Thesis progression

Thesis title: Diagnostic potential of urine FTIR spectroscopy combined with machine learning for active tuberculosis detection

Thesis progression title: Sample preparation

Student: Kanyanat Pattanang Student ID: 685070020-0

Advisor: Asst. Prof. Dr. Wises Namwat

Co-advisor: Asst. Prof. Dr.Sakawrat Kanthawong

Date: 15/03/2026

1. Introduction

T uberculosis (TB), caused by Mycobacterium tuberculosis (MTB), remains a major global public health concern, as MTB is a facultative intracellular pathogen capable of infecting and persisting in humans for decades despite an intact immune system.(1,2). Approximately 5–10% of people infected with TB will eventually get symptoms and develop TB disease. This organism predooms innately causes pulmonary tuberculosis. However, it is also capable of spreading to and affecting various other organs in the body (3). The global increase in TB incidence that emerged during the COVID-19 pandemic has slowed and begun to stabilize (4). In 2024, an estimated 10.7 million people (95% uncertainty interval [UI]: 9.9–11.5 million) fell ill with TB incident cases and 1.23 million died from the disease (95% UI: 1.13–1.33 million). The TB incidence rate was 131 cases per 100,000 population per year (95% UI: 122–141), with a case fatality rate of 11.5% (1). During 2022–2024, there were 1,268 TB clusters involving 3,794 cases. Clusters of 2 or 3 cases comprised 85% of all clusters. There were 33 large clusters (≥10 cases) that represented 3% of all clusters and contained 599 cases (16% of all clustered cases)(5). Consistent with previous years, approximately 90% of TB cases had HIV test results, among whom 5% were reported to have TB–HIV co-infection, highlighting HIV infection as a key medical risk factor for TB (5) In Thailand, TB remains a significant public health burden, had 123,000 new tuberculosis cases in 2022 (6).

To reduce the burden of tuberculosis, it is necessary to move away from traditional diagnostic methods such as culture, which takes a relatively long time due to the slow growth of the pathogen, and the acid-fast bacillus (AFB) test cannot between active tuberculosis (TB) and latent TB infection (7,8) or PCR-based assays (GeneXpert) may yield false-positive or false-negative results, particularly in samples with a low bacillary load (Song et al., 2025). Therefore, there is a critical need for novel diagnostic methods that enable rapid detection while minimizing patient discomfort and avoiding invasive procedures.

Attenuated total reflection–Fourier transform infrared (ATR-FTIR) spectroscopy is an effective option due to its sensitive and non-destructive technique. ATR-FTIR spectroscopy can extract intricate biochemical information from the vibrational energy inherent in chemical bonds found within biomolecules, such as lipids, nucleic acids, proteins, and carbohydrates when applied to liquid matrices (9). There is a rapidly increasing number of reports for the analysis of biofluids (saliva, blood, serum, etc) (9–13). Urine represents a promising alternative to blood or sputum, as it contains more than 3,000 metabolites or metabolic species that can be exploited for diagnostic purposes. The most abundant organic constituents of urine include urea, creatinine, hippuric acid, and citric acid (14). Several studies using urine have focused on disease investigation such as autism (15), early-stage chronic kidney disease (16) and Diabetes (9). ATR-FTIR has been successfully used in diagnosis as well.

While FTIR offers high sensitivity and molecular specificity, challenges such as overlapping bands and matrix effects can complicate spectral interpretation. To address these limitations, chemometric methods and machine learning algorithms, such as Principal Component Analysis (PCA), Random Forest (RF), and Support Vector Machines (SVM), are employed to extract meaningful patterns from complex spectral data.18 These tools enhance the diagnostic potential of FTIR spectroscopy, making it a powerful approach for applications such as identifying disease-specific biomarkers and developing predictive diagnostic models (17).

Accordingly, this study aims to distinguish individuals with active TB from healthy controls using urine samples analyzed by FTIR spectroscopy in combination with machine learning approaches.

2. Objective

1. To calculate sample size

2. To collect urine from active tuberculosis patients and healthy people

3. Materials and methods

3.1 Sample size

$$n = \frac{\left(Z^2_{1-\frac{\alpha}{2}}\right) P(1-P)}{e^2}$$

$$n = \frac{(1.96)^2(0.157)(1-0.157)}{0.05^2}$$

$$n \approx 204$$

n = Sample size

$Z^2_{1-\frac{\alpha}{2}}$ = The confidence level (1-alpha) 95% is 1.96

P = The prevalence of tuberculosis disease in Thailand is 0.157

e = The maximum tolerated error 5%

Accordingly, a total of 204 samples were required to achieve 95% accuracy with a 5% error margin. These were equally divided into 102 healthy controls and 102 active TB cases.

3.2 Sample collection

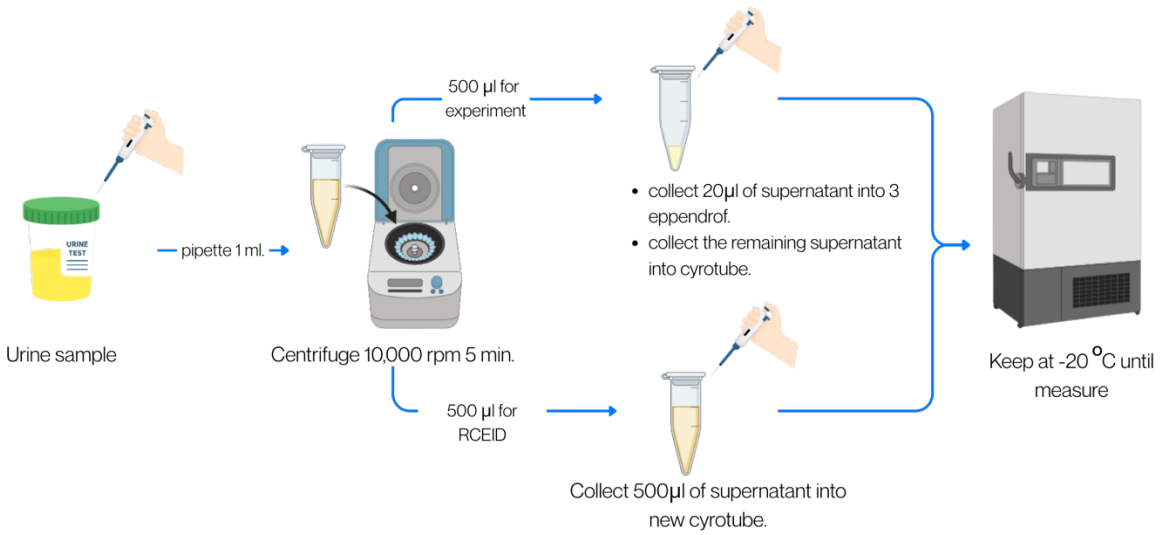
1. Collect urine samples from volunteers who have passed the initial screening using a structured questionnaire based on established criteria.

	Active TB	**Healthy control**
Inclusion criteria	1. Aged 18-70 years old 2. Identified based on clinical symptoms, including chronic cough (≥2 weeks), hemoptysis, fever, loss of appetite, or weight loss. 3. Abnormal chest X-ray findings and smear-positive status, either before treatment or after no more than 2 weeks of treatment. 4. IGRA-positive 5. GeneXpert-positive	1. Aged 18-70 years old 2. were defined as healthy individuals with no confirmed exposure to tuberculosis. 3. Normal chest X-ray findings, no history of tuberculosis, and no history of close contact with tuberculosis patients or household exposure to TB. 4. IGRA-negative 5. GeneXpert-negative
Exclusion criteria (applied to all participants)	1. Age under 18 years or over 70 years. 2. HIV co-infection.	

	3. Presence of underlying diseases, including hepatitis or cirrhosis, anemia or thalassemia, diabetes, immunodeficiency, other chronic lung diseases, or kidney disease. 4. Pregnancy. 5. Infection with nontuberculous mycobacteria (NTM)

All samples, approximately 30 mL of mid-stream urine was collected into sterile containers and immediately placed on ice (4 °C) until sample preparation (18), along with their clinical data, were collected as leftover specimens from the Research and Diagnostic Center for Emerging Infectious Diseases (RCEID), Faculty of Medicine, Khon Kaen University.

3.3: ATR-FTIR spectral acquisition



1. The urine samples were centrifuged at 10,000 rpm for 5 minutes.
2. A volume of 500 µL was aliquoted into two microcentrifuge tubes.
3. The supernatant was collected, and the pellet was discarded.
4. One aliquot was transferred to the Research and Emerging Infectious Diseases Center (REICD) for storage.
5. The remaining aliquot was further subdivided into 20 µL portions in triplicate and stored at −20 °C until further analysis.

4. Results

5. Conclusion

6. References

1. World Health Organization. Global tuberculosis report 2025 [Internet]. 2025 [cited 2025 Dec 29]. Available from: https://iris.who.int/server/api/core/bitstreams/e97dd6f4-b567-4396-8680-717bac6869a9/content

2. Kalscheuer R, Palacios A, Anso I, Cifuente J, Anguita J, Jacobs WR, et al. The Mycobacterium tuberculosis capsule: a cell structure with key implications in pathogenesis. Biochem J. 2019 Jul 18;476(14):1995. doi:10.1042/BCJ20190324 PubMed PMID: 31320388.

3. Buhas BA, Muntean LAM, Ploussard G, Feciche BO, Andras I, Toma V, et al. Renal Cell Carcinoma Discrimination through Attenuated Total Reflection Fourier Transform Infrared Spectroscopy of Dried Human Urine and Machine Learning Techniques. Int J Mol Sci. 2024 Sep 1;25(18):9830. doi:10.3390/IJMS25189830/S1 PubMed PMID: 39337322.

4. World Health Organization. Global tuberculosis report 2024. Geneva; 2024. Report.

5. Centers for Disease Control and Prevention (CDC). Reported Tuberculosis in the United States, 2024 [Internet]. Atlanta, GA; 2025. Report. Available from: https://www.cdc.gov/tb-surveillance-report-2024/executive-commentary/index.html

6. World Health Organization. Global tuberculosis report 2022. Annual Global TB Report of WHO [Internet]. 2022 [cited 2026 Jan 17];8(1):1–68. Available from: https://www.who.int/teams/global-tuberculosis-programme/tb-reports/global-tuberculosis-report-2022%0Ahttps://www.who.int/teams/global-tuberculosis-programme/tb-reports/global-tuberculosis-report-2022#:~:text=context of global...-,Download,-Read More%0Ahtt

7. Tobin EH, Tristram D. Tuberculosis Overview. StatPearls. 2024 Dec 22. PubMed PMID: 28722945.

8. Tedjo A, Karuniawati A, Prawoto P, Riyanto SA, Ibrahim AS. Examination of Acid-Fast Bacilli in Sputum using Modified Light Microscope with Homemade Light Emitting Diode Additional Attachment. Makara Journal of Health Research. 2012 Jan 1;15(2):81–5. doi:10.7454/MSK.V15I2.941

9. Farooq S, Zezell DM. Diabetes Monitoring through Urine Analysis Using ATR-FTIR Spectroscopy and Machine Learning. Chemosensors. 2023 Nov 1;11(11). doi:10.3390/chemosensors11110565

10. Tiernan H, Byrne B, Kazarian SG. ATR-FTIR spectroscopy and spectroscopic imaging for the analysis of biopharmaceuticals. Spectrochim Acta A Mol Biomol Spectrosc. 2020 Nov 5;241:118636. doi:10.1016/J.SAA.2020.118636 PubMed PMID: 32610215.

11. Sitnikova VE, Kotkova MA, Nosenko TN, Kotkova TN, Martynova DM, Uspenskaya M V. Breast cancer detection by ATR-FTIR spectroscopy of blood serum and multivariate data-analysis. Talanta. 2020 Jul 1;214. doi:10.1016/j.talanta.2020.120857 PubMed PMID: 32278436.

12. Lee HC, Delrue C, De Bruyne S, Speeckaert MM. Academic Editors: Nicusor Iftimia The Promise of Infrared Spectroscopy in Liquid Biopsies for Solid Cancer Detection [Internet]. 2025. doi:10.3390/diagnostics

13. Alajaji SA, Sabzian R, Wang Y, Sultan AS, Wang R. A Scoping Review of Infrared Spectroscopy and Machine Learning Methods for Head and Neck Precancer and Cancer Diagnosis and Prognosis. Cancers. Multidisciplinary Digital Publishing Institute (MDPI); 2025. doi:10.3390/cancers17050796

14. Jeng MJ, Sharma M, Lee CC, Lu YS, Tsai CL, Chang CH, et al. Raman Spectral Characterization of Urine for Rapid Diagnosis of Acute Kidney Injury. J Clin Med. 2022 Aug 1;11(16). doi:10.3390/jcm11164829

15. Sarigul N, Bozatli L, Kurultak I, Korkmaz F. Using urine FTIR spectra to screen autism spectrum disorder. Scientific Reports 2023 13:1. 2023 Nov 9;13(1):19466-. doi:10.1038/s41598-023-46507-z PubMed PMID: 37945643.

16. Rachawangmuang P, Chatchawal P, Tippayawat P, Jusakul A, Kraiklang R, Lert-itthiporn W, et al. Discrimination of urine infrared spectral biomarkers for early-stage chronic kidney disease patients using attenuated total reflectance fourier transform infrared spectrometry. Clinica Chimica Acta. 2026 Jan 15;579. doi:10.1016/j.cca.2025.120665 PubMed PMID: 41101440.

17. Lin M, Lu HC, Lin HW, Pan SW, Cheng BM, Tseng TR, et al. Fast Screening of Tuberculosis Patients Based on Analysis of Plasma by Infrared Spectroscopy Coupled with Machine Learning Approaches [Internet]. Vol. 10. 2025;10:11817–27. doi:10.1021/acsomega.4c07990

18. Organization WH. Guidelines for the collection of clinical specimens during field investigation of outbreaks [Internet]. World Health Organization; 2000 [cited 2026 Feb 8]. Available from: https://iris.who.int/handle/10665/66348