MATHEMATICS AND STATISTICS OF, FOR AND AGAINST TUBERCULOSIS

GAURAV D. SANKHE, PH.D., RESEARCH ASSOCIATE, IMMUNOLOGY PROGRAM, SLOAN KETTERING INSTITUTE, 1275 YORK AVE, NEW YORK, USA

SEAR - UNION

Abstract

After a brief dip in Tuberculosis (TB) incidence and mortality driven by Covid-19 pandemic, Mycobacterium tuberculosis (Mtb) emerged yet again as one of the leading cause of deaths in humans across globe. The rise of Multi Drug Resistant TB (MDR-TB) exacerbates the success of “end TB strategy”. Thus, the statistics of TB clearly highlight the requirement for stepping up efforts for comprehensive health initiatives and concerted action plan to combat TB. Two-pronged approach could be by, on one hand, hitting the pathogen Mtb at its adaptability by gleaning into its mechanisms of adaptation and on the other, by reinforcing the host immune onslaught and pressures against Mtb. With two research stories, one involving mathematics and statistics for TB and other against TB, this article accentuates the merit of integrating mathematics and statistics with the biology of host as well as pathogen for performing data-backed theoretical investigations to not only further mechanistic insights but also suggest potential novel targets against TB.

Tuberculosis recently regained its position of the leading cause of deaths in humans from single infectious agent. Till date it is the most prevalent disease across the globe. A quick glance over the recent UN world tuberculosis report [1] brings in some hope but with much more disappointment. After the Covid-19 driven dip or pause in Tuberculosis incidence and mortality across all demographics, the sharp rise in sheer numbers has stretched the “End tuberculosis” timelines beyond the predicted temporal plots. In 2022, there were approximately 10.6 million new TB cases globally, with the disease claiming around 1.6 million lives. India, which is one of the high TB burden countries, reported about 2.9 million new TB cases and an estimated 0.5 million TB-related deaths in the same year. These statistics highlight the urgent need for comprehensive public health initiatives and concerted efforts to combat tuberculosis, particularly in regions with a high burden like India [1]. The rise of multidrug-resistant TB (MDR-TB) reported as ~5% of total global incidence ushers scare and despair. These statistics of TB burden strongly suggest that we are falling short of our efforts, and we have to step up our “endgame” against the TB causing pathogen Mycobacterium tuberculosis (Mtb). Broadly, the strategies of intervention stretch from finding druggable targets against vital mechanisms of adaptation and survival of the pathogen on one side and exploring avenues to reinforce host immune pressure and pathogen specific onslaught on the other. Mathematics and Statistics comes in handy for pursuing both strategies. The following research stories would, presumably, highlight the significance of knowledge of mathematics and statistics that could be fused with underlying biology to sharpen our tools in the battle against Tuberculosis.

The current regimen for DSTB (Drug Sensitive TB) includes Rifampicin (RIF), Isoniazid (INH), Pyrazinamide (PZA), and Ethambutol (EMB). Treatment typically spans 6 to 9 months, starting with an intensive phase of all four drugs for 2 months, followed by a continuation phase with RIF and INH for 4 to 6 months. MDR-TB is defined as resistance to at least two of the four, Rifampicin (RIF) and Isoniazid (INH), the two most potent anti-tubercular drugs. These antibiotics exert collateral damage on the other symbiotic bacteria in the gut, collectively called the gut microbiome. Recently a clinical study on gut microbiome dysbiosis and dynamics performed over time for DS-TB and MDR-TB patients [2] revealed that the TB-driven inflammation lead strong interrelation between gut microbiome and TB clearance. Microbiome disruption leading to pathobiont (troublemaker bacteria) dominance and evolution of resistance in commensals (gut friendly bacteria) further establishing their dominance in the microbiome provides a mechanistic understanding of the temporal fate of TB antibiotic driven gut microbiome dysbiosis. Fecal-microbiota transplantation of antimicrobial resistant commensal microbiome presents a unique solution for avoiding or mitigating TB drugs-driven inflammation thus reducing inertia for TB drug administration by patients over the prolonged due course. Now, in this clinical trial to determine the effect of antibiotics used in TB treatment on the composition of fecal microbiome (proxy for gut microbiome), and peripheral blood transcriptomics (signatures for TB-driven inflammation) used to model TB disease progression and resolution, involved hardcore mathematics and statistics. For instance fecal microbiome characterization informed enrichment in gut commensals (measured as NES, normalized enrichment score) as a function of microbiome dynamics, and mycobacterial load reduction (TTP) was defined for a particular hallmark pathway or TB signature obtained from blood peripheral transcriptomics yielding expression levels of common inflammatory signatures like interferon (IFN)-α, IFN-γ, interleukin-6 (IL-6) and Janus Kinase(JAK)-signal transducer and activator of transcription 3 (STAT3). Briefly, NES for a particular TTP is considered a general nonlinear function (the random forest) is applied to TTP and species relative abundances X in a particular sample say i as fixed effects and 1/ID (Internal expression differences) indicates random effects to account multiple samples from the same patient. Mixed-effects random forest regression modeling where NES for particular i was fit as a nonlinear function of (TTP and X) i + 1/ID, was applied to determine model-inferred associations [2]. Linear mixed-effects models were also run to identify significant associations between sex, age, and treatment kinetics on microbiome diversity to microbiome composition. These models were used to identify microbial species, pathways, and host genes associated with sex, age and treatment duration. This study appreciates the power of mathematics and statistics for correlating and delineating the underlying effects of a complex phenomenon, resulting in crucial insights presented by tangible results [2].

he study explored changes in the gastrointestinal microbiome associated with MDR TB treatment by using metagenomic sequencing of stool samples. Principal coordinate analysis on centered log-ratio transformed data was used to address compositionality in species and metabolic pathway abundance. Significant differences in microbial composition and gene content were identified between baseline and treatment completion samples compared to those collected at 2 weeks, 2 months, and 6 months. This was evidenced by PERMANOVA ($P < 0.05$) [3]. Notably, 23.6% of [Fig.1B] the variation in species abundance and 40.1% [Fig.1C] in pathway abundance on the first axis were linked to treatment-induced microbiome changes in the first 6 months. The second axis highlighted variability in early treatment response among different individuals.

The study explored the impact of MDR TB treatment on microbial species and metabolic pathways, comparing findings to pretreatment and drug-sensitive TB treatment results. MDR TB therapy, with the exception of increasing Flavonifractor plautii, significantly depleted the GI microbiome within the first six months (FDR < 0.05). Key species such as Bacteroides uniformis, Blautia spp., Clostridium bolteae, and others, known for producing bioactive metabolites and transforming bile acids, were notably reduced [2, 4].

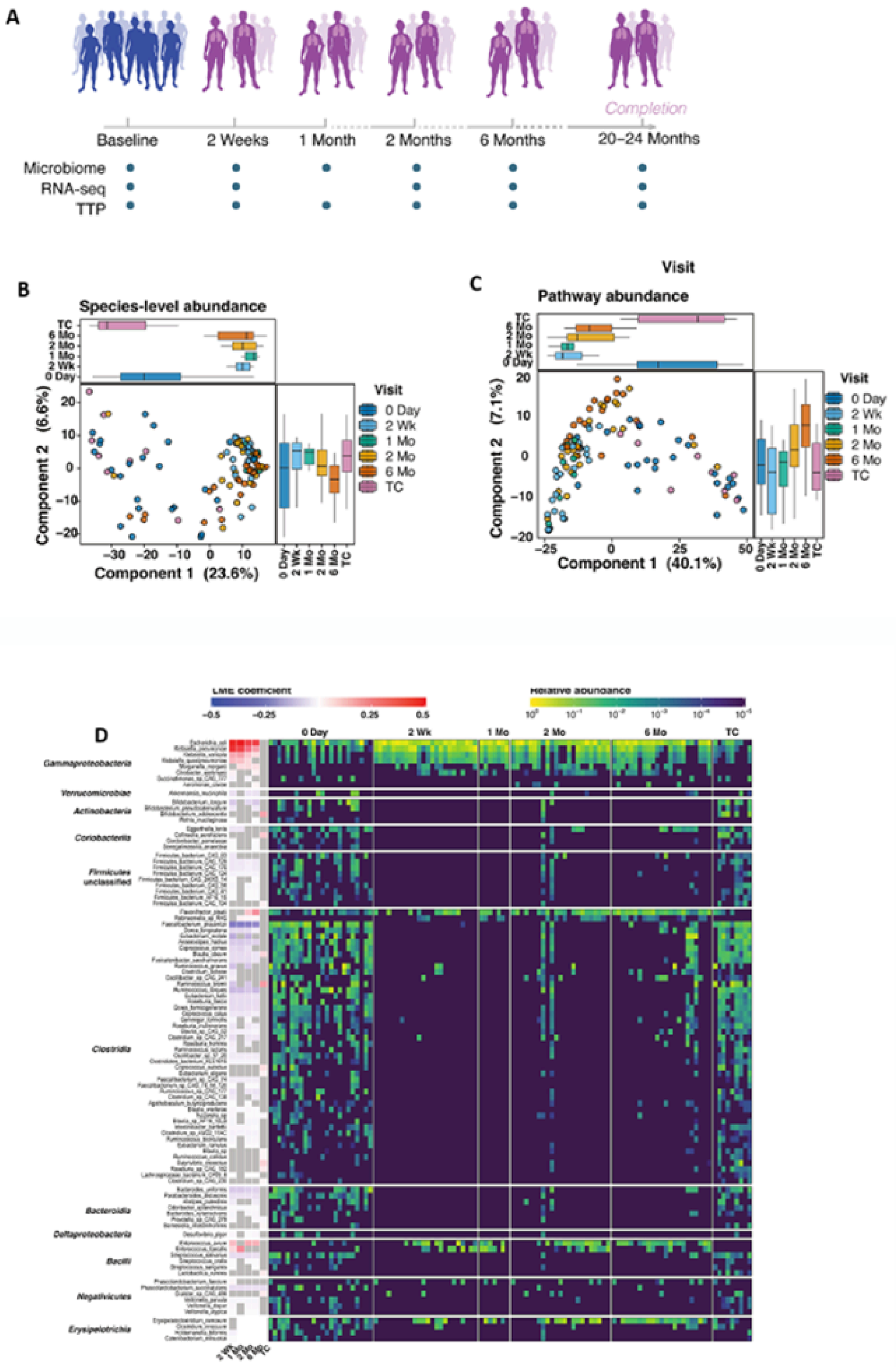
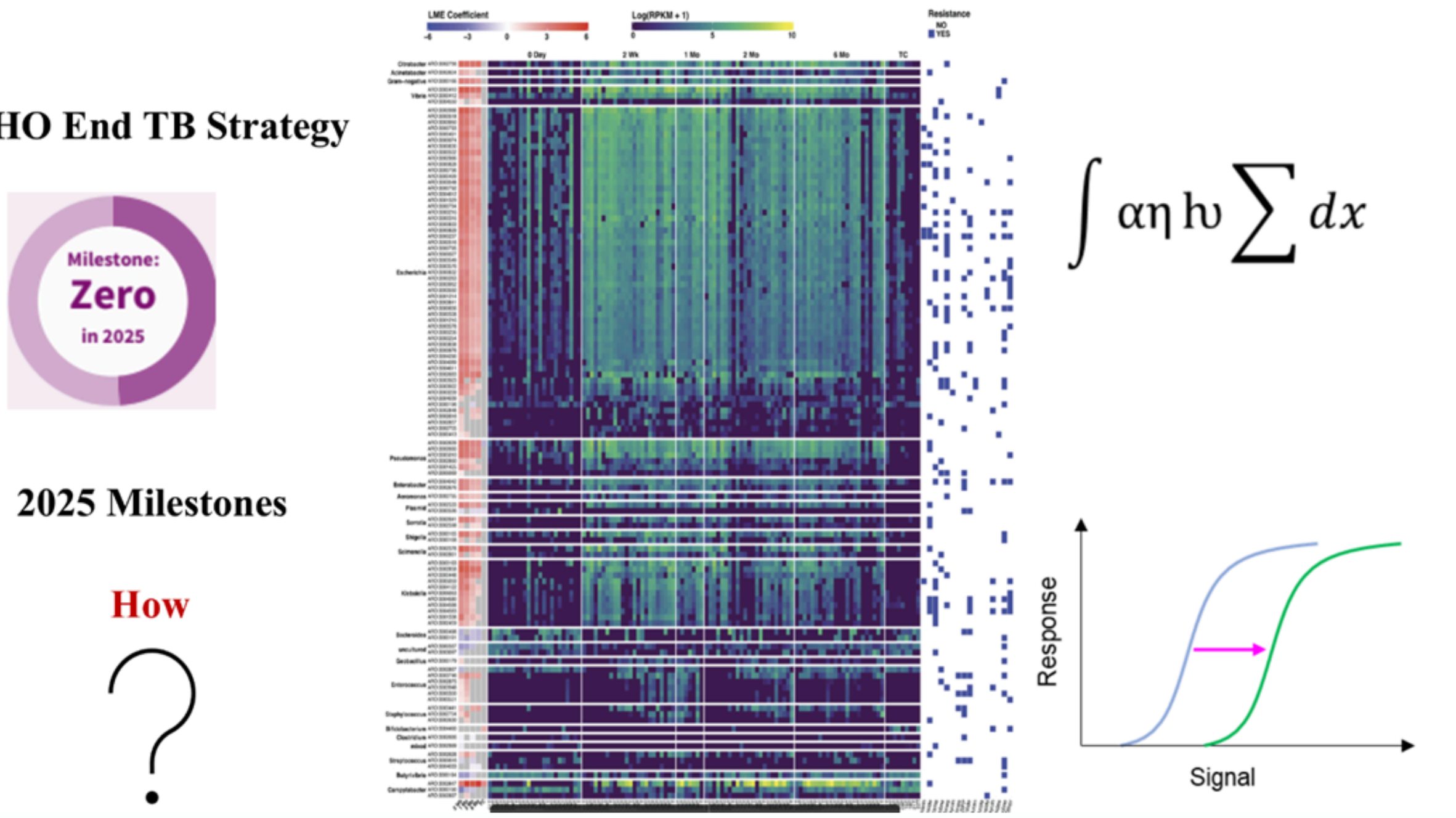


Fig. 1. MDR TB treatment induces Mtb lung sterilization and causes a temporary perturbation in the microbiome, which recovers by treatment cessation. Schematic of the MDR TB treatment observational cohort. (B and C) Principal coordinate analysis on center-log-ratio transformed data for species abundances (B) and functional pathway abundance (C)from metagenomic sequencing. (D)Taxonomic abundance of the microbiome during MDR treatment. Bhattarai et al., Sci. Transl. Med. 16, eadi9711 (2024)

Emerging MDR TB strains have invigorated the need of finding novel mechanisms of adaptation of the pathogen in order to suggest strategies of intervention against it. Two component signaling system (TCS), comprising of a sensing component known as histidine kinase (HK) and a downstream responsive component named as response regulator (RR), is a predominant toolbox used by bacteria and lower eukaryotes for sensing and adaptation [5]. Phosphorylation at histidine residue of HK (hence the name) upon incoming signal and physical transfer of the phosphoryl group to aspartate residue of RR thereafter, forms the atypical “talk” to inform and respond towards the incoming signal. HK and RR are present on the same operon in the bacterial genome like a couple made in heaven and hence termed as a cognate pair. Mtb TCSs are recently shown to have non-cognate interactions as well informally known as “crosstalk” [6]. TCSs are a playground for mathematicians and theoreticians for identifying various signaling landscapes and unveiling their unique and vivid output features. In another recent study, sequestration of HK by other non-cognate RRs establishes a signal detection threshold below which all incoming signals for the HK get diffused and ignored [5]. The pathogen uses this simple design principle to define the quantum of signal versus noise thereby distinguishing what to ignore and what to respond. This sequestration phenomenon where non cognate RR tightly binds to a signal-active HK could be imagined similar to our close friend acting like a sequester by hugging us tightly in the middle of an altercation posed by a bully so that we don’t respond to it and lose our mind over it. This power of ignorance is blissful for Mtb allowing it to avoid mounting disproportionately large and biologically costly responses against frivolous signals. Since incoming signals or ligands for many HKs in Mtb are unknown, without employing mathematical framework and statistical tools, it would have been an arduous task to achieve both (a) visualize and formalize this underlying phenomenon and (b) proving it quantitatively and robustly. Considering a theoretical signal of various strengths (Concentration and time) as input and expression levels of the downstream genes for the signaling cascade as output, the input-output profiles with and without sequestration were mathematically modeled using all phosphorylation and other chemical reactions leading to the expression events. Most of the kinetic parameters of the reactions were obtained by biochemical and biophysical techniques in vitro (in test tubes) and others were obtained through literature. The enthralling moment in the study is witnessing the mathematical model employing invitro kinetic parameters predicting the extent decrease in expression of downstream genes due to sequestration given the relative increase in the sequester amount in vivo (in mycobacterial cell) overlapped robustly with the expression levels of the downstream genes in vivo obtained by qRT PCR. Phosphohistidine and phosphoaspartate in signal activated HKs and RRs respectively, being unique and rare in humans, make TCSs potent drug targets. Targeting such a sequestration interaction can render the bug to get bugged even with small stress signals posed by host immunity and or drug regimens. Exploiting this finding to pin a sequestration interaction, in principle, can exhaust the pathogen enough to succumb to host-immune pressures. Pinning down various protein complexes using molecular scaffolds has been a tried and tested hack by pharma companies which could be easily extended in this case. Here, we again observe that mathematics and statistics rendered envisioning a phenomenon of HK sequestration by non-cognate RRs, thus providing a potential target against the tuberculosis pathogen.

Mathematics and statistics have been savior to simulate clinical trials for various potent vaccine candidates during the COVID-19 pandemic in order to proceed with the expensive and exhaustive multiple phase clinical trial of the best candidate thereby accelerating vaccine development when it was most required (Trial simulator software by Certara Inc, USA), Further to compute optimal dosage amount and regimen for the vaccine mathematical models defining immunological mechanisms are extensively used [7]. Mathematical modeling has also helped elucidate and provide the possible mechanistic insights of natural and post-treatment control and quantified effects of antiretroviral therapy used against HIV [8]. Given the extensive utilization of mathematics and statistics in not only in TB research but also to other infectious diseases, I honestly feel that having some level of training in mathematics from high school till graduation is imperative and should be ensured through government policies. The western world has realized this in its full glory where a former British PM recently enforced mathematics as a compulsory subject in its high school curriculum rather than being optional. Indeed, generative AI tools like Bard, ChatGPT, Perplexity AI will exist to assist us for the future, yet, they won’t help in developing the underlying logic and rationale, essential for building algorithms. Better learn a subject for a few productive years than to be a lifetime slave or subscriber of AI, right?



Abstract art highlighting ambitious WHO End TB Strategy target (left), a possible method to reach that goal could be utilization of mathematics and statistics in TB research as showcased in a representative heat map relative abundance of various bacterial families in gut microbiome [2] (center) and elucidative equation with graphical representation of Stimulus-Response profile with sequestration.

References

1. Global tuberculosis report 2023: Geneva: World Health Organization: 2023. https://www.who.int/teams/global-tuberculosis-programme/tb-reports/global-tuberculosis-report-2023.
2. Bhattarai, S. K., Du, M., Zeamer, A. L., M Morzfeld, B., Kellogg, T. D., Firat, K., Benjamin, A., Bean, J. M., Zimmerman, M., Mardi, G., Vilbrun, S. C., Walsh, K. F., Fitzgerald, D. W., Glickman, M. S., & Bucci, V. (2024). Commensal antimicrobial resistance mediates microbiome resilience to antibiotic disruption. Science translational medicine, 16(730), eadi9711. https://doi.org/10.1126/scitranslmed.adi9711.
3. Lloyd-Price, J., Arze, C., Ananthakrishnan, A. N., Schirmer, M., Avila-Pacheco, J., Poon, T. W., Andrews, E., Ajami, N. J., Bonham, K. S., Brislawn, C. J., Casero, D., Courtney, H., Gonzalez, A., Graeber, T. G., Hall, A. B., Lake, K., Landers, C. J., Mallick, H., Plichta, D. R., Prasad, M., … Huttenhower, C. (2019). Multi-omics of the gut microbial ecosystem in inflammatory bowel diseases. Nature, 569(7758), 655–662. https://doi.org/10.1038/s41586-019-1237-9.
4. Sankhe, G. D., Dixit, N. M., & Saini, D. K. (2018). Activation of Bacterial Histidine Kinases: Insights into the Kinetics of the cis Autophosphorylation Mechanism. mSphere, 3(3), e00111-18. https://doi.org/10.1128/mSphere.00111-18.
5. Sankhe, G. D., Raja, R., Singh, D. P., Bheemireddy, S., Rana, S., Athira, P. J., Dixit, N. M., & Saini, D. K. (2023). Sequestration of histidine kinases by non-cognate response regulators establishes a threshold level of stimulation for bacterial two-component signaling. Nature communications, 14(1), 4483. https://doi.org/10.1038/s41467-023-40095-2.
6. Agrawal, R., Sahoo, B., & Saini, D.K. (2016). Cross-talk and specificity in two-component signal transduction pathways. Future microbiology, 11, 685-97, https://doi.org/10.2217/fmb-2016-0001.
7. Giorgi, M., Desikan, R., van der Graaf, P. H., & Kierzek, A. M. (2021). Application of quantitative systems pharmacology to guide the optimal dosing of COVID-19 vaccines. CPT: pharmacometrics & systems pharmacology, 10(10), 1130–1133. https://doi.org/10.1002/psp4.12700.
8. Vemparala, B., Chowdhury, S., Guedj, J., & Dixit, N. M. (2024). Modelling HIV-1 control and remission. NPJ systems biology and applications, 10(1), 84. https://doi.org/10.1038/s41540-024-00407-8.