Sex-Based Differences in Tuberculosis Immune Response: The Role of Leptin

**Authors**

* Guozheng Yang (ORCID: 0000-0001-7361-1847)

**Author affiliations**

1. College of Public Health, University of Georgia, Athens, GA, USA.

Corresponding author: guozheng0224@uga.edu

# Abstract

## Background:

In tuberculosis infection, higher body fat storage in females may better support immune activation and response compared to males. Leptin may mediate this process, bridging metabolism and the immune system.

## Methods:

We conducted a cross-sectional study in Kampala, Uganda, recruiting 60 participants with the first episode of tuberculosis. Demographic characteristics, body composition indices, and immune response measures were compared between 30 male and 30 female participants. A structural equation-based mediation analysis examined the association between CD4+ cell levels and body fat, as well as leptin’s mediation effect. The analysis was performed for all participants and separately by sex. To incorporate material from this course, we further evaluated the performance of random forest and regression tree models in predicting leptin levels using body composition indices. One-fifth of the data set was reserved as a test set for model evaluation, while hyperparameter tuning was conducted on the training set using five-fold cross-validation.

## Results:

Compared to male participants, females had a higher fat mass index (median: 2.46 vs. 4.41), higher leptin levels (median: 1.16 ng/ml vs. 1.56 ng/ml), and higher CD4+ cell levels. Leptin levels were positively associated with body fat storage, with a stronger effect observed in females. Among all participants, 82% of the effect of body fat storage on immune activation was mediated by leptin, with a more pronounced mediation effect in females than in males. The random forest model is better than the regression tree in predicting leptin levels using body composition indices, with a lower root mean squared error (0.40 vs 0.41) and a higher R-square (0.53 vs. 0.23) on the test set.

## Conclusions:

Enhancing body fat storage may support immune activation in tuberculosis infection. Leptin could serve as a therapeutic target for sex-specific immune modulation in tuberculosis.

# Keywords:

**Tuberculosis, leptin, body fat, immune response**

# 1. Introduction

As a global threat, tuberculosis (TB) caused 10.8 million new cases in 2023 with Africa contributing 24% of these cases (1). Countries with high annual incidence rates are mostly in Africa, where TB incidence is significantly higher than in other regions. TB prevention and control strategies have shown progress in recent years. However, further efforts are needed to eliminate TB infection and transmission.

Males are more likely to develop TB disease than females, a difference that cannot be fully explained by socioeconomic status or access to healthcare (2,3). Research has shown that this sexual bias is linked to differences in immune systems, influenced by chromosome-encoded genes and hormones (4). Previous studies indicate that body fat mass could also affect immune responses (5,6). Since women generally have higher body fat storage, their immune systems may receive sufficient energy to support a stronger response to infections compared to men. In diseases like TB that often cause body wasting (7), differences in energy reserves may contribute to variations in immune response between men and women.

Leptin, a key link between metabolism and the immune system, plays an important role in activating and regulating immune responses (8,9). It functions as an energy indicator, signaling whether the body has enough energy to activate and sustain immune responses (10,11). For individuals with similar body fat mass, females generally have higher serum leptin levels than males (12). However, leptin levels may be lower in tuberculosis patients (13), suggesting a suppressed role in regulating metabolism and immune system stability. Since women typically have higher body fat stores and leptin levels, the impact on women may be less pronounced compared to men. However, there is a lack of evidence to demonstrate this sex difference.

This paper presents findings on sexual differences in metabolism and immune responses in individuals initially infected with TB. The primary hypothesis is that women have a stronger immune response to TB due to higher body fat storage. Additionally, leptin is hypothesized to mediate immune response based on body energy storage.

# 2. Methods

## 2.1 Study design and participants

From March to April 2017, a cross-sectional study was conducted in Kampala, Uganda. Sixty participants were recruited from Mulago Hospital, including 30 males and 30 females aged 15 years or older. All participants had a first episode of TB confirmed by culture or molecular methods. To prevent interference with metabolism and immune responses, individuals with co-morbidities such as asthma, kidney disease, liver disease, cancer, HIV, or diabetes were not enrolled. Those receiving hormonal or immune therapies were also excluded.

## 2.2 Body index measures

At enrollment, basic body indices, including age, sex, height (m), and weight (kg), were measured. Participants who provided informed consent underwent total body water (TBW, L) estimation using deuterium dilution. Specifically, participants received a precisely weighed oral dose of deuterium oxide (~25-30g , 99.8% purity) after an overnight fast. Participants were asked to refrain from eating and to consume only the minimum amount of water needed to quench thirst during the 4-hour equilibration period. Saliva samples collected before and after a 4-hour equilibration period were analyzed using Fourier transform infrared spectrophotometry to determine deuterium concentration, and TBW was estimated by extrapolating deuterium dilution space, adjusted by a factor of 1.041 (14). Based on TBW estimates, lean body mass (LBM, kg) was further estimated assuming a hydration fraction of 0.732 (14). Body fat mass (FM, kg) was calculated by substituting LBM from body weight. The percentage of TBW, LBM, and FM were calculated based on body weight. Fat mass index (FMI) and Fat-free mass index (FFMI) were calculated using FMI = FM (kg) / height () and FFMI = LBM (kg) / height (). With consent, venous blood samples were collected to measure leptin levels (ng/ml). All measurements were conducted at the Institute of Infectious Diseases at Makerere University.

## 2.3 Immune response measures

To assess immune activation, flow cytometry was used to measure CD4+ and CD8+ T cell counts in venous blood samples (7 ml of plasma). The percentages of CD4+ and CD8+ cells were calculated by dividing their counts by the total lymphocyte count. Levels of C-reactive protein (CRP, mg/dl), interferon-gamma (INF-gamma, pg/ml), tumor necrosis factor-alpha (TNF-alpha, pg/ml), and interleukin-10 (IL-10, pg/ml) were also measured as indicators of immune response.

## 2.4 Data cleaning

Due to measurement instability, a small number of participants had total body water (TBW) values recorded as less than 0 or greater than 0.70, which are biologically implausible. As TBW serves as a foundational parameter for calculating other body composition indices, participants with TBW values outside the plausible range (<0 or >0.73) were excluded from downstream analyses to ensure data quality and the validity of derived measures. In addition, percentage-based indices such as body water and body fat were originally recorded in decimal format and were converted to percentage format (e.g., 0.45 to 45%) to enhance interpretability. CD4+ and CD8+ cell counts were rounded to the nearest integer to reflect the typical reporting standard in clinical and immunological research.

## 2.5 Statistical analysis

In the descriptive analysis, we compared the distribution of demographic characteristics (e.g., age, height, weight) and body composition indices (e.g., body fat, lean body mass, leptin level) between male and female participants using the median and interquartile range (IQR). Differences in immune responses were visualized with boxplots showing the distributions of available measures, including CD4+ cell counts, CRP levels, and INF-gamma levels. The association between log-transformed leptin levels and body fat, lean body mass, percentage of CD4+ cells, and percentage of CD8+ cells was assessed using scatterplots with Pearson correlation coefficients.

We examined the predictability of fat mass index to leptin level using a simple linear regression. A Poisson regression model was applied to CD4+ cell counts using a log link, with fat mass index and leptin level as predictors. All leptin levels in regression models used the logarithmic transformation. We estimated standardized regression coefficients and calculated the total, direct, and indirect effects of fat mass index on CD4+ cell counts (15). The mediation effect of leptin was quantified as the indirect effect divided by the total effect, expressed as the percentage of the total effect mediated by leptin (16). Confidence intervals (CI) at the 95% confidence level for all estimated effects were obtained using the bootstrap method with 500 replicates. The mediation analysis was first conducted for all participants. We then performed a subgroup analysis by applying the mediation analysis separately to male and female participants.

We further compared the performance of random forest and regression tree models in predicting log-transformed leptin levels using body composition indices. The predictors included gender, age, BMI, percentage of lean body mass, and percentage of fat mass. One-fifth of the dataset was set aside as the test set, with the remaining data used for training. Five-fold cross-validation was applied to the training set for hyperparameter tuning. For the random forest model, a grid search was conducted to optimize the number of randomly sampled predictors and the minimum number of observations required to split a node, targeting the lowest root mean squared error (RMSE). For the regression tree model, cost complexity and tree depth were tuned to minimize RMSE as well. Model performance was evaluated on the test set using RMSE and R-squared.

Last, we simulated a data set with log-transformed leptin level, sex, age, and BMI. This simulated data set was then used to examine the performance of random forest model in predicting leptin level using the other variables as predictors. The sex variable was generated using a Bernoulli distribution with a probability of 0.6 of being females. Age was assumed to follow a normal distribution with a mean of 35 and standard deviation of 3. BMI was also assumed a normal distribution with a mean of 25 and standard deviation of 1.5. We generated log-transformed leptin levels assuming the relationship between it and the other predictors as: . Besides the association structure as described above, we also added an error term following a standard normal distribution. The model evaluation process was the same as conducted on the real-world data set.

# 3. Results

## 3.1 Difference in body composition index by sex

The 30 male and 30 female participants had a similar age distribution ([Table 1](#tbl-body)). Males were taller (median = 168.00 cm, IQR: 158.00 cm-175.00 cm) and heavier (median = 55.20 kg, IQR: 50.00 kg-60.00 kg) than females, but BMI was similar between sexes (median: 19.48 vs. 19.20). Males had a higher body water proportion (median = 63.67%, IQR: 59.92%-68.09%) and lean body mass composition (median = 86.98%, IQR: 81.86%-93.01%), while females had a higher body fat composition (median = 21.37%, IQR: 14.74%-27.09%). After adjusting for height, males had a lower FMI (median: 2.46 vs. 4.41) but a higher FFMI (median: 16.80 vs. 15.19) than females. Leptin levels were higher in females than in males (median: 1.16 ng/ml vs. 1.56 ng/ml).

|  |
| --- |
| Table 1: Body characteristics and composition by sex. |

## 3.2 Difference in immune response by sex

CRP, INF-gamma, TNF-alpha, and IL-10 levels showed no clear differences between males and females ([Figure 1](#fig-boxplot)). However, females had more compact distributions for CRP, TNF-alpha, and IL-10 levels. Both sexes had INF-gamma levels well above zero. Lymphocyte subsets varied between sexes. CD4+ levels were lower in males but showed less variation. The difference was not evident regarding CD8+ level. Among males, the median CD4+ count was around 300/uL, with a median percentage of about 20%, while females had a median count above 400/uL. Both sexes had CD4+ and CD8+ counts as low as nearly zero. The maximum CD4+ count reached 800/uL, whereas the maximum CD8+ count was over 400/uL in males and nearly 700/uL in females.

|  |
| --- |
| Figure 1: Comparison of immune response by sex. The boxplots denote immune response measures among male (blue) and female (orange) participants. |

## 3.3 Association between body composition and immune response

Leptin levels were strongly associated with body fat among females, with a positive correlation (Pearson correlation coefficient = 0.6, [Figure 2](#fig-sct)). In contrast, leptin levels showed a negative correlation with lean body mass in females (Pearson correlation coefficient = -0.4). Males had consistently lower leptin levels than females, with no evident correlation with body fat or lean body mass. No clear association was observed between leptin levels and lymphocyte subsets. The correlation between leptin levels and the percentage of CD4+ cells was weak (Pearson correlation coefficient = 0.09). Similarly, the correlation with the percentage of CD8+ cells was slightly negative but not significant (Pearson correlation coefficient = -0.13).

|  |
| --- |
| Figure 2: Association between log-transformed leptin level and body fat, lean body mass, percentage of CD4+ cells, and percentage of CD8+ cells. Male and female participants are denoted by blue and pink points, respectively. The size of point denotes body weight of each participant. R marks the Pearson correlation coefficients. |

## 3.4 Mediation effect of leptin in immune response activation

When fitting linear regression for all participants, leptin level was positively associated with FMI (=0.100, 95% CI: 0.064-0.135, [Table 2](#tbl-coef)). A similar positive association was observed among female participants (=0.113, 95% CI: 0.062-0.165), but not among males (=0.000, 95% CI: -0.003-0.002). When predicting CD4+ cell counts using FMI and leptin level for all participants, both predictors showed a positive association (=0.008, 95% CI: 0.000-0.017 for FMI; =0.364, 95% CI: 0.314-0.413 for leptin level). Among males, both FMI (=0.025, 95% CI: 0.011-0.040) and leptin level (=6.995, 95% CI: 4.489-9.500) were positively associated with CD4+ cell counts. However, among females, CD4+ cell counts were positively associated only with leptin level (=0.322, 95% CI: 0.262-0.381).

|  |
| --- |
| Table 2: Coefficient estimation (95% CI) from linear regression and Poisson regression. Fat mass index is only included in the linear regression model, while both predictors are included in the Poisson regression model to assess medication effects. |

Among all participants, FMI had a notable effect on promoting CD4+ cell counts (total effect=0.25, 95% CI: -0.15-0.57, [Table 3](#tbl-ma)), with a large proportion (indirect effect=0.20, 95% CI: -0.14-0.55) mediated by leptin level. The mediation effect of leptin accounted for 82% of the total effect of FMI. A similar pattern was observed among female participants, where the effect of FMI (total effect=0.24, 95% CI: -0.35-0.67) was almost entirely mediated by leptin (indirect effect=0.24, 95% CI: -0.30-0.80). Among male participants, CD4+ cell counts were primarily influenced by the direct effect of FMI (total effect=0.09, 95% CI: -0.55-0.76; direct effect=0.10, 95% CI: -0.57-0.89).

|  |
| --- |
| Table 3: Mediation analysis with estimated effects and 95% CI. For negative direct effect or indirect effect, the percentage of mediated effect cannot be calculated due to model setting. |

## 3.5 Prediction leptin level using body compositions

The tuning results indicate that the random forest model achieved optimal performance with five randomly sampled predictors and a minimum split size of approximately 15 ([Figure 3](#fig-rf)). This led to an RMSE of 0.40 and an R-square of 0.53 on the test set. In contrast, the regression tree model demonstrated minimal variation in RMSE across different values of tree depth and cost complexity ([Figure 4](#fig-rt)), with the optimal parameter set resulting in an RMSE of 0.41 and an R-square of 0.23 in the test set. This lack of sensitivity to parameter changes suggests that the regression tree may have reached its performance ceiling due to the limited sample size and the model’s inherent simplicity. Unlike random forest, which benefits from averaging multiple trees and capturing more complex patterns, the single-tree structure of the regression tree model may not sufficiently capture the underlying variability in the data.

|  |
| --- |
| Figure 3: Parameter tuning results of random forest. Darker square means lower root mean squared error. |

|  |
| --- |
| Figure 4: Parameter tuning results of regression tree. Darker square means lower root mean squared error. |

The plots comparing predicted values and observed values further illustrate the performance differences between the two models ([Figure 5](#fig-rfrt)). Predicted values from the random forest model are more clustered, generally following the diagonal line. This indicates a better agreement between predicted and observed leptin levels. In contrast, the regression tree model shows clear signs of underfitting, with predicted values appearing in discrete horizontal bands and failing to capture the continuous nature of leptin levels. This is one of the characteristics of single-tree models, which produces piece-wise constant predictions and lacks the flexibility to model associations. Overall, the random forest model provides more accurate and nuanced predictions, benefiting from its more flexible structure and ability to reduce variance.

|  |
| --- |
| Figure 5: Comparison of random forest and regression tree: predicted values vs. observed values. The dashed black line on the diagonal denotes the reference where predicted values are equal to observed values. |

# 4. Simulation study using random forest

The performance of the random forest model on the simulated data set appears to be highly accurate and optimized ([Figure 6](#fig-rfsim), [Figure 7](#fig-rfpredsim)). The parameter tuning results demonstrate that the lowest RMSE was achieved when the number of randomly sampled predictors was set to 3 and the minimum number of split points was 37. In contrast, the model exhibited higher RMSEs as the number of randomly sampled predictors decreased, suggesting that incorporating a larger subset of predictors improves predictive performance. Additionally, the scatterplot comparing predicted and observed values shows that most points lie very close to the diagonal line, indicating strong agreement between predictions and actual outcomes. These findings collectively suggest that the random forest model effectively captured the complex, non-linear relationships embedded in the simulated data, resulting in robust and reliable predictions.

|  |
| --- |
| Figure 6: Parameter tuning results of random forest model using simulated data. Darker square means lower root mean squared error. |

|  |
| --- |
| Figure 7: Fitting results of random forest on simulated data: predicted values vs. observed values. The dashed black line on the diagonal denotes the reference where predicted values are equal to observed values. |

# 5. Discussion

This study examined the role of leptin in metabolism and immune response regulation. The sample included 30 males and 30 females with the first TB episode from Uganda. Our findings revealed that females had higher body fat storage, higher leptin levels, and stronger immune responses. Body fat was positively associated with leptin levels, and leptin mediated the effect of body fat on immune activation.

Our findings align with previous studies showing that females tend to have a stronger immune response than males (17). This difference in immunity may influence susceptibility and the speed of infection clearance between sexes (18). This study focused on immune activation and identified a sex-based difference in CD4+ cell levels, consistent with findings from a previous study (19). Recent research also indicates that changes in body fat in males can significantly reduce CD8+ cell levels compared to females (20). Our study, along with previous research, provides evidence of sex-specific differences in the initiation and dynamic response of the immune system to TB infection. Our study also identified leptin as a bridge between metabolism and the immune system. This suggests its potential as a therapeutic target for regulating immune responses based on specific metabolic conditions (8). Animal-based preclinical studies have demonstrated its practical significance (21), but further research is needed to assess its clinical applicability in human populations. The extended two sections demonstrate the better performance of the random forest model in predicting leptin levels using body composition indices. Its strong performance is further demonstrated when non-linear relationships exist.

This study has several limitations. First, the small sample size reduced the statistical power of the analysis. Sample loss due to the instability of the deuterium dilution approach in measuring body water further limited the detection of significant associations. Second, non-biological factors such as professional exposure, alcohol use, and smoking status (22) were not controlled, which could have influenced immune responses. Given their higher prevalence among males (23), these factors may have confounded the findings. the machine learning component of this study serves primarily as a methodological exercise due to the limited sample size. A substantially larger data set would be required to enable a more robust and convincing comparison between the random forest and regression tree models. However, recruiting a sufficiently large sample of patients experiencing their first episode of TB in high-prevalence regions remains a significant challenge. Despite these limitations, this study provides valuable insights that complement previous research. Future studies should explore the dynamic regulatory role of leptin at different stages of immune responses.

# References

1. Organization WH. Global tuberculosis report 2024. World Health Organization; 2024.

2. Horton KC, MacPherson P, Houben RM, et al. Sex differences in tuberculosis burden and notifications in low-and middle-income countries: A systematic review and meta-analysis. *PLoS medicine*. 2016;13(9):e1002119.

3. Ledesma JR, Ma J, Vongpradith A, et al. Global, regional, and national sex differences in the global burden of tuberculosis by HIV status, 1990–2019: Results from the global burden of disease study 2019. *The Lancet Infectious Diseases*. 2022;22(2):222–241.

4. Gupta M, Srikrishna G, Klein SL, et al. Genetic and hormonal mechanisms underlying sex-specific immune responses in tuberculosis. *Trends in immunology*. 2022;43(8):640–656.

5. Womack J, Tien PC, Feldman J, et al. Obesity and immune cell counts in women. *Metabolism*. 2007;56(7):998–1004.

6. Ilavská S, Horváthová M, Szabová M, et al. Association between the human immune response and body mass index. *Human immunology*. 2012;73(5):480–485.

7. Mupere E, Malone L, Zalwango S, et al. Wasting among uganda men with pulmonary tuberculosis is associated with linear regain in lean tissue mass during and after treatment in contrast to women with wasting who regain fat tissue mass: Prospective cohort study. *BMC infectious diseases*. 2014;14:1–10.

8. Naylor C, Petri WA. Leptin regulation of immune responses. *Trends in molecular medicine*. 2016;22(2):88–98.

9. Francisco V, Pino J, Campos-Cabaleiro V, et al. Obesity, fat mass and immune system: Role for leptin. *Frontiers in physiology*. 2018;9:640.

10. Lord G. Role of leptin in immunology. *Nutrition reviews*. 2002;60(suppl\_10):S35–S38.

11. Cava AL, Matarese G. The weight of leptin in immunity. *Nature Reviews Immunology*. 2004;4(5):371–379.

12. Hickey MS, Israel RG, Gardiner SN, et al. Gender differences in serum leptin levels in humans. *Biochemical and molecular medicine*. 1996;59(1):1–6.

13. Buyukoglan H, Gulmez I, Kelestimur F, et al. Leptin levels in various manifestations of pulmonary tuberculosis. *Mediators of inflammation*. 2007;2007(1):064859.

14. IAEA., IAEA. Introduction to body composition assessment using the deuterium dilution technique with analysis of saliva samples by fourier transform infrared spectrometry. International Atomic Energy Agency; 2011.

15. Sobel ME. Asymptotic confidence intervals for indirect effects in structural equation models. *Sociological methodology*. 1982;13:290–312.

16. Baron RM, Kenny DA. The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of personality and social psychology*. 1986;51(6):1173.

17. Klein SL, Flanagan KL. Sex differences in immune responses. *Nature Reviews Immunology*. 2016;16(10):626–638.

18. Takahashi T, Iwasaki A. Sex differences in immune responses. *Science*. 2021;371(6527):347–348.

19. Bothamley GH. Male sex bias in immune biomarkers for tuberculosis. *Frontiers in immunology*. 2021;12:640903.

20. Dhanyalayam D, Thangavel H, Sidrat T, et al. The influence of body fat dynamics on pulmonary immune responses in murine tuberculosis: Unraveling sex-specific insights. *International Journal of Molecular Sciences*. 2024;25(13):6823.

21. Fantuzzi G, Faggioni R. Leptin in the regulation of immunity, inflammation, and hematopoiesis. *Journal of leukocyte biology*. 2000;68(4):437–446.

22. Wigger GW, Bouton TC, Jacobson KR, et al. The impact of alcohol use disorder on tuberculosis: A review of the epidemiology and potential immunologic mechanisms. *Frontiers in immunology*. 2022;13:864817.

23. Soh AZ, Chee CBE, Wang Y-T, et al. Alcohol drinking and cigarette smoking in relation to risk of active tuberculosis: Prospective cohort study. *BMJ open respiratory research*. 2017;4(1).