Who is truly left behind? Diagnostic bias and the implication of GeneXpert inequity on tuberculosis reporting

ITB 3P Research Program 2026 (Postdoctoral Fellowship 3P)

Nuning **Nuraini**, Arrofiatuz **Zahra**, Kamal Khairudin **Sukandar**, Heni **Widayani**, Ilham Saiful **Fauzi**, Syaribah Noor **Brice**, Swasono **Rahardjo**, Bony Wiem **Lestari**

Introduction

Tuberculosis (TB) remains **a serious global health threat**, and in 2023 it surpassed COVID-19 as the world's leading cause of death from a single infectious agent. TB is not only a medical infection but also a complex socio-economic problem, particularly in Indonesia, which has the **second highest TB burden worldwide**. Often called a **"poverty disease",** TB disproportionately affects vulnerable populations: overcrowded slum areas with poor air circulation allow its bacteria to spread unchecked, while economic hardship leads to food insecurity and malnutrition that weaken immunity. The World Health Organization (WHO) endorses rapid molecular tests such as GeneXpert MTB as the initial diagnostic for people with TB symptoms. In Indonesia's decentralized healthcare system, the unequal distribution of diagnostic facilities can lead to disparities in disease detection and access to treatment. Presidential Regulation No. 67 of 2021 aims to expand TB diagnostic coverage toward the 2030 elimination target, yet implementation varies across regions. West Java, as the most populous province in Indonesia, has heterogeneous health services, ranging from dense urban areas to remote rural regions. This disparity poses both a challenge and an opportunity to understand the real distribution and impact of existing diagnostic capacity.

Keyword:
Tuberculosis, GeneXpert, Underreporting, Diagnostic access, Bayesian hierarchical model, Network centrality, Spatial vulnerability, West Java, Indonesia

Research Objectives

This study aims to fill gap by:

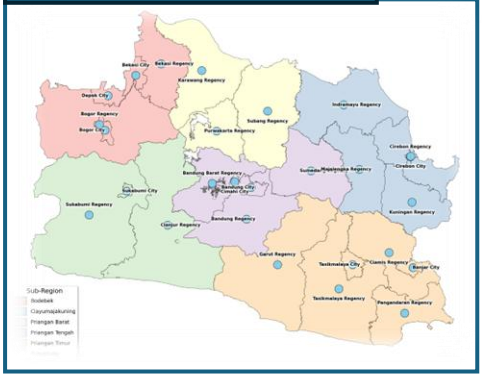
(i) **Evaluating the spatial distribution of GeneXpert** facilities in West Java while accounting for socio-economic factors,

(ii) **Applying a Bayesian inference** approach to correct for reporting biases, thereby identifying regions with a high underlying TB burden, and

(iii) **Mapping inter-regional vulnerability networks** driven by the socio-economic determinants of TB spread.

By correcting this data bias, this research has aimed to refine the estimated disease burden so it can support the intervention strategy for better-targeted interventions and when allocating the GeneXpert in West Java in the future.

Study Settings



Methods

Stage 1: Evaluating Diagnostic Resource Allocation

The Poisson Regression model was used to identify which socio-economic, environmental, and demographic factors are likely to serve as a foundation for locating GeneXpert tools in West Java.

$$GX_i \sim Poisson(\mu_i)$$
$$\log(\mu_i) = a_0 + \sum_{k=1}^{s} a_k x_{ik}$$

where a_0 is the baseline intercept for the natural log of the expected GeneXpert count, and s represents the number of variable included in the model. Predictors screened with VIF (VIF > 10 excluded) to avoid multicollinearity

Stage 2: Bayesian Hierarchical Modeling for Diagnostic Bias Correction

A model hierarchical structure can be expressed as

$$n_i \sim Poisson(\gamma_i N_i)$$
$$\text{logit}(\gamma_i) = \sigma_i + \alpha_i GX_i$$
$$y_i \sim Binomial(n_i, \beta_i)$$

From the model, we identify two important parameters for further analysis: β and γ, which represent the positivity rate and testing rate, respectively.

Regional differences tested with Kruskal-Wallis and Dunn's post hoc test. Then, classified into 4 quadrants using the posterior probability that its β and γ exceed the median.

Stage 3: Spatial Vulnerability Network and Centrality Analysis

Epicenters were identified by constructing a complete weighted network of regions, where edge weights were based on socioeconomic similarity alone or combined with geographical distance, and then ranking regions using Betweenness, Eigenvector, and PageRank centrality.

$$SI(i,j) = \frac{1}{1 + d(i,j)}$$

where $d(i,j)$ represents the Euclidean distance of the socioeconomic profiles between regions i and j, namely $d(i,j) = \sqrt{\sum_k (Z_{ik} - Z_{jk})^2}$ where Z_{ik} is the standardized value of the k-th feature in region i.

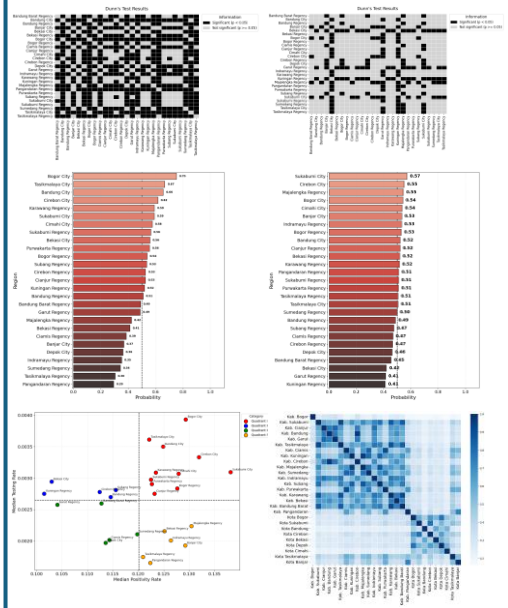
Two scenarios for edge weights:

(1) Socioeconomic only, $W(i,j) = SI(i,j)$

(2) Socioeconomic and geography,

$$W(i,j) = \frac{SI(i,j)}{1 + d_{geo}(i,j)}$$

Results



Key Findings:

- GeneXpert allocation in West Java **appeared to follow population size and poverty levels**.
- Reported TB cases may be partly shaped by uneven testing: **positivity rates were largely uniform across the province while testing rates varied spatially**.
- Most regions fell into Quadrant I (high testing, high positivity), whereas Quadrants III and IV (low testing) may **conceal underreporting or undetected transmission** and may **warrant active case finding and expanded diagnostics**.
- In the spatial vulnerability network, **regencies tended to resemble other regencies, and cities other cities.**
- **Ciayumajakuning emerged as a potential priority area**, with Majalengka (Quadrant IV) ranking highest in Eigenvector and PageRank centrality, while Sumedang, Garut, and Depok appeared as potential spatial brokers.

Output

1. Dayat Hidayata, Ilham Saiful Fauzi, **Nuning Nuraini**, *Mathematical model of cellular and humoral immune responses of dengue virus infection with hematocrit dynamics*, Microbial Pathogenesis, Volume 218,108641,(2026)
2. Stefanus Raynaldo Septyawan, **Nuning Nuraini**, Muhammad Fakhruddin, Chai Jian Tay, Asep K. Supriatna, Edy Soewono, *Modeling dengue transmission in Bandung: Integrating vehicle mobility data forenhanced insights*, Applied Math Modelling (156) ,11682 (2026).

This research is funded by the ITB 3P Research Program 2026 (Post-doctoral Fellowship Scheme) through the Directorate of Research and Innovation, Institut Teknologi Bandung