Predicting Wound Healing: A Machine-learning, Partial Least Squares Discriminant Analysis Model Utilizing Microbiome, Metabolome, and Clinical Marker Data Sets

Catherine B. Anders Ph.D.[1,2 *], Hannah Smith[2], Jeremy Boyd, Ph.D.[2], Michael C. Davis, M.S.[1], Tyler M.W. Lawton[2], Margarette M. Doucette, DO[1,3], Mary Cloud B. Ammons Ph.D.[1,2 ⊤]

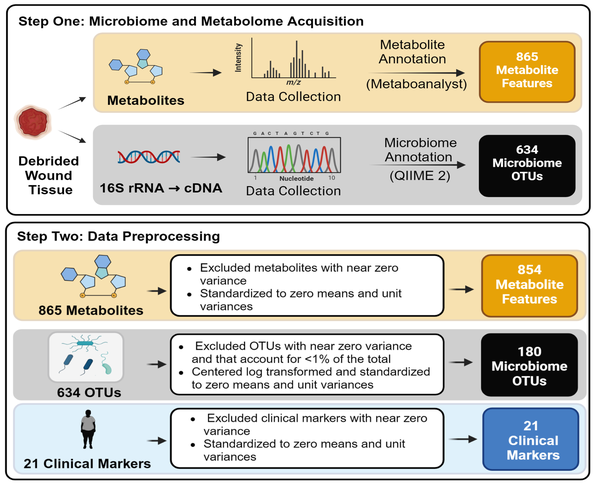
1) Research Service, Boise VA Medical Center (BVAMC), 2) Idaho Veterans Research & Education Foundation (IVREF), 3) Department of Physical Medicine & Rehabilitation BVAMC

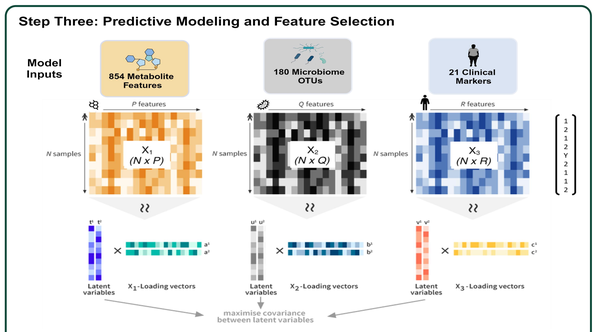
*Presenting Author, ⊤Corresponding Author

Significance

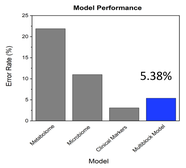
Type 2 Diabetes affects more than 37 million people in the United States and is the number one cause of non-traumatic lower-limb amputation in adults due to diabetic foot ulcers (DFU). The chronic wound microenvironment consists of a complex milieu of host cells, microbial species, and metabolites. While much is known about the wound microbiome, our knowledge of the metabolic landscape and its influence on microbial diversity and wound healing is limited. Furthermore, the integration of these complex datasets into a predictive model with relevance to clinical outcome is almost non-existent. Here, we present a multi-omics data analysis coupled with machine-learning cross validation of microbiome and metabolome profiles from human chronic wounds. The model was then integrated with patient metadata to determine predictive correlation to clinical outcome. The final model selected a total of 527 features (N = 16 clinical, 91 microbiome, and 420 metabolome), and was able to predict the clinical outcome with an overall error rate of 5.38%. These results indicate that the integration of wound microbiome and metabolomics data with patient clinical metadata can be utilized to predict clinical outcomes regarding wound healing and with low error rates. Furthermore, the biomarkers selected within the model may offer novel insights into wound microenvironment composition and improve treatment efficacy in difficult to heal wounds.

Methods





Model performance and selected number of features:



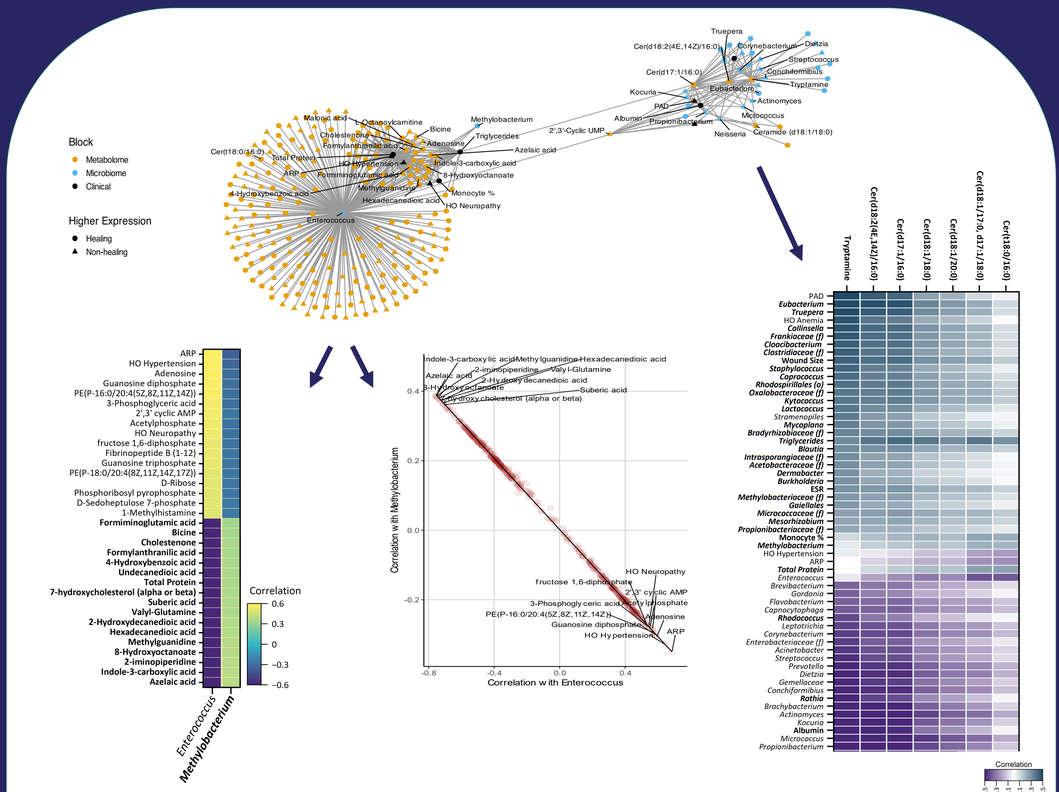
Block	Number of Features
Metabolome	420
Microbiome	91
Clinical Markers	16
Total	527

References: 1) Lê Cao K-A, et al. *PLOS ONE*. 2016;11(8):e0160169.; 2) Lê Cao K-A, Welham M. CRC Press; 2022.; 3) Liquet B, et al. *BMC Bioinformatics*. 2012;13(1):325.; 4) Singh A, et al. *Bioinformatics*. 2019;35(17):3055–3062.; 5) Rohart F, et al. *PLOS Computational Biology*. 2017;13(11):e1005752.

Acknowledgements

This work was supported in part by the U.S. Department of Veterans Affairs, Office of Research and Development Biomedical Laboratory Research Program, the Idaho INBRE Program (NIH NIGMS P20 GM103408), a NIAID-NIH award (RO3AI135998; PI Ammons), the IVREF Center of Biomedical Research Excellence in Emerging/Reemerging Infectious Diseases (NIH NIGMS P20GM109007), the Wound Healing Foundation, and a VA Merit Award (I01BX005344). This content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health, U.S. Department of Veterans Affairs, or the United States Government.

Clinically-Relevant Biological Insight is Discovered Through Multiblock Integrative Predictive Modeling



Clinical Insights

- Much of the clinical microbiology data is limited to culturable bacteria and lacks information on the impact of metabolic interaction between host and pathogen; this multi-omics approach, integrated with clinical outcome, provides a much more comprehensive view of host-pathogen interaction in a clinically-relevant model.
- While clinical metrics alone are predictive of outcome, integration of the multiblock model provides impactful insight into the fundamental mechanisms driving clinical outcome.
- Multiblock predictive modeling further supported the importance of investigating beyond the most **abundant** colonizing microbes and delving into the most **impactful** microbes through contribution to the metabolic landscape.

IDAHO VETERANS RESEARCH & EDUCATION FOUNDATION

VA HEALTH CARE Defining EXCELLENCE in the 21st Century

ID² COBRE Idaho Center of Biomedical Research Excellence in Emerging/Reemerging Infectious Disease

Results

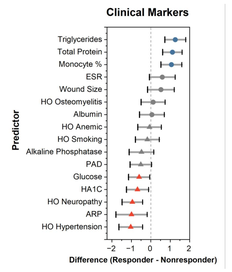
For this project, we collected debrided wound tissue from patients who were enrolled in the Wound Care Clinic at the Boise VA Medical Center. A total of 45 samples were collected. 20 samples were taken from 6 different patients who responded to treatment and are designated as Responders. 25 samples were from 7 patients who failed to respond to treatment and are designated as Non-Responders. The table (A) details patient demographics and clinical markers collected for each sample. The difference plot (B) highlights those clinical markers that were significantly associated with Responders (blue) and Non-Responders (red).

A)

Total Samples (N = 45)	Responders (n = 20) Mean ± SE or n (%)	Non-Responders (n = 25) Mean ± SE or n (%)
Avg Age (years)	70.7 ± 3.1	70.4 ± 1.6
BMI (kg/m^2)	34.6 ± 1.8	31.1 ± 0.9
Glucose (mg/dL)*	152 ± 8.9	194 ± 9.2
Wound Size (cm^2)	4.4 ± 1.6	1.2 ± 0.2
Wound Depth (cm)	0.5 ± 0.1	0.2 ± 0.02
Avg AMR Pathogens*	1.7 ± 0.4	3.8 ± 0.5
HA1C*	6.1 ± 0.3	7.0 ± 0.3
Albumin	3.7 ± 0.1	3.7 ± 0.1
Alkaline Phosphatase	83.3 ± 4.6	91.7 ± 2.5
CRP	2.4 ± 0.8	1.1 ± 0.3
ESR	33.7 ± 4.2	24.1 ± 3.2
Monocyte %*	11.3 ± 0.7	7.7 ± 0.5
Total Protein*	7.4 ± 0.1	6.9 ± 0.1
Triglycerides*	180 ± 15	90 ± 8.0
HO Amputation	0 (0)	29 (72.5)
HO Smoking	11 (55)	13 (52)
HO Hypertension*	12 (60)	24 (96)
HO Anemia	5 (25)	8 (32)
HO Hyperlipidemia	13 (65)	19 (76)
HO Neuropathy*	4 (20)	16 (64)
HO PAD	1 (5)	6 (24)
HO Osteomyelitis	5 (25)	5 (20)

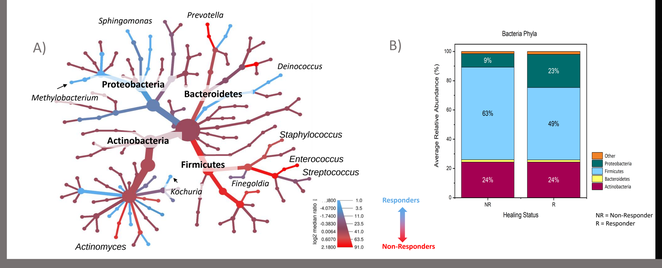
*significant at p < 0.05

B)



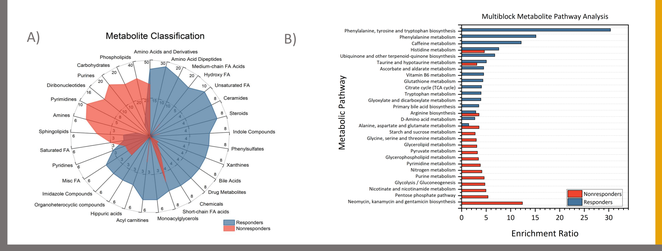
Clinical Markers

The heat graph taxonomic tree (A) shows the hierarchical structure of taxonomic ranks of the 91 unique bacterial genera that were selected into the model. Different colors are used to highlight differentially abundant taxa, based on $\log_2$ median ratio between Non-Responders in red and Responders in blue. Among all phyla (B), Non-Responders had more Firmicutes present in their wounds whereas Responders had more Proteobacteria. When we look at specific genera, we can see that three common and dominant wound-associated pathogens are present in Firmicutes phylum.



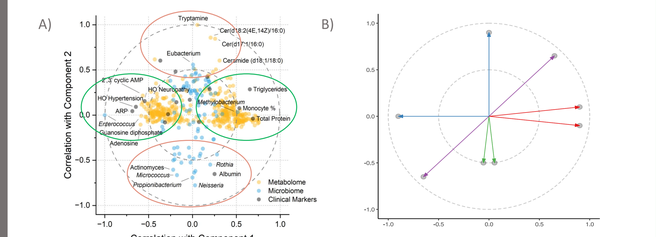
Microbiome

For the metabolome, 420 features were selected into the model. The radar plot depicts the number of metabolite features that are present within the identified metabolite subclasses. On this plot each radiating line of the plot represents a metabolite subclass with the outer potion of the colored section corresponding to the number of metabolites identified within each subclass for both Responders and Non-Responders in red. Most notably, Responder wounds (blue) contained more metabolites associated with amino acids and their derivatives and metabolites associated with anti-inflammatory pathways in contrast to metabolite subclasses represented in Non-Responders (red) which had an over abundance of fatty acids and metabolites related to carbohydrate metabolism and glycolysis. When these metabolites are mapped onto metabolic pathways, we see that the metabolic pathways enriched in Responders are primarily associated with pathways that are critical for wound resolution; however, the pathways enriched in Non-Responders are skewed towards acute and chronic inflammation.



Metabolome

The circle plot includes 527 features selected into the model which displays features selected into the model, as well as relationships among features (A). Each dot represents a model features plotted in 2D space, indicating their correlation with model components one (x-axis), and two (y-axis). The dashed circles are guides representing correlations of ±0.5 (inner circle) and ±1 (outer circle). Features closer to the outer circle are more important for predicting the outcome. The location of individual features relative to other features indicate whether those feature are positively or negative correlated with one another (B). The angle made by connecting two features through the origin gives the sign of their correlation. Acute angles indicate positive correlations (red and green), obtuse angles represent negative correlations (purple), and right angles indicate no correlation (blue). Moreover, the length of the connecting lines gives the magnitude of the correlation. For example, the angle made by the red and green lines is identical, but the red lines are longer, indicating a stronger positive correlation between the red features relative to the green features. Along component one, *Enterococcus* and *Methylobacterium* (green circles) are negatively associated with one another and have mirror-image relationships to the metabolite clusters on the left and right of the figure. Along component two, tryptamine and a ceramide group (orange circles) are positively correlated with each other and negatively correlated to the microbiome features towards the bottom of the graph.



Multiblock Model