UC San Diego
UC San Diego Previously Published Works

Title

Integrative omics to detect bacteremia in patients with febrile neutropenia.

Permalink

https://escholarship.org/uc/item/2r85r39v

Journal

PLoS One, 13(5)

Authors

Kelly, Rachel

Lasky-Su, Jessica

Yeung, Sai-Ching

et al.

Publication Date

2018

DOI

10.1371/journal.pone.0197049

Copyright Information

Peer reviewed

Integrative omics to detect bacteremia in patients with febrile neutropenia

Rachel S. Kelly[1,2‡]**, Jessica Lasky-Su**[1,2‡]**, Sai-Ching J. Yeung**[3]**, Richard M. Stone**[2,4]**, Jeffrey M. Caterino**[5]**, Sean C. Hagan**[6]**, Gary H. Lyman**[7,8]**, Lindsey R. Baden**[2,4]**, Brett E. Glotzbecker**[2,4]**, Christopher J. Coyne**[9]**, Christopher W. Baugh**[2,6‡]**, Daniel J. Pallin**[2,6‡]*

1 Channing Division of Network Medicine, Brigham and Women's Hospital, Boston, MA, United States of America, **2** Harvard Medical School, Boston, MA, United States of America, **3** Department of Emergency Medicine, The University of Texas MD Anderson Cancer Center, Houston, TX, United States of America, **4** Dana Farber Cancer Institute, Boston, MA, United States of America, **5** Ohio State University Medical School, Wexner Medical Center Department of Emergency Medicine, Columbus, OH, United States of America, **6** Department of Emergency Medicine, Brigham and Women's Hospital, Boston, MA, United States of America, **7** Fred Hutchinson Cancer Research Center, Seattle, WA, United States of America, **8** School of Medicine, University of Washington, Seattle, WA, United States of America, **9** University of California, San Diego, School of Medicine and Department of Emergency Medicine, San Diego, CA, United States of America

‡ RSK and JL-S are co-first authors. CWB and DJP are co-senior authors.
* dpallin@bwh.harvard.edu

Abstract

Background

Cancer chemotherapy-associated febrile neutropenia (FN) is a common condition that is deadly when bacteremia is present. Detection of bacteremia depends on culture, which takes days, and no accurate predictive tools applicable to the initial evaluation are available. We utilized metabolomics and transcriptomics to develop multivariable predictors of bacteremia among FN patients.

Methods

We classified emergency department patients with FN and no apparent infection at presentation as bacteremic (cases) or not (controls), according to blood culture results. We assessed relative metabolite abundance in plasma, and relative expression of 2,560 immunology and cancer-related genes in whole blood. We used logistic regression to identify multivariable predictors of bacteremia, and report test characteristics of the derived predictors.

Results

For metabolomics, 14 bacteremic cases and 25 non-bacteremic controls were available for analysis; for transcriptomics we had 7 and 22 respectively. A 5-predictor metabolomic model had an area under the receiver operating characteristic curve of 0.991 (95%CI: 0.972,1.000), 100% sensitivity, and 96% specificity for identifying bacteremia. Pregnenolone steroids were more abundant in cases and carnitine metabolites were more abundant in controls. A 3-predictor gene expression model had corresponding results of 0.961 (95% CI: 0.896,1.000), 100%, and 86%. Genes involved in innate immunity were differentially expressed.

Conclusions

Classifiers derived from metabolomic and gene expression data hold promise as objective and accurate predictors of bacteremia among FN patients without apparent infection at presentation, and can provide insights into the underlying biology. Our findings should be considered illustrative, but may lay the groundwork for future biomarker development.

Introduction

Bacteremia due to chemotherapy-associated febrile neutropenia (FN) is one of the most deadly oncologic emergencies, with mortality rates of up to 50%.[1, 2] Therefore, presumptive broad-spectrum antibiotic treatment is recommended for all FN patients within 1 hour of onset of symptoms, and the vast majority are admitted to the hospital.[3–6] However, fewer than 25% of FN patients are found to have bacteremia.[1, 3, 4] Antibiotic treatment carries risks including *C. difficile* colitis, selection of drug-resistant strains, drug toxicity, allergic reactions, and drug-drug interactions, while hospitalization confers risks of acquisition of nosocomial pathogens, exposure to medical errors, catheter-associated infections, thromboembolism, and financial burdens.[3, 4, 7–9]

Currently, detection of bacteremia depends on culture, which rules out bacteremia only after several days; too late to inform initial decisions regarding hospitalization and therapy. Guidelines recommend that initial treatment decisions be based on clinical evaluation including the Multinational Association for Supportive Care in Cancer (MASCC) score.[4, 10] However, this score was not designed to detect bacteremia, and is insufficiently accurate even for its intended use, prediction of safe discharge, with a negative predictive value for complications of only 83%.[11] Consequently, many clinicians do not rely on it, admitting all FN patients by default.[5, 6] A newer score, the Clinical Index of Stable Febrile Neutropenia, is also inadequate, with a 9.1% rate of bacteremia in the low-risk group.[12] Similarly, PCR for bacterial DNA, and measurement of host markers such as procalcitonin, lack sufficient sensitivity.[13–15] Objective tests are needed to detect bacteremia during the initial evaluation, so that patient-specific management strategies can be employed.

High throughput 'omics' profiling is a powerful tool for the discovery of biomarkers for various conditions, including infectious diseases.[16] Metabolomics, which provides an integrated profile of biological status, reflecting the "net results" of genetic, transcriptomic, proteomic, and, environmental interactions, represents a particularly powerful research tool.[17] Differentially-abundant metabolites may be utilized as biomarkers to discriminate between those with and without bacteremia.[18, 19] By integrating gene-expression data with metabolomic data, we can identify changes in upstream regulators of the metabolites of interest, strengthening pathway analyses and facilitating deeper understanding of underlying biology.[20, 21]

The aim of this study, was to develop methods for the development of omic biomarkers that can accurately identify bacteremia among FN patients without apparent infection. Eventually, we hope to build upon the most promising results from this project to develop a multivariable biomarker that can be advanced into a test for routine clinical use, whilst informing on underlying biology.

Methods

Study population

Between October 2015 and October 2016, we recruited patients with fever (defined as a single oral temperature ≥38.3˚C or an oral temperature ≥38˚C sustained over a 1-hour period)[4,

10] and FN (defined as <1000 neutrophils/μL_2), from the Brigham and Women's Hospital emergency department (receiving emergency department for the Dana Farber Cancer Institute). All patients had detailed clinical evaluations including history, physical exam, chest X-ray, urinalysis, and, blood and urine cultures. Information on antibiotic use in the 24 hours prior to blood collection was also recorded. Patients presenting with focal bacterial infection (pneumonia, skin infection, urinary tract infection, intra-abdominal infection) or evidence of sepsis in the judgment of two investigators were excluded, as the goal of this study was to find an indicator of bacteremia in patients without initially-detectable bacterial infection. Trained research assistants screened for eligible patients between the hours of 7am and 11pm daily. Once subjects were identified, they collected two blood samples (in PAXgene RNA preservation tubes and lithium heparin tubes) and clinical data. Two investigators then independently categorized each subject, relying upon the diagnostic investigation conducted as part of routine care and the recommendations of the Infectious Diseases Society of America and the American Society of Clinical Oncology.[4, 10, 22] The investigators placed each subject into one of three categories: (i) No evidence of focal infection or sepsis at presentation and ultimately found to have been bacteremic (cases), (ii) No evidence of focal infection or sepsis at presentation and ultimately found not to have been bacteremic (controls), or (iii) Evidence of focal infection or sepsis at presentation (excluded),. The Partners Health Care Human Subjects Research Committee approved this study, and all participants provided written informed consent.

Metabolomic and gene expression profiling

Mass spectrometry-based metabolomic profiling was performed on plasma samples by Metabolon, Inc. (Durham, NC), as described previously [23, 24] Briefly, the global biochemical profiling analysis was composed of four unique arms covering a broad range of the metabolome; (i) reverse-phase chromatography positive ionization methods optimized for hydrophilic compounds (LC–MS Pos Polar) and (ii) hydrophobic compounds (LC–MS Pos Lipid); (iii) reverse-phase chromatography with negative ionization conditions (LC–MS Neg), and (iv) a HILIC chromatography method coupled to negative (LC–MS Polar). Metabolites were annotated based on an iterative process of matching on mass to charge ratio, retention time and spectral fragmentation signature, followed by manual curation to confirm biochemical identification. Further details are provided in S1 Methods.

RNAseq could not be used in this population due to the paucity of white blood cells. We used EdgeSeq from HTG Molecular, Inc (Tucson, AZ).[25] to quantify the expression of 2,560 genes relevant to cancer and immunology. This technology does not require an RNA isolation step. The panel of genes was chosen by an iterative process of literature review and key opinion leader feedback, and included 24 gene groups and pathways.[26]. Analysis was performed as described previously [25, 26]. Briefly, the EdgeSeq assay couples quantitative nuclease protection with next-generation sequencing. After allowing nuclease protection probes (NPPs) to hybridize to their target RNAs, S1 nuclease is added to remove excess unhybridized NPPs and RNA, leaving behind only NPPs hybridized to their target RNAs, resulting in a stoichiometric conversion of target RNA to the NPPs and producing a 1:1 ratio of NPP to RNA. The quantitative nuclease protection steps are automated on the EdgeSeq processor, followed by PCR to add sequencing adaptors and tags. The labeled samples are pooled, cleaned, and sequenced on a next-generation sequencing platform using standard protocols. The resulting data are processed and reported by EdgeSeq parsing software. Further details are provided in S1 Methods.

Quality control and data processing

Metabolomics. We used a previously published [27] QC and processing pipeline to clean the metabolomic dataset. We excluded metabolites with zero abundance in all samples, then imputed all remaining missing observations with half the lowest detected value for that metabolite. We considered metabolites with zero variance uninformative and excluded them. We then *pareto* scaled and log-transformed the data.

Gene expression. We performed data processing and normalization according to Edge-Seq manufacturer's standards.[25, 26] Data were transferred from the Illumina MiSeq sequencer as demultiplexed FASTQ files, with one file per original well of the 96-well sample plate. The HTG EdgeSeq Parser was used to align the FASTQ files to the probe list to collate the data, which were then median-normalized.[28]

Statistical analysis

Prediction. We used unsupervised principal components analysis (PCA) to assess the ability of the metabolome and gene expression data to discriminate cases (with bacteremia) from controls (without bacteremia). We further interrogated these plots to determine if other clinical factors may be driving the metabolomic profiles. We then used supervised partial least squares discriminant analysis (PLS-DA) to assess predictive accuracy for bacteremia. Next, we attempted to identify metabolomics and genetic predictive profiles using two approaches; (i) We used independent logistic regression models adjusting for age, sex, body mass index (BMI), and, tumor type (solid/liquid) to identify the metabolites most strongly associated with the presence or absence of bacteremia. Differential gene expression analysis was used for the gene expression data. (ii) We employed least absolute shrinkage and selection (LASSO) sparsity-inducing logistic regression to identify more parsimonious metabolomic and genomic predictors. We ran two logistic models; one containing all metabolites, and one containing all genes to identify the subset of metabolites and genes retained in the models. These were then selected as the predictors. We used the lambda that produced the minimum mean cross validated error.[29]

We then created metabolite and gene summary scores based on (i) the first principal component of metabolites/expressed genes identified as differentially abundant in cases vs. controls in the regression models; and (ii) metabolites/genes selected in the LASSO model. We used receiver operating characteristic (ROC) curve analysis to evaluate the predictive ability of these summary scores for bacteremia, and employed the method of DeLong to compare areas under the ROCcurves for the different scores.[30] We determined a cutoff that maximized sensitivity, and calculated specificity at this cutoff. The currently recommended approach to risk stratification is to classify a patient as high-risk if the MASCC score is <21 or if any of the Infectious Diseases Society of America/American Society of Clinical Oncology high risk criteria are met.[3, 4] While this approach to risk stratification was not designed to detect bacteremia, no other method currently exists to classify these patients. Therefore, we compared the accuracy of this classifier as a predictor of bacteremia to the accuracy of our omic predictors of bacteremia.

Analysis of underlying biology. We explored metabolic pathways using MetaboAnalyst (http://www.metaboanalyst.ca/), which takes both overrepresentation and pathway topology into account, assigning more weight to metabolites that form key components or 'hubs' of specific pathways. For gene expression, we performed gene set enrichment analysis using the g. GOSt tools from the g.profiler package (http://biit.cs.ut.ee/gprofiler/).

In order to combine the findgins from the metabolomics and genetic analysis, we used IMPaLA (Integrated Molecular Pathway Level Analysis; http://impala.molgen.mpg.de/) [31]

to identify pathways that were jointly dysregulated at the level of both metabolites and gene expression. IMPaLA performs over-representation analysis considering both genes and metabolites to provide a combined pathway p-value as well as a q-value that accounts for multiple testing.

Results

Study population

Metabolomic profiling was performed for 58 subjects who were classified by two investigators with no disagreements. Fourteen (24%) had bacteremia (cases); 25 (43%) had no evidence of bacterial infection (controls); and, 19 (33%) had evidence of a focal infection without bacteremia (excluded from analysis), resulting in a total of 39 analyzed subjects (**Table 1**). Gene expression data were available for only seven of the cases and 21 of the controls, due to logistical issues. A further control had gene expression profiling only. In total seven cases and 22 controls were included in the gene expression analysis.

Cases demonstrated a significantly higher maximum temperature, lower neutrophil and lymphocyte counts, and a lower MASCC score as expected. Patients with both solid and liquid tumors originating from a variety of organs were included. Cases were more likely to have a liquid tumor than controls, but there was no significant difference in tumor site. Antibiotic treatment was initiated prior to sample collection in 85% of subjects; there was no significant difference in the proportion of cases and controls who received antibiotics (**Table 1**).

Metabolomics

A total of 1,296 metabolites were measured. After exclusion of metabolites that were missing for all subjects and those with no variance across the population, 1,204 metabolites remained for analysis. These included amino acids, carbohydrates, lipids, nucleotides, vitamins, peptides, energy metabolites, and 163 xenobiotics. PCA based on all 1,204 metabolites revealed separation between cases and controls along the first two components which together explained 27% of the variance in the data (**Figure A in S1 File**). To determine if these metabolomic profiles were driven by other clinical factors we also interrogated the PCA plot in terms of tumor type (liquid or solid); tumor site and antibiotic use prior to blood draw. Among the bacteremic cases only, we also explored the bacteria type subsequently identified in the culture (Gram-positive, Gram-negative or both). These plots indicated no clustering based on any of these variables; and therefore provided no evidence that these factors were driving the metabolomic profiles (**Figure B in S1 File**). Regression models confirmed that both PC1 (p = 8.8×10^{-4}) and PC2 (p = 0.024) were significantly associated with case status.

Partial least squares discriminant analysis (**Fig 1**) suggested that a metabolomic classifier could distinguish between cases and controls, with $R^2 = 0.650$, and a cross-validated Q^2 of 0.410 for the first component. Interrogation of the variable importance in the projection plot (VIP; a measure of the relative importance of each feature in the PLS-DA) identified 17α-hydroxypregnanolone glucuronide, estrone 3-sulfate, 5α-pregnan-3, 20β-diol disulfate and pregn steroid monosulfate (C21H3405S) as the top metabolites driving the discrimination. Carnitines, also had high VIP scores.

Permutation testing revealed that the model was not robust (p = 0.366). Therefore, a refined discriminatory profile was sought by identifying metabolites significantly associated with bacteremia using multivariable logistic regression. After adjustment for age, sex, BMI and, tumor type (liquid or solid), a total of 177 metabolites were significant at p<0.05 and 19 were significant at a p<0.01 (**Fig 2** and **Table A in S1 File**). A majority of the significant metabolites were lipids. We observed upregulation of pregnenolone steroids and downregulation of carnitine

Table 1. Baseline characteristics of the study population.

		Cases (n = 14)		Controls (n = 25)		p-value
Sex	Female	4	28.6%	13	52.0%	0.193
	Male	10	71.4%	12	48.0%	
Age (years)	mean (SD)	55.1 (11.6)		47.0 (16.4)		0.082
BMI	mean (SD)	26.4 (2.7)		25.3 (4.3)		0.310
Tmax (°F)	mean (SD)	102.2 (0.8)		101.3 (0.8)		2×10^{-3}*
Absolute Neutrophil Count	mean (SD)	0.12 (0.22)		0.35 (0.31)		0.012*
Absolute Lymphocyte Count	mean (SD)	0.21 (0.20)		0.56 (0.51)		0.005*
MASCC risk score	mean (SD)	15.0 (4.3)		18.1 (3.8)		0.032
MASCC high risk	≥21 (low risk)	1	7.1%	8	32.0%	0.120
	<21 (high risk)	13	92.9%	17	68.0%	
ASCO	high	14	100.0%	22	88.0%	0.540
	low	0	0.0%	3	12.0%	
IDSA	high	13	92.9%	20	80.0%	0.391
	low	1	7.1%	5	20.0%	
Tumor type	liquid	12	85.7%	11	44.0%	0.017
	solid	2	14.3%	14	56.0%	
Cancer type[a]	Breast	0	0.0%	3	12.0%	0.441
	Esophageal	0	0.0%	2	8.0%	
	Gynecological	1	7.1%	2	8.0%	
	Hematological	12	85.7%	10	40.0%	
	Lung	0	0.0%	1	4.0%	
	Male reproductive	0	0.0%	2	8.0%	
	Other	1	7.1%	4	16.0%	
	Skin	0	0.0%	1	4.0%	
Antibiotics prior to blood draw[b]	Yes	14	100.0%	19	76.0%	0.071
	No	0	0.0%	6	24.0%	
Organism[c]	Gram-Negative	8	57.1%	-		
	Gram-Positive	5	35.7%	-		
	Both	1	7.1%	-		
Gene Expression Data[d]	Available	7	52.9%	21	84.0%	0.033*

*Indicates a significant difference between cases and controls at a 95% confidence level

Tmax–maximum temperature

MASCC–Multinational Association for Supportive Care in Cancer

ASCO–American Society of Clinical Oncology binary classifier

IDSA–Infectious Diseases Society of America binary classifier

SD–Standard deviation

[a] Information on cancer type was not available for one control

[b] The patient received antibiotics in the 24 hours prior to blood draw

[c] Type of bacteria ultimately identified in the culture samples from bacteremic cases

[d] Transcriptomic analysis was conducted on 7 cases and 22 controls–one of these controls did not have metabolomic profiling available

metabolites among bacteremia cases. The relative metabolite intensities in cases and controls for the top eight upregulated and top eight downregulated metabolites are shown in **Figure C in S1 File**. Pathway analysis identified six metabolic pathways that were enriched among these significant metabolites: pyrimidine metabolism (p = 0.002), ascorbate and aldarate metabolism (p = 0.003), purine metabolism (p = 0.017), sphingolipid metabolism (p = 0.018), pantothenate metabolism (p = 0.022) and valine, leucine and isoleucine metabolism (p = 0.022).
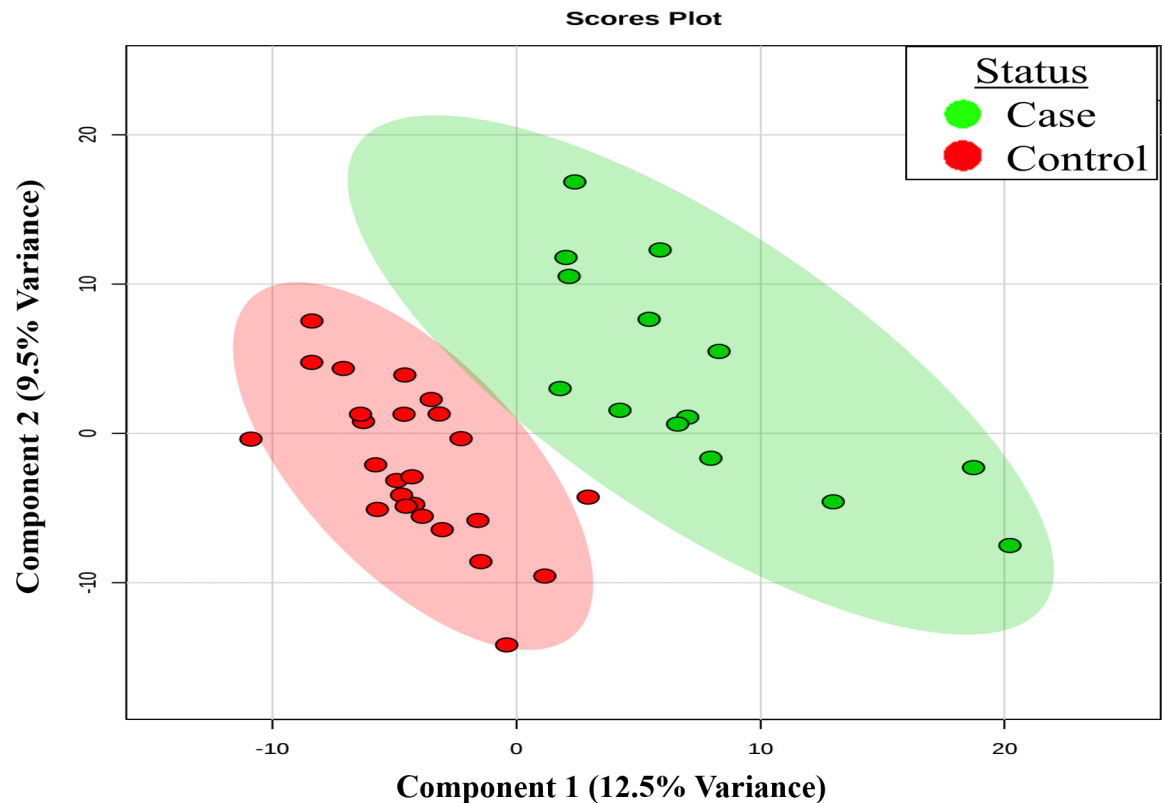
Scores Plot

Fig 1. Partial least square discriminant analysis demonstrating metabolomic differences in bacteremia cases (n = 14) and controls (n = 25). The first two components and the corresponding percentage of the total variance in the metabolome explained by these two components are presented.

https://doi.org/10.1371/journal.pone.0197049.g001

We generated a summary score based on the first principal component of the 177 metabolites, and used ROC curve analysis to explore the discriminatory ability of this score (**Fig 3**). To identify the most parsimonious model, we then used LASSO regression, and identified a five-metabolite score. The AUCs of the two currently available clinical classifiers; MASCC and MASCC plus an indicator of risk ("high-risk" classifier), were 0.624 (95%CI: 0.508–0.741) and 0.540 (95%CI: 0.486–0.594), respectively. Both metabolite scores significantly out-performed these classifiers with AUCs of 0.969 (95%CI: 0.918–1.000) (*p dif* MASCC classifier = 2.3×10^{-8}; *p dif* high risk classifier $<2.2 \times 10^{-16}$) for the standard logistic score and 0.991 (95%CI: 0.972–1.000) (*p dif* MASCC classifier = 8.0×10^{-10}; *p dif* high risk classifier $<2.2 \times 10^{-16}$) for the LASSO score. Furthermore, while the sensitivity of the binary MASCC and "high risk" classifiers was high (93% and 100%, respectively), the corresponding specificities were only 32% and 8%. In contrast, at the cutoff required to achieve 100% sensitivity, the specificity of the standard logistic classifier was 88%, and the parsimonious LASSO classifier was 96% (**Table 2**).

Sensitivity analyses. A sensitivity analysis additionally adjusting for antibiotic use in the 24 hours prior to blood draw identified 118 significant metabolites; 112 (95%) of which were also among the 177 identified in the original analyses. We ran a further sensitivity analysis adjusting for the difference between absolute neutrophil and absolute leukocyte count. Again, of the 169 metabolites that retained significance in this model, 159 (94%) were among the original 177 metabolites., demonstrating the robustness of these findings.
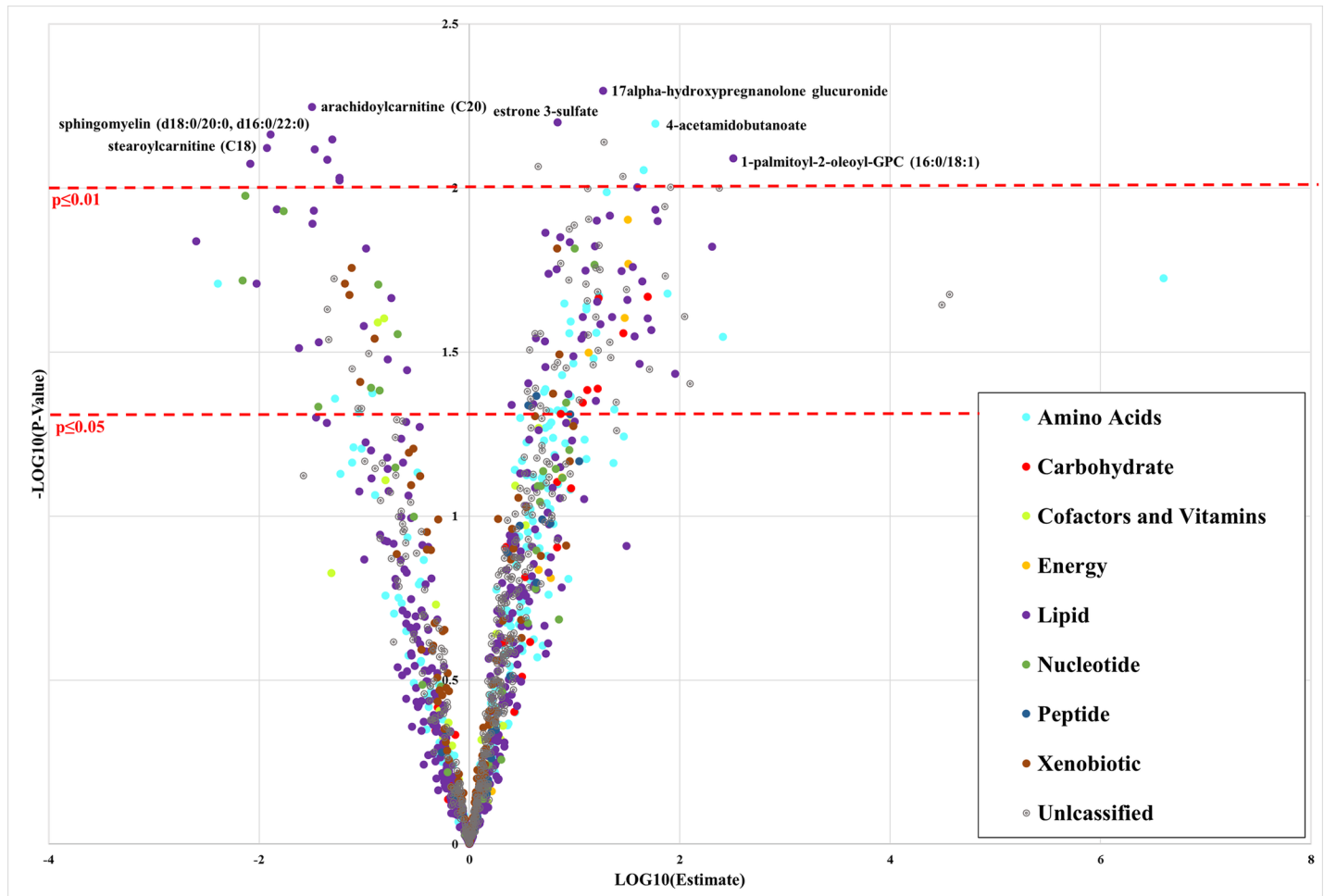
Fig 2. Association between metabolites and bacteremia using a multivariable logistic regression model after adjustment for age, sex, BMI and tumor type (liquid or solid). Metabolites are colored according to their super pathway assignment and top metabolites are named; nominal significance levels of 95% and 99% are indicated with dashed red line; the x axis represents the strength of the association and the y-axis the significance–metabolites to the right of the plot are at higher levels in cases than controls, metabolites to the left are at higher levels among the controls.

https://doi.org/10.1371/journal.pone.0197049.g002

Similarly, we wanted to determine whether the type of bacteria (Gram-negative or Gram-positive) responsible for the bacteremia influenced the results. Although the numbers were too small (n = 5 cases) for the model to converge when considering only at Gram-positive bacteria, we identified 129 metabolites significantly associated with Gram-negative bacteria (n = 8 cases). Ninety-nine (77%) of these were also among the 177 metabolites; including many of the most significant hits such as the carnitines and pregnenolone steroids. Furthermore, when comparing the relative levels of these top metabolites in controls versus cases stratified by Gram-status, both the Gram-negative and Gram positive bacteria cases were distinct from the controls (**Figure D in S1 File**). This again suggests that case-control status rather than bacteria type among the cases was the biggest driver of the differential metabolite abundance.

Gene expression profiling

Evidence of separation in the gene expression profile of cases (n = 7) and controls (n = 22) was also suggested by a PCA model based on 2,560 mRNAs. The PLS-DA model (**Fig 4**) resulted in an R^2 of 0.60, but a Q^2 of only 0.09, and again permutation testing indicated that this model
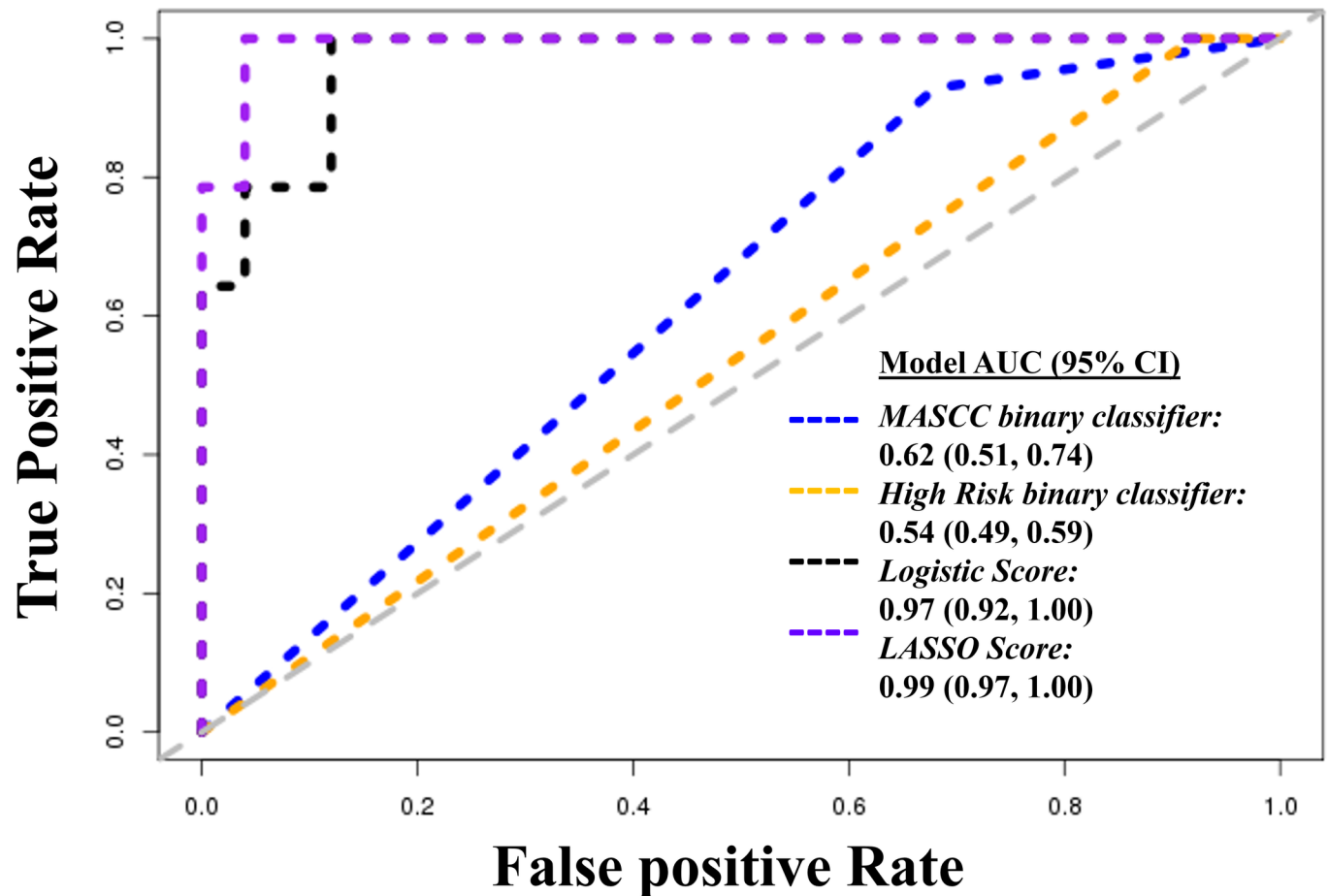
Fig 3. Receiver operating characteristic curves showing the performance of metabolomic (logistic and LASSO) predictors compared to existing clinical (MASCC and high-risk) classifiers. *AUC–Area under the receiver operating characteristic curve. MASCC binary classifier-Multinational Association for Supportive Care in Cancer score categorized into <21 (high risk) and ≥21(low risk). High risk binary classifier- defines a patient as high risk if the MASCC score is <21 or if any of the* Infectious *Diseases Society of America/American Society of Clinical Oncology high risk criteria are met. Logistic score–summary score based on the 137 metabolites associated with bacteremia under a multivariable logistic regression model. LASSO score–summary score based on the seven metabolites associated with bacteremia under a penalized LASSO model.*

https://doi.org/10.1371/journal.pone.0197049.g003

was not robust (p = 0.974). Differential expression analysis identified 150 nominally significant genes, of which three were significant at p<0.01 (**Fig 5**). The top genes included *CTSS* (p = 0.005), *FGL2* (p = 0.006), *LYZ* (p = 0.009), *PPBP* (p = 0.010), *CYBB* (p = 0.011) and *CD86* (p = 0.011). The expression levels of the top eight over-expressed genes and the top eight under expressed-genes in cases versus controls are shown in **Figure D in S1 File.** G.profiler analysis determined that the significant genes were enriched for 24 biological terms, including a number relating to vesicle mediated transport, and cytokines (**Table B in S1 File**). When LASSO regression was employed, only three genes were retained in the model: *RAD18*, which encodes a DNA repair protein, *MAPKAPK3*, a kinase activated in response to cellular stress and *JAG1*, which has a reported role in hematopoiesis. Sensitivity analyses were not performed on the gene-expression data due to sample size limitations.

In this subset of patients the AUC for the MASCC classifier was 0.610 (95%CI: 0.437–0.784) and for the high-risk classifier0.546 (95% CI: 0.484–0.607). Again, the omics-based scores out-performed these metrics. The summary score based on the 150 genes had an AUC of 0.974 (95%: 0.923–1.000) (*p dif* MASCC classifier = 1.21×10^{-4}; *p dif* high risk classifier< 2.20×10^{-16}),

Table 2. Sensitivity and specificity of the metabolomic and gene expression classifiers compared to existing clinical predictors in terms of area under the curve, sensitivity and specificity.

Classifier	AUC	95%CI	Sensitivity	Specificity
Metabolomic Data				
MASSC ≤21	0.624	(0.508, 0.741)	93%	32%
High Risk Classification	0.54	(0.486, 0.594)	100%	8%
Logistic Score optimal cutoff	0.969	(0.918, 1.000)	100%	88%
LASSO Score optimal cutoff	0.991	(0.972, 1.000)	100%	96%
Gene Expression Data				
MASSC ≤21	0.61	(0.437, 0.784)	89%	36%
High Risk Classification	0.546	(0.484, 0.607)	100%	9%
Logistic Score optimal cutoff	0.974	(0.926 1.000)	100%	86%
LASSO Score optimal cutoff	0.961	(0.896, 1.000)	100%	86%

AUC–Area under the receiver operating characteristic curve

MASCC binary classifier–Multinational Association for Supportive Care in Cancer score categorized into <21 (high risk) and ≥21(low risk)

High risk binary classifier–defines a patient as high-risk if the MASCC score is <21 or if any of the Infectious Diseases Society of America/American Society of Clinical Oncology high risk criteria are met

Logistic and LASSO score cutoffs were chosen to obtain 100% sensitivity

https://doi.org/10.1371/journal.pone.0197049.t002

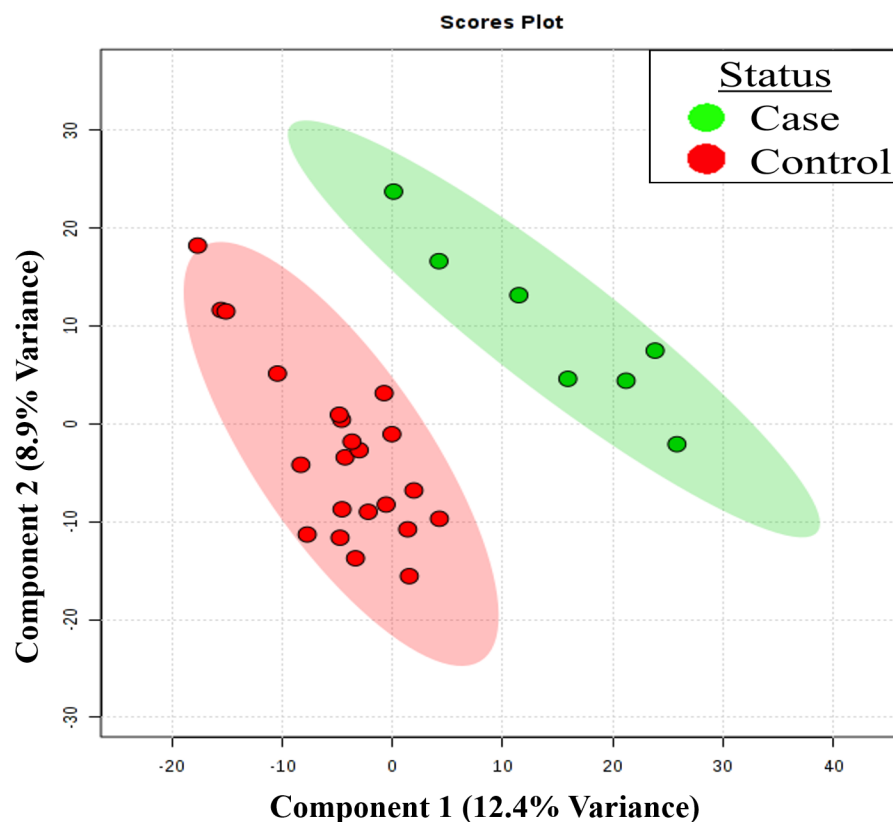**Fig 4. Partial least square discriminant analysis demonstrating differential gene expression in bacteremia cases (n = 9) and controls (n = 21).** Gene expression profiles based on 2560 genes quantified using EdgeSeq; the first two components and the corresponding percentage of variances in the gene expression profile explained by these two components are presented.
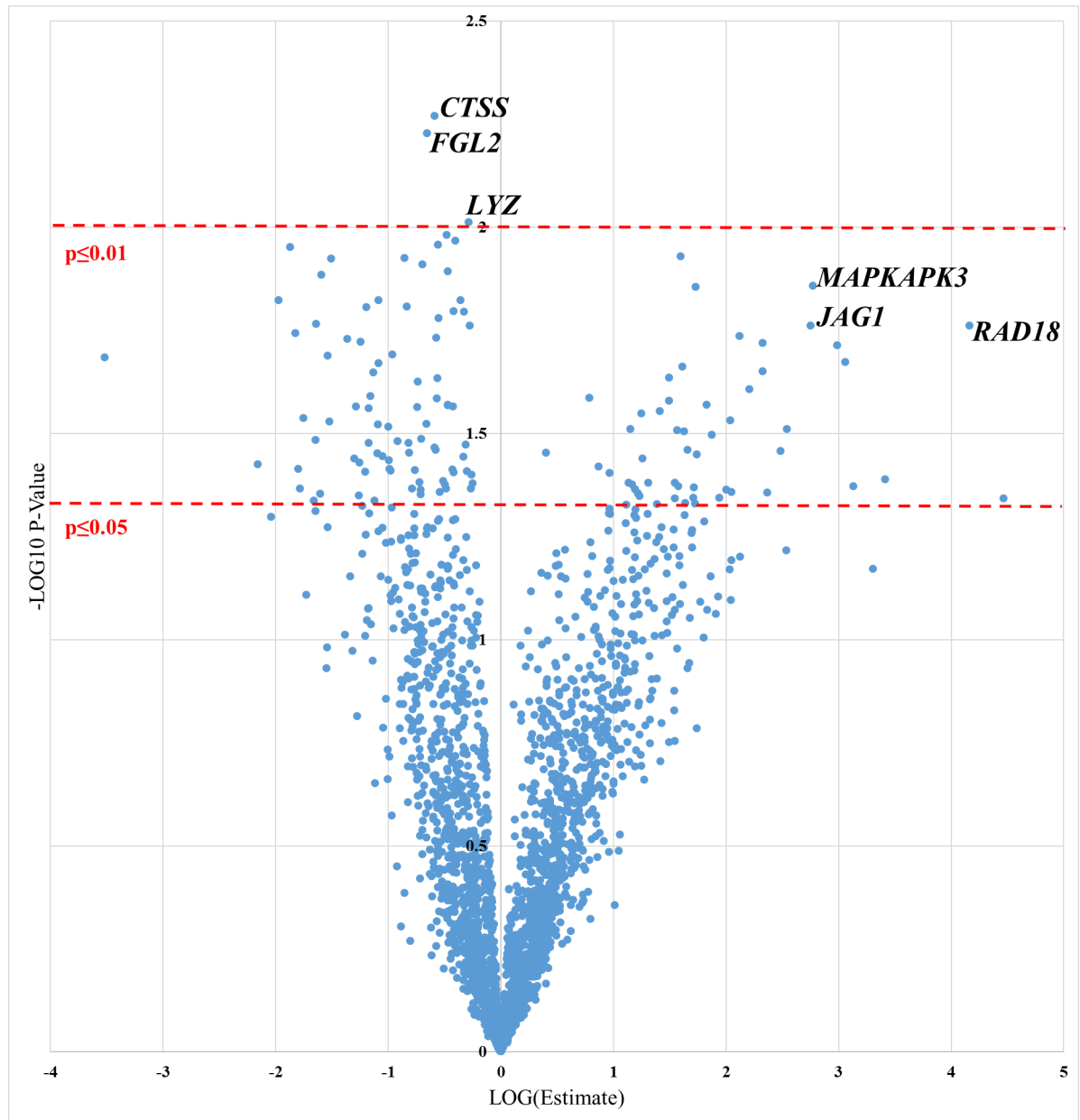
https://doi.org/10.1371/journal.pone.0197049.g004

Fig 5. Association between gene expression and bacteremia using a logistic regression model. Top genes are names; nominal significance levels of 95% and 99% are indicated with dashed red line; the x axis represents the strength of the association and the y-axis the significance–genes to the right of the plot are more highly expressed in cases than controls, genes to the left are more highly expressed among the controls.

https://doi.org/10.1371/journal.pone.0197049.g005

while the AUC for a model including *RAD18* (p in logistic model = 0.017), *MAPKAPK3* (p in logistic model = 0.014) and *JAG1* (p in logistic model = 0.017) as predictors was 0.961 (95%CI: 0.896 1.000) (*p dif* MASCC classifier = 5.51×10^{-6}; *p dif* high risk classifier $<2.20 \times 10^{-16}$) (**Fig 6**). The binary MASCC and high-risk classifiers in this gene expression subpopulation were sensitive (89% and 100%, respectively), but had low specificities (36% and 9%). In contrast, with 100% sensitivity, the specificities was 86% for both our standard logistic and LASSO genomic classifiers (**Table 2**).
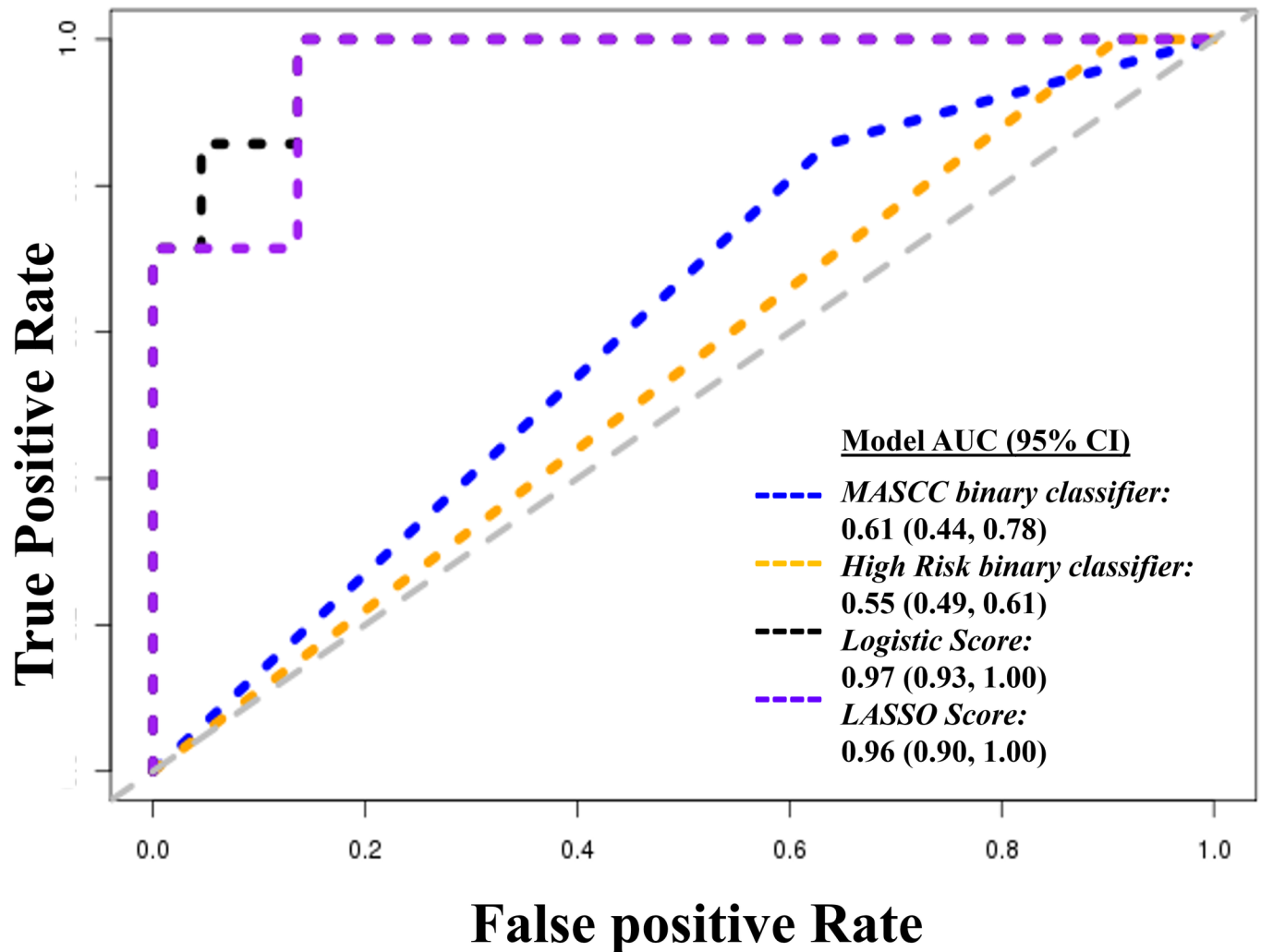
Fig 6. Receiver operating characteristic curves showing the performance of gene expression based (logistic and LASSO) predictors compared to existing clinical (MASCC and high-risk) classifiers. AUC–Area under the receiver operating characteristic curve. MASCC binary classifier–Multinational Association for Supportive Care in Cancer score categorized into <21 (high risk) and ≥21(low risk). High risk binary classifier–defines a patient as high risk if the MASCC score is <21 or if any of the Infectious Diseases Society of America/American Society of Clinical Oncology high risk criteria are met. Logistic score–summary score based on the 153 genes associated with bacteremia under a logistic regression model. LASSO score–summary score based on the 2 genes associated with bacteremia under a penalized LASSO model.

Integrative analysis

When we combined the metabolomic and gene expression data into a single data set, there were too few observations to demonstrate predictive ability, because we had both datasets for only 28 participants. This caused the number of parameters to exceed the number of observations by too wide a margin for our models to converge on a solution. However, integrative analysis using IMPaLA revealed that a number of the genes and metabolites independently identified as significant in the previous analyses are involved in the same biological pathways and processes (**Table 3**). Forty eight pathways were identified as being differentially perturbed in terms of both metabolomics and gene expression (joint q-value<0.05). These included pathways relating to the immune system (joint q-value = 6.29×10^{-8}) insulin regulation (including the Insulin receptor signaling cascade; joint q-value = 6.29×10^{-8}, IRS-related events triggered by IGF1R;joint q-value = 4.33×10^{-3}), the MAPK signaling pathway (joint q-value = 8.38×10^{-3})

Table 3. Pathways differentially enriched at metabolomic and transcriptomic levels.

Process	# Overlapping Genes	Genetic p-value	Genetic q-value	# Overlapping metabolites	Metabolomic p-value	Metabolomic q-value	Joint p-value	Joint q-value
Immune System	52	1.28E-10	5.51E-07	5	8.39E-03	3.52E-01	3.06E-11	6.29E-08
Innate Immune System	37	3.36E-08	2.90E-05	3	1.08E-01	8.46E-01	7.40E-08	7.61E-05
Signaling by NGF	16	1.90E-05	3.53E-03	4	1.56E-03	2.01E-01	5.44E-07	3.73E-04
Cytokine Signaling in Immune system	20	3.36E-07	1.45E-04	1	2.83E-01	1.00E+00	1.64E-06	6.73E-04
Signaling by EGFR	15	2.12E-06	7.64E-04	2	6.94E-02	7.59E-01	2.47E-06	7.25E-04
PI3K-Akt signaling pathway—Homo sapiens (human)	15	4.82E-06	1.39E-03	1	6.43E-02	7.26E-01	4.95E-06	1.27E-03
HIF-1 signaling pathway—Homo sapiens (human)	7	1.37E-04	9.67E-03	3	3.73E-03	3.16E-01	7.88E-06	1.80E-03
Transmembrane transport of small molecules	7	5.36E-01	1.00E+00	15	1.20E-06	4.80E-03	9.84E-06	1.97E-03
Chemokine signaling pathway—Homo sapiens (human)	11	6.61E-06	1.72E-03	1	1.05E-01	8.26E-01	1.05E-05	1.97E-03
Insulin receptor signaling cascade	12	2.05E-05	3.53E-03	2	4.11E-02	6.14E-01	1.26E-05	2.16E-03
Integrated Breast Cancer Pathway	6	8.33E-05	7.26E-03	2	1.53E-02	4.04E-01	1.86E-05	2.83E-03
Adaptive Immune System	20	4.70E-04	2.23E-02	4	2.82E-03	2.55E-01	1.92E-05	2.83E-03
NGF signaling via TRKA from the plasma membrane	14	2.14E-05	3.55E-03	2	7.57E-02	7.59E-01	2.32E-05	3.18E-03
Signal Transduction	45	3.49E-04	1.80E-02	10	5.30E-03	3.24E-01	2.63E-05	3.38E-03
Signaling by Insulin receptor	12	5.19E-05	6.77E-03	2	4.11E-02	6.14E-01	3.00E-05	3.42E-03
Downstream signal transduction	13	3.18E-05	4.91E-03	2	6.33E-02	7.26E-01	2.84E-05	3.42E-03
DAP12 signaling	13	3.64E-05	5.42E-03	2	6.33E-02	7.26E-01	3.22E-05	3.49E-03
Central carbon metabolism in cancer—Homo sapiens (human)	6	9.07E-05	7.26E-03	3	4.24E-02	6.14E-01	5.18E-05	4.33E-03
IRS-mediated signaling	11	9.00E-05	7.26E-03	2	4.11E-02	6.14E-01	5.00E-05	4.33E-03
IRS-related events triggered by IGF1R	11	1.04E-04	8.01E-03	2	4.11E-02	6.14E-01	5.70E-05	4.33E-03
IGF1R signaling cascade	11	1.04E-04	8.01E-03	2	4.11E-02	6.14E-01	5.70E-05	4.33E-03
Signaling by Type 1 Insulin-like Growth Factor 1 Receptor (IGF1R)	11	1.08E-04	8.01E-03	2	4.11E-02	6.14E-01	5.89E-05	4.33E-03
DAP12 interactions	13	5.92E-05	6.91E-03	2	6.33E-02	7.26E-01	5.06E-05	4.33E-03
Signaling by PDGF	13	6.91E-05	7.18E-03	2	6.33E-02	7.26E-01	5.84E-05	4.33E-03
GPCR signaling-G alpha s PKA and ERK	12	7.06E-05	7.18E-03	1	8.48E-02	7.59E-01	7.80E-05	5.53E-03
Diseases of signal transduction	12	2.05E-05	3.53E-03	1	3.30E-01	1.00E+00	8.70E-05	5.81E-03

(Continued)

Table 3. (*Continued*)

Process	# Overlapping Genes	Genetic p-value	Genetic q-value	# Overlapping metabolites	Metabolomic p-value	Metabolomic q-value	Joint p-value	Joint q-value
Chemokine signaling pathway	9	8.44E-05	7.26E-03	1	1.05E-01	8.26E-01	1.12E-04	7.18E-03
MAPK Signaling Pathway	8	5.27E-04	2.42E-02	1	2.19E-02	4.38E-01	1.43E-04	8.38E-03
Pathways in cancer—Homo sapiens (human)	14	1.18E-04	8.55E-03	2	9.55E-02	8.26E-01	1.40E-04	8.38E-03
UMP Synthase Deficiency (Orotic Aciduria)	3	1.96E-03	5.47E-02	5	7.26E-03	3.29E-01	1.73E-04	8.66E-03
Pyrimidine Metabolism	3	1.96E-03	5.47E-02	5	7.26E-03	3.29E-01	1.73E-04	8.66E-03
MNGIE (Mitochondrial Neurogastrointestinal Encephalopathy)	3	1.96E-03	5.47E-02	5	7.26E-03	3.29E-01	1.73E-04	8.66E-03
Dihydropyrimidinase Deficiency	3	1.96E-03	5.47E-02	5	7.26E-03	3.29E-01	1.73E-04	8.66E-03
Beta Ureidopropionase Deficiency	3	1.96E-03	5.47E-02	5	7.26E-03	3.29E-01	1.73E-04	8.66E-03
GPCR signaling-cholera toxin	11	1.90E-04	1.22E-02	1	8.48E-02	7.59E-01	1.94E-04	9.51E-03
Ca-dependent events	1	2.82E-01	8.54E-01	4	5.98E-05	5.40E-02	2.02E-04	9.68E-03
Pyrimidine nucleotides nucleosides metabolism	3	1.94E-02	2.20E-01	6	9.34E-04	1.49E-01	2.16E-04	9.86E-03
Ascorbate and aldarate metabolism—Homo sapiens (human)	1	2.58E-01	8.12E-01	5	8.12E-05	5.40E-02	2.47E-04	1.08E-02
Metabolism	19	8.04E-01	1.00E+00	32	3.36E-05	4.46E-02	3.11E-04	1.33E-02
PLC beta mediated events	2	9.41E-02	4.94E-01	4	3.91E-04	1.20E-01	4.12E-04	1.73E-02
G-protein mediated events	2	9.75E-02	5.04E-01	4	3.91E-04	1.20E-01	4.26E-04	1.75E-02
Hemostasis	20	6.59E-05	7.18E-03	1	6.17E-01	1.00E+00	4.52E-04	1.82E-02
Disease	15	5.10E-04	2.37E-02	3	9.87E-02	8.26E-01	5.48E-04	2.09E-02
HTLV-I infection—Homo sapiens (human)	10	5.54E-04	2.47E-02	1	1.05E-01	8.26E-01	6.24E-04	2.29E-02
GPCR signaling-G alpha s Epac and ERK	10	8.33E-04	3.30E-02	1	8.48E-02	7.59E-01	7.45E-04	2.60E-02
SLC-mediated transmembrane transport	4	3.82E-01	1.00E+00	11	1.98E-04	8.75E-02	7.92E-04	2.67E-02
Pyrimidine catabolism	1	1.24E-01	5.67E-01	5	9.23E-04	1.49E-01	1.15E-03	3.54E-02
PKB-mediated events	3	9.52E-03	1.41E-01	2	1.53E-02	4.04E-01	1.43E-03	4.21E-02

Overlapping–indicates gene/metabolite are found both in the named pathway and among the set identified as significantly associated with bacteremia in this population

Enrichment p and q values are provided for the metabolites alone, the genes alone and the joint enrichment

Q-value–p-value adjusted for the false discovery rate

and multiple pathways involved in the signaling processes necessary for cell survival, differentiation and apoptosis, (such as signaling by NGF; joint q-value = 3.73×10^{-45}, signaling by EGFR; joint q-value = 7.25×10^{-4} andDAP12 signaling; joint q-value = 4.49×10^{-3}).

Discussion

We have demonstrated methods for discovery of a multi-omics-based predictor for detection of bacteremia among FN patients without apparent infection. Omic profiles are increasingly being leveraged as predictive, diagnostic and prognostic biomarkers, and can provide insight into the underlying biology.[32, 33] Metabolomic and transcriptomic diagnostics have already been deployed clinically. For example, real-time analysis of the metabolome is a clinical reality, as exemplified by the iKnife, which performs instantaneous analysis of the mucosal lipidome to phenotype colorectal cancer during surgery.[34] Similarly, transcriptomics have yielded diagnostic tests that are approaching clinical implementation for detection of serious bacterial infection.[35, 36] To our knowledge, this was the first study to develop multi-omic based bio-markers of bacteremic FN; one prior study used metabolomics alone to investigate infection in the setting of FN.[37] Our future work will apply these methods to a larger sample size, sepa-rated into derivation and validation sets, with the goal of developing a clinically-applicable tool for detection of bacteremia during the initial evaluation.

For detection of bacteremia, the two existing clinical predictors; MASCC and the high-risk classifier, were sensitive but demonstrated poor specificity. In contrast, although derived from only a small population, our metabolomic and transcriptomic predictors maintained impres-sive specificities with 100% sensitivity. These results demonstrate that derivation of omics-based predictors of bacteremia in the setting of FN is feasible, and justify further research in a larger sample.

We found that pregnenolone steroids, which are cortisol precursors, were upregulated in cases relative to controls. Prior research has linked sepsis to an overexpression of cortisol pre-cursors despite normal cortisol levels, implying decreased activity of 3-beta-hydroxysteroid dehydrogenase.[38] Carnitines were down-regulated in cases, which is in agreement with pre-vious evidence for a differential abundance of carnitines in bacteremic compared to non-bac-teremic patients.[39, 40] L-carnitine has entered clinical trials as a therapeutic agent for patients with sepsis, and metabolomic analysis has been used to identify the subset of patients responding.[41] Other pathways involved in amino acid metabolism were also perturbed, which could relate to the enhanced extraction of amino acids by the liver noted in patients with sepsis and systemic inflammatory response syndrome.[42] For example, pyrimidine metabolism was differentially regulated, possibly due to the *de novo* synthesis of pyrimidines required for successful proliferation of pathogens in blood.[43] Furthermore, the dysregulation of the ascorbate and aldarate metabolism pathway may relate to the lower circulating levels of ascorbate reported in patients with sepsis.[44]

A number of genes showed differential expression between cases and controls. These included, *PPBP*, an antimicrobial protein with bactericidal and antifungal activity; *CYBB*, which is essential for the microbicidal oxidase system of phagocytes; *LYZ*, which encodes a component of the innate immune system that cleaves peptidoglycan bonds in the bacterial cell wall;,and *CD86* and *FGL2* which have been associated with severity and worse outcomes in sepsis[45, 46] These results suggest that a weaker innate immune response might predispose to bacteremia after depletion of the adaptive immune system by chemotherapy. On a pathway level, the differentially-expressed genes were enriched for a number biological processes including some relating to the release of cytokines, supporting the idea of a unique immuno-logic milieu in FN patients with bacteremia.

Integrative analysis supported the findings of the individual omics analysis, and provided mechanistic links between the metabolites and genes independently associated with bacter-emia. For example, MAPK signaling pathways were perturbed. MAPKs play an important role in the cascade of cellular responses evoked by extracellular stimuli such as pro-inflammatory

cytokines or physical stress, and they have been shown to be activated in the setting of bacterial challenge.[47] These integrative results should be viewed with caution due to the limited sample size, which was constrained by funding. However they do provide a good indication of potential gene-metabolite relationships, and a demonstration that in a larger sample, dysregulation at multiple omic levels can be explored simultaneously to interrogate underlying mechanisms.

Regarding limitations, we caution the reader that the sample size, particularly for the transcriptomic data was limited, and we were unable to stratify by potential effect modifiers such as sex.These results should therefore be considered illustrative, and the specific biological entities identified as differentially abundant should be considered exploratory.RNAseq could not be used to quantify expression of the entire transcriptome, because leukocyte-poor blood has insufficient cells for this technique. Therefore, we used EdgeSeq, which does not require an RNA isolation step, but quantifies relative abundance of only 2,560 transcripts. However, the genes included in the EdgeSeq panel were selected on the basis of their relevance to both cancer and the immune system, making them ideally suited to this study population. Because participants had FN as an unscheduled emergency, their dietary intake prior to presenting in the Emergency room may have influenced their metabolome, but this represents the scenario in which a predictor of bacteremia would ultimately be utilized. However, we were able to provide evidence that prior antibiotic use was not influencing our results; nor was the type or site of the initial tumor. Furthermore, within the limited sample size we were able to demonstrate that the metabolomic profile appeared to be similar for Gram-negative and Gram-positive bacteremia,although it would be of interest to explore this further in the larger sample.

In conclusion, we generated metabolomic and transcriptomic predictors of bacteremia, among FN patients without apparent infection at presentation. With overfitting as a caveat, these predictors significantly outperformed currently-recommended risk-stratification tools, with markedly improved specificity and perfect sensitivity. Interrogation of differentially-abundant biomarkers revealed biologically-plausible roles in bacteremia within the setting of FN. This is the first such study within a field that is in dire need of novel biomarkers and management strategies. Via further study in a larger sample with discovery and validation sets, development of a biologically meaningful objective predictor that utilizes both clinical and omics data is feasible, and will facilitate early diagnosis. This, in turn, will enable appropriate aggressive treatment for patients predicted to have bacteremia, and appropriate conservative treatment of those predicted not to have bacteremia.

Supporting information

S1 Methods. Supplementary methods.
(DOCX)

S1 File. Table A: Metabolites significantly (p≤0.01) associated with bacteremia. *Effect estimates adjusted for age, sex, BMI and tumor type (liquid or solid). Metabolites with an X- prefix are awaiting annotation.* **Table B; Biological processes and pathways enriched among 150 genes significantly associated with bacteremia.** *Enrichment analysis performed using the g. GOSt tools from the g.profiler package (http://biit.cs.ut.ee/gprofiler/) p-values are Bonferonni corrected.* **Figure A: Metabolomic PCA by Case-Control Status. Figure B: Metabolomic PCA by Other Clinical Variables. Figure C: Relative Metabolite Intensity levels in Bacteremia Cases and Controls for the Top Eight Upregulated and Top Eight Downregulated Metabolites. Figure D: Relative Metabolite Intensity levels in Bacteremia Cases stratified by Gram status (negative or positive) and Controls (CO) for the Top Eight Upregulated and Top Eight Downregulated Metabolites. Figure E: Expression levels in Bacteremia Cases and**

Controls for the Top Eight Overexpressed and Top Eight Unexpressed Genes.
(DOCX)

S1 Dataset. Gene expression data/metabolomics data and the relevant epidemiological and clinical variables for 29 subjects.
(CSV)

S2 Dataset. Metabolomics data and the relevant epidemiological and clinical variables for 39 subjects.
(CSV)

S3 Dataset. Biochemical names of the metabolites profiled.
(CSV)

Acknowledgments

Author Contributions

References

1. Weycker D, Barron R, Kartashov A, Legg J, Lyman GH. Incidence, treatment, and consequences of chemotherapy-induced febrile neutropenia in the inpatient and outpatient settings. Journal of oncology pharmacy practice: official publication of the International Society of Oncology Pharmacy Practitioners. 2014; 20(3):190–8. Epub 2013/07/05. https://doi.org/10.1177/1078155213492450 PMID: 23824496.

2. Thursky KA, Worth LJ. Can mortality of cancer patients with fever and neutropenia be improved? Current opinion in infectious diseases. 2015; 28(6):505–13. Epub 2015/09/17. https://doi.org/10.1097/QCO.0000000000000202 PMID: 26374951.

3. Freifeld AG, Bow EJ, Sepkowitz KA, Boeckh MJ, Ito JI, Mullen CA, et al. Clinical practice guideline for the use of antimicrobial agents in neutropenic patients with cancer: 2010 Update by the Infectious Diseases Society of America. Clinical infectious diseases: an official publication of the Infectious Diseases

Society of America. 2011; 52(4):427–31. Epub 2011/01/06. https://doi.org/10.1093/cid/ciq147 PMID: 21205990.

4. Flowers CR, Seidenfeld J, Bow EJ, Karten C, Gleason C, Hawley DK, et al. Antimicrobial prophylaxis and outpatient management of fever and neutropenia in adults treated for malignancy: American Society of Clinical Oncology clinical practice guideline. J Clin Oncol. 2013; 31(6):794–810. Epub 2013/01/16. https://doi.org/10.1200/JCO.2012.45.8661 PMID: 23319691.

5. Baugh CW, Brooks GA, Reust AC, Wang TJ, Caterino JM, Baker ON, et al. Provider familiarity with specialty society guidelines for risk stratification and management of patients with febrile neutropenia. The American journal of emergency medicine. 2016; 34(8):1704–5. Epub 2016/06/06. https://doi.org/10.1016/j.ajem.2016.05.047 PMID: 27262602.

6. Baugh CW, Wang TJ, Caterino JM, Baker ON, Brooks GA, Reust AC, et al. Emergency Department Management of Patients With Febrile Neutropenia: Guideline Concordant or Overly Aggressive? Academic emergency medicine: official journal of the Society for Academic Emergency Medicine. 2017; 24 (1):83–91. Epub 2016/09/10. https://doi.org/10.1111/acem.13079 PMID: 27611638.

7. Ahn J, Yu K, Stolzenberg-Solomon R, Simon KC, McCullough ML, Gallicchio L, et al. Genome-wide association study of circulating vitamin D levels. Hum Mol Gene. 2010; 19. https://doi.org/10.1093/hmg/ddq155 PMID: 20418485

8. Lathia N, Mittmann N, DeAngelis C, Knowles S, Cheung M, Piliotis E, et al. Evaluation of direct medical costs of hospitalization for febrile neutropenia. Cancer. 2010; 116(3):742–8. Epub 2009/12/24. https://doi.org/10.1002/cncr.24773 PMID: 20029970.

9. Day E, Kim S, Hughes-Davies L. Barriers to protocol-led early discharge of low-risk febrile neutropenia patients. Clinical oncology (Royal College of Radiologists (Great Britain)). 2014; 26(8):516. Epub 2014/06/10. https://doi.org/10.1016/j.clon.2014.05.003 PMID: 24909700.

10. Freifeld AG, Bow EJ, Sepkowitz KA, Boeckh MJ, Ito JI, Mullen CA, et al. Clinical practice guideline for the use of antimicrobial agents in neutropenic patients with cancer: 2010 update by the infectious diseases society of america. Clin Infect Dis. 2011; 52(4):e56–93. https://doi.org/10.1093/cid/cir073 PMID: 21258094.

11. Bitar RA. Utility of the Multinational Association for Supportive Care in Cancer (MASCC) Risk Index Score as a Criterion for Nonadmission in Febrile Neutropenic Patients with Solid Tumors. The Permanente journal. 2015; 19(3):37–47. Epub 2015/07/16. https://doi.org/10.7812/TPP/14-188 PMID: 26176568; PubMed Central PMCID: PMCPMC4500479.

12. Coyne CJ, Le V, Brennan JJ, Castillo EM, Shatsky RA, Ferran K, et al. Application of the MASCC and CISNE Risk-Stratification Scores to Identify Low-Risk Febrile Neutropenic Patients in the Emergency Department. Ann Emerg Med. 2016. https://doi.org/10.1016/j.annemergmed.2016.11.007 PMID: 28041827.

13. Yanagihara K, Kitagawa Y, Tomonaga M, Tsukasaki K, Kohno S, Seki M, et al. Evaluation of pathogen detection from clinical samples by real-time polymerase chain reaction using a sepsis pathogen DNA detection kit. Critical care (London, England). 2010; 14(4):R159. Epub 2010/08/25. https://doi.org/10.1186/cc9234 PMID: 20731880; PubMed Central PMCID: PMCPMC2945143.

14. Idelevich EA, Silling G, Niederbracht Y, Penner H, Sauerland MC, Tafelski S, et al. Impact of multiplex PCR on antimicrobial treatment in febrile neutropenia: a randomized controlled study. Medical microbiology and immunology. 2015; 204(5):585–92. Epub 2015/01/13. https://doi.org/10.1007/s00430-014-0385-7 PMID: 25573349.

15. Reers Y, Idelevich EA, Patkau H, Sauerland MC, Tafelski S, Nachtigall I, et al. Multiplex PCR assay underreports true bloodstream infections with coagulase-negative staphylococci in hematological patients with febrile neutropenia. Diagnostic microbiology and infectious disease. 2016; 85(4):413–5. Epub 2016/05/26. https://doi.org/10.1016/j.diagmicrobio.2016.04.024 PMID: 27220608.

16. Pacchiarotta T, Deelder AM, Mayboroda OA. Metabolomic investigations of human infections. Bioanalysis. 2012; 4(8):919–25. Epub 2012/04/27. https://doi.org/10.4155/bio.12.61 PMID: 22533566.

17. Nicholson JK, Wilson ID. Opinion: understanding 'global' systems biology: metabonomics and the continuum of metabolism. Nat Rev Drug Discov. 2003; 2(8):668–76. https://doi.org/10.1038/nrd1157 PMID: 12904817.

18. Psychogios N, Hau DD, Peng J, Guo AC, Mandal R, Bouatra S, et al. The human serum metabolome. PloS one. 2011; 6(2):e16957. Epub 2011/03/02. https://doi.org/10.1371/journal.pone.0016957 PMID: 21359215; PubMed Central PMCID: PMC3040193.

19. Rhee EP, Gerszten RE. Metabolomics and cardiovascular biomarker discovery. Clin Chem. 2012; 58 (1):139–47. https://doi.org/10.1373/clinchem.2011.169573 PMID: 22110018.

20. Bartel J, Krumsiek J, Schramm K, Adamski J, Gieger C, Herder C, et al. The Human Blood Metabolome-Transcriptome Interface. PLoS Genet. 2015; 11(6):e1005274. Epub 2015/06/19. https://doi.org/10.1371/journal.pgen.1005274 PMID: 26086077; PubMed Central PMCID: PMCPMC4473262.

21. Robles AI, Harris CC. Integration of multiple "OMIC" biomarkers: A precision medicine strategy for lung cancer. Lung Cancer. 2017; 107:50–8. Epub 2016/06/28. https://doi.org/10.1016/j.lungcan.2016.06.003 PMID: 27344275; PubMed Central PMCID: PMCPMC5156586.

22. Baron EJ, Miller JM, Weinstein MP, Richter SS, Gilligan PH, Thomson RB Jr., et al. A guide to utilization of the microbiology laboratory for diagnosis of infectious diseases: 2013 recommendations by the Infectious Diseases Society of America (IDSA) and the American Society for Microbiology (ASM)(a). Clinical infectious diseases: an official publication of the Infectious Diseases Society of America. 2013; 57(4): e22–e121. Epub 2013/07/13. https://doi.org/10.1093/cid/cit278 PMID: 23845951; PubMed Central PMCID: PMCPMC3719886.

23. Sha W, da Costa KA, Fischer LM, Milburn MV, Lawton KA, Berger A, et al. Metabolomic profiling can predict which humans will develop liver dysfunction when deprived of dietary choline. FASEB journal: official publication of the Federation of American Societies for Experimental Biology. 2010; 24(8):2962–75. https://doi.org/10.1096/fj.09-154054 PMID: 20371621; PubMed Central PMCID: PMC2909293.

24. Guo L, Milburn MV, Ryals JA, Lonergan SC, Mitchell MW, Wulff JE, et al. Plasma metabolomic profiles enhance precision medicine for volunteers of normal health. Proceedings of the National Academy of Sciences of the United States of America. 2015; 112(35):E4901–E10. https://doi.org/10.1073/pnas.1508425112 PMC4568216. PMID: 26283345

25. Girard L, Rodriguez-Canales J, Behrens C, Thompson DM, Botros IW, Tang H, et al. An Expression Signature as an Aid to the Histologic Classification of Non-Small Cell Lung Cancer. Clin Cancer Res. 2016; 22(19):4880–9. Epub 2016/06/30. https://doi.org/10.1158/1078-0432.CCR-15-2900 PMID: 27354471; PubMed Central PMCID: PMCPMC5492382.

26. HTGMolecular. HTG EdgeSeq Oncology Biomarker Panel 2017 [7/13/2017]. Available from: https://www.htgmolecular.com/products/mrna-oncology-biomarkers

27. Blighe K, Chawes BL, Kelly RS, Mirzakhani H, McGeachie M, Litonjua AA, et al. Vitamin D prenatal programming of childhood metabolomics profiles at age 3 y. Am J Clin Nutr. 2017; 106(4):1092–9. Epub 2017/08/25. https://doi.org/10.3945/ajcn.117.158220 PMID: 28835366; PubMed Central PMCID: PMCPMC5611786.

28. Risso D, Ngai J, Speed TP, Dudoit S. Normalization of RNA-seq data using factor analysis of control genes or samples. Nat Biotech. 2014; 32(9):896–902. https://doi.org/10.1038/nbt.2931 http://www.nature.com/nbt/journal/v32/n9/abs/nbt.2931.html#supplementary-information. PMID: 25150836

29. Tibshirani R. Regression shrinkage and selection via the lasso: a retrospective. Journal of the Royal Statistical Society: Series B (Statistical Methodology). 2011; 73(3):273–82. https://doi.org/10.1111/j.1467-9868.2011.00771.x

30. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics. 1988; 44(3):837–45. PMID: 3203132.

31. Kamburov A, Cavill R, Ebbels TM, Herwig R, Keun HC. Integrated pathway-level analysis of transcriptomics and metabolomics data with IMPaLA. Bioinformatics (Oxford, England). 2011; 27(20):2917–8. Epub 2011/09/07. https://doi.org/10.1093/bioinformatics/btr499 PMID: 21893519.

32. Koen N, Du Preez I, Loots du T. Metabolomics and Personalized Medicine. Advances in protein chemistry and structural biology. 2016; 102:53–78. Epub 2016/02/02. https://doi.org/10.1016/bs.apcsb.2015.09.003 PMID: 26827602.

33. van Rensburg IC, Loxton AG. Transcriptomics: the key to biomarker discovery during tuberculosis? Biomarkers in medicine. 2015; 9(5):483–95. Epub 2015/05/20. https://doi.org/10.2217/bmm.15.16 PMID: 25985177.

34. Alexander J, Gildea L, Balog J, Speller A, McKenzie J, Muirhead L, et al. A novel methodology for in vivo endoscopic phenotyping of colorectal cancer based on real-time analysis of the mucosal lipidome: a prospective observational study of the iKnife. Surg Endosc. 2017; 31(3):1361–70. https://doi.org/10.1007/s00464-016-5121-5 PMID: 27501728; PubMed Central PMCID: PMCPMC5315709.

35. Kaforou M, Herberg JA, Wright VJ, Coin LJM, Levin M. Diagnosis of Bacterial Infection Using a 2-Transcript Host RNA Signature in Febrile Infants 60 Days or Younger. JAMA. 2017; 317(15):1577–8. https://doi.org/10.1001/jama.2017.1365 PMID: 28418473.

36. Herberg JA, Kaforou M, Wright VJ, Shailes H, Eleftherohorinou H, Hoggart CJ, et al. Diagnostic Test Accuracy of a 2-Transcript Host RNA Signature for Discriminating Bacterial vs Viral Infection in Febrile Children. JAMA. 2016; 316(8):835–45. https://doi.org/10.1001/jama.2016.11236 PMID: 27552617.

37. Richter ME, Neugebauer S, Engelmann F, Hagel S, Ludewig K, La Rosée P, et al. Biomarker candidates for the detection of an infectious etiology of febrile neutropenia. Infection. 2016; 44(2):175–86. https://doi.org/10.1007/s15010-015-0830-6 PMID: 26275448

38. Khashana A, Ojaniemi M, Leskinen M, Saarela T, Hallman M. Term neonates with infection and shock display high cortisol precursors despite low levels of normal cortisol. Acta paediatrica (Oslo, Norway: 1992). 2016; 105(2):154–8. Epub 2015/11/06. https://doi.org/10.1111/apa.13257 PMID: 26537554.

39. Nanni G, Pittiruti M, Giovannini I, Boldrini G, Ronconi P, Castagneto M. Plasma carnitine levels and urinary carnitine excretion during sepsis. JPEN Journal of parenteral and enteral nutrition. 1985; 9(4):483–90. Epub 1985/07/01. https://doi.org/10.1177/0148607185009004483 PMID: 3928925.

40. Famularo G, de Simone C, Trinchieri V, Mosca L. Carnitines and Its Congeners: A Metabolic Pathway to the Regulation of Immune Response and Inflammation. Annals of the New York Academy of Sciences. 2004; 1033(1):132–8. https://doi.org/10.1196/annals.1320.012 PMID: 15591010

41. Leite HP, de Lima LFP. Metabolic resuscitation in sepsis: a necessary step beyond the hemodynamic? Journal of Thoracic Disease. 2016; 8(7):E552–E7. https://doi.org/10.21037/jtd.2016.05.37 PMC4958886. PMID: 27501325

42. Druml W, Heinzel G, Kleinberger G. Amino acid kinetics in patients with sepsis. Am J Clin Nutr. 2001; 73(5):908–13. Epub 2001/05/03. https://doi.org/10.1093/ajcn/73.5.908 PMID: 11333844.

43. Samant S, Lee H, Ghassemi M, Chen J, Cook JL, Mankin AS, et al. Nucleotide Biosynthesis Is Critical for Growth of Bacteria in Human Blood. PLoS Pathogens. 2008; 4(2):e37. https://doi.org/10.1371/journal.ppat.0040037 PMC2242838. PMID: 18282099

44. Wilson JX. Mechanism of action of vitamin C in sepsis: Ascorbate modulates redox signaling in endothelium. BioFactors (Oxford, England). 2009; 35(1):5–13. https://doi.org/10.1002/biof.7 PMC2767105. PMID: 19319840

45. Nolan A, Kobayashi H, Naveed B, Kelly A, Hoshino Y, Hoshino S, et al. Differential role for CD80 and CD86 in the regulation of the innate immune response in murine polymicrobial sepsis. PLoS One. 2009; 4(8):e6600. Epub 2009/08/13. https://doi.org/10.1371/journal.pone.0006600 PMID: 19672303; PubMed Central PMCID: PMCPMC2719911.

46. Cazalis M-A, Lepape A, Venet F, Frager F, Mougin B, Vallin H, et al. Early and dynamic changes in gene expression in septic shock patients: a genome-wide approach. Intensive Care Medicine Experimental. 2014; 2:20. https://doi.org/10.1186/s40635-014-0020-3 PMC4512996. PMID: 26215705

47. Hotson AN, Hardy JW, Hale MB, Contag CH, Nolan GP. The T cell STAT signaling network is reprogrammed within hours of bacteremia via secondary signals. Journal of immunology (Baltimore, Md: 1950). 2009; 182(12):7558–68. Epub 2009/06/06. https://doi.org/10.4049/jimmunol.0803666 PMID: 19494279; PubMed Central PMCID: PMCPMC4136495.