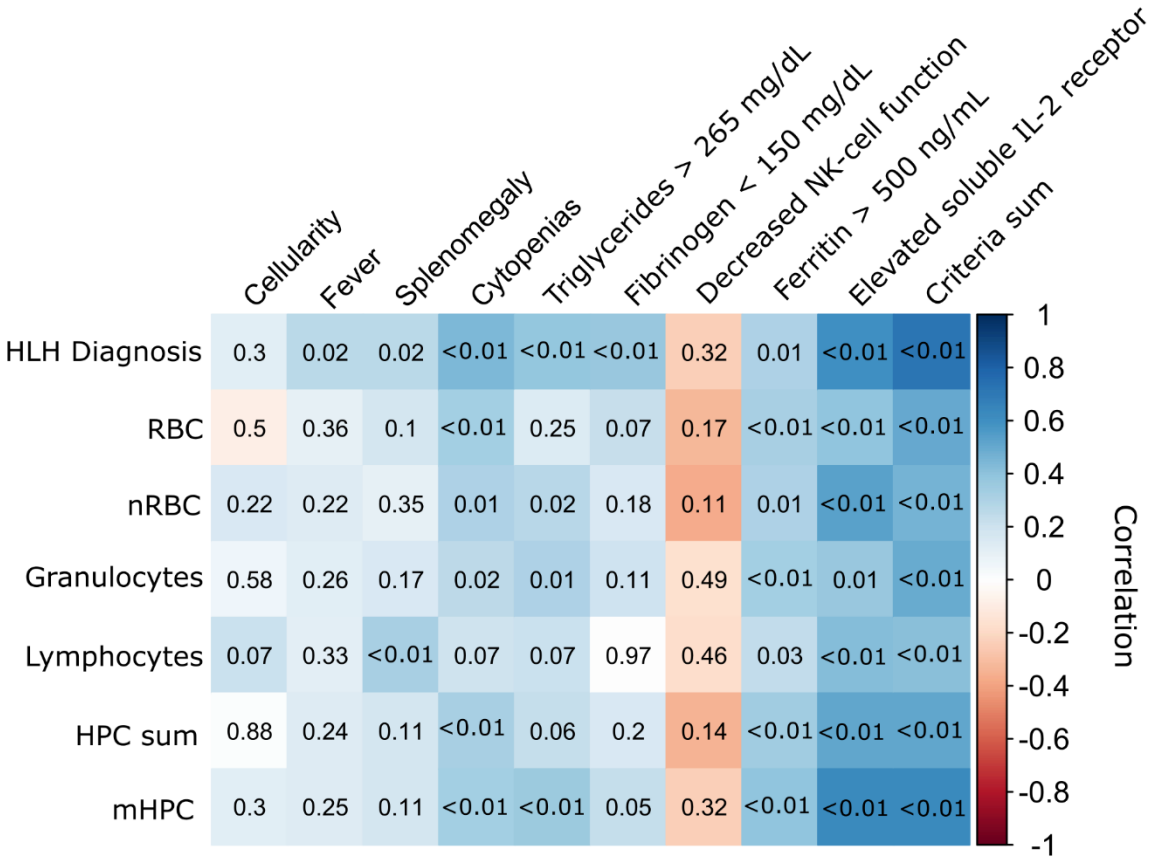
In order to determine a threshold for each lineage which best distinguished between HLH and non-HLH patients, R's OptimalCutpoints package was implemented based on Youden's index[1] Youden's index maximizes the difference between the true positive and false positive rate; this index produces an optimal cut point independent of the prevalence.[2] The area under the curve (AUC) and 95% confidence intervals (CI) were reported for each lineage based on the determined threshold.

Classification and regression trees (CART) were implemented to determine whether a combination of lineages was better able to distinguish between HLH and non-HLH than each lineage individually. CART is a tree-based learning technique for classifying observations.[3] All four dichotomized lineages (based on the thresholds from the Youden analysis) were entered into CART using the ctree function in the party package, and only important lineages were included in the final tree.[4] Two CARTs were fit: one which included all four dichotomized lineages, and one excluding lymphocytes. To ensure that the final models were valid, we conducted a 5 times repeated 10-fold cross validation (CV) for both CARTs.[5] The CV AUCs and 95% CIs were reported.

A supplemental analysis was also conducted entering the lineages as continuous variables into two CART models, with and without lymphocytes to see whether CART would choose different cut points for lineages selected than cut points determined in the Youden analysis. A p-value < 0.05 was considered statistically significant. All analyses were conducted using R software v3.2.2.[6]

Supplemental references

1. OptimalCutpoints: An R Package for Selecting Optimal Cutpoints in Diagnostic Tests | López-Ratón | Journal of Statistical Software [Epub ahead of print].
2. Youden WJ. Index for rating diagnostic tests. Cancer 1950;3(1):32–35.
3. Breiman L, Friedman J, Stone C, Olshen R. Classification and Regression Trees. Belmont, CA: Wadsworth; 1984.
4. Hothorn T, Hornik K, Zeileis A. Unbiased Recursive Partitioning: A Conditional Inference Framework. J Comput Graph Stat 2006;15(3):651–674.
5. Kohavi R. A study of cross-validation and bootstrap for accuracy estimation and model selection. In: Proceedings of the 14th International Joint Conference on Artificial Intelligence. 1995. p.
6. R Core Team. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing.



Supplemental Figure 1: The value inside of each box represents the significance (p-value) of the correlation (Spearman's rank). The color of the box represents the direction of the correlation, negative (red) to positive (blue). Lighter shades represent weaker correlations, while darker shades represent stronger correlations.
RBC: red blood cells, *nRBC:* nucleated red blood cells, *HPC sum:* total count from all four lineages. *mHPC:* multiple nucleated cells within a single HPC