Journal of Public Health Indonesian
Volume.2 Issue.5, (January, 2026) Pages 72-78
E-ISSN: 3048-1139
DOI : https://doi.org/10.62872/zjyppe33
https://nawalaeducation.com/index.php/JHH

Biostatistical Approach to Predict Disease Risk Using Public Health Data

Loso Judijanto[1], Isah Fitriani[2]
IPOSS Jakarta, Indonesia[1], Universitas Ahmad Dahlan, Indonesia[2]

Received: October 12, 2025
Revised: November 03, 2025
Accepted: November 20, 2025
Published: November 29, 2025

Corresponding Author:
Author Name*: Loso Judijanto
Email*:
losojudijantobumn@gmail.com

Abstract: *The increasing complexity of public health issues demands an analytical approach capable of optimally utilizing data to support disease prevention efforts. The increasing availability of public health data opens up opportunities for the development of evidence-based predictive approaches. This study aims to examine the role of biostatistical approaches in predicting disease risk using public health data and its implications for preventive efforts and health policy. The study employed a qualitative approach using literature review methods, including journal articles, academic books, and relevant policy documents. Data analysis was conducted thematically to identify the role of biostatistics in risk factor analysis, predictive model development, and the associated methodological and policy challenges. The study results indicate that biostatistical approaches play a crucial role in identifying multifactorial relationships between health determinants and disease incidence at the population level. Disease risk prediction models have been shown to support the identification of high-risk groups and the planning of more efficient preventive interventions. Key challenges include data quality, limited human resources, and gaps in the translation of analysis results into health policy. Overall, the biostatistical approach is a strategic foundation for the development of a data- and evidence-based public health system oriented towards disease prevention.*
Keywords : Biostatistics; Public Health Data; Disease Risk Prediction

How to cite:
Judijanto L, Fitriani I (2026). Biostatistical Approach to Predict Disease Risk Using Public Health Data. *Journal of Public Health Indonesian,* 2(5), 72-78. DOI: https://doi.org/10.62872/zjyppe33

INTRODUCTION

Changing disease patterns in modern populations indicate a significant shift from the dominance of infectious diseases to an increase in chronic and degenerative non-communicable diseases. This situation is exacerbated by the emergence of new infectious diseases characterized by rapid spread and widespread impacts on public health. The complex interactions between biological, behavioral, environmental, and socioeconomic factors shape disease risk profiles that can no longer be explained simply. The inability of conventional approaches to capture this complexity has the potential to result in ineffective and poorly targeted health policies. The need for analytical methods capable of integrating various risk factors becomes increasingly urgent as the burden of disease at the population level increases. Disease prevention efforts require a strong scientific foundation so that interventions can be implemented before disease progresses further (Olowe et al., 2024). Predictive approaches are a key element in shifting the paradigm of health care from curative to preventive. Population-based disease risk prediction opens up opportunities for more proactive and sustainable health planning (Lestari, 2024).

Journal of Public Health Indonesian
Volume.2 Issue.5, (January, 2026) Pages 72-78
E-ISSN: 3048-1139
DOI : https://doi.org/10.62872/zjyppe33
https://nawalaeducation.com/index.php/JHH

Advances in information technology have driven health systems to generate vast amounts of diverse data. Public health data now encompasses demographic information, health status, lifestyle behaviors, environmental factors, and access to healthcare services. This abundance of data holds enormous potential to comprehensively describe the health condition of populations (Kogevinas et al., 2022; Li, 2024). The use of data, which is still descriptive in nature, means that much important information has not been processed into strategically valuable knowledge. This gap between data availability and its analytical use poses a unique challenge for health system managers. Data that is not properly processed risks remaining merely an administrative archive without any real contribution to improving public health. Transforming raw data into predictive information requires a systematic and structured scientific approach. The use of appropriate analytical methods is a key prerequisite for health data to have a real impact.

Biostatistics exists as a scientific discipline capable of bridging the needs of complex health data analysis. Biostatistical principles enable objective, measurable, and probability-based data processing. The relationship between various risk factors and disease incidence can be analyzed quantitatively through valid statistical approaches (Shafira & Haritsm 2025). Biostatistics' ability to identify hidden patterns makes it relevant for large-scale public health data analysis. Appropriate analytical methods can reduce bias and improve the accuracy of health research results. Interpretation of biostatistics-based analysis results provides a strong scientific basis for decision-making. The use of biostatistics also allows for a more measurable evaluation of the effectiveness of health interventions. The integration of biostatistics with public health data is a crucial foundation for the development of disease risk prediction models.

The biostatistical approach serves not only as an analytical tool but also as a means to build applicable predictive models. Disease risk prediction models allow estimating the probability of disease occurrence in specific population groups. This information can be used to identify high-risk groups earlier. Accurate risk identification provides opportunities for more targeted preventive interventions. The efficiency of health resource use can be improved through data-driven risk mapping. The accuracy of predictive models is highly dependent on the quality of the data and the analytical methods used. The biostatistical approach offers a systematic framework for testing the validity and reliability of models. The development of robust models contributes to improving the quality of public health planning (Hayati & Pawenang, 2021).

The application of disease risk prediction models has strategic implications for health systems. The resulting risk information can support the formulation of evidence-based health policies. Preventive policies designed based on prediction results have the potential to reduce morbidity and mortality. Data-driven decision-making increases the accountability and transparency of public policy. This approach also helps governments and stakeholders prioritize health programs. Interventions can be focused on the risk factors that most contribute to disease incidence. The impact of health policies is expected to be more effective and sustainable. Integrating risk prediction into the health system strengthens promotive and preventive efforts.

Challenges in implementing a biostatistical approach still require serious attention. Variations in the quality of public health data can impact the accuracy of analysis results. Incomplete data or recording errors have the potential to bias predictive models. Human resource competency in statistical analysis remains a challenge in many regions. Limited understanding of statistical interpretation can hinder the utilization of research findings. The need for consistent methodological standards is increasingly crucial to ensure the validity of results. Capacity building for health data analysts is integral to implementing this approach. Institutional support is needed for optimal implementation of the biostatistical approach. Addressing these challenges will determine the success of public health data utilization (Saraswati, 2024; Izza & Rizmayanti, 2024).

The biostatistical approach also opens up opportunities for integration with other disciplines. Collaboration between epidemiology, data science, and health policy enriches disease risk analysis. A multidisciplinary approach enables a more holistic understanding of health determinants. The complexity of public health problems demands the synergy of diverse expertise. This integration strengthens the predictive and interpretative capabilities of analysis results. The resulting models become more relevant to real-world conditions. The collaborative approach also increases the acceptance of research results by policymakers. Interdisciplinary synergy broadens the scientific and practical impact of the biostatistical approach.

The development of a biostatistical approach to predicting disease risk is an unavoidable necessity. Optimizing the use of public health data offers significant opportunities for improving the quality of population health. Accurate risk prediction supports a paradigm shift toward more effective disease prevention. A data-driven approach strengthens the scientific foundation for health program planning and evaluation. The successful implementation of this approach depends on data quality, analytical methods, and human resource capacity. Policy and infrastructure support are key factors in ensuring its sustainable implementation. The biostatistical approach plays a strategic role in addressing increasingly complex public health challenges. The scientific and practical contributions of this approach are expected to improve overall public health.

METHODOLOGY

This study uses a qualitative approach with a literature review method to examine the role of biostatistics in predicting disease risk based on public health data. The research focuses on conceptual understanding, methodological developments, and the implications of applying biostatistics in evidence-based health decision-making. Data were obtained from various scientific publications, academic books, health organization reports, and relevant policy documents, which were searched through reputable scientific databases. The literature was systematically selected based on criteria of relevance and methodological quality to minimize bias. Data collection was carried out through critical reading to identify key concepts, methods, and findings related to disease risk prediction. Data analysis used thematic analysis techniques to group and interpret key themes in an integrative manner. The validity of the study was maintained through source triangulation, methodological transparency, and consistency of interpretation. The results of the study are presented narratively and analytically to map the contributions, benefits, and challenges of the biostatistics approach, and provide a conceptual basis for policy development and further research.

RESULTS AND DISCUSSION

The Role of Biostatistical Approaches in Identifying and Analyzing Disease Risk Factors Based on Public Health Data

Biostatistical approaches play a central role in unraveling the complex relationships between risk factors and disease incidence at the population level. Statistical analysis allows the identification of risk patterns that are not always apparent through descriptive observations alone. Heterogeneous public health data can be processed into structured information through the application of appropriate statistical methods. Relationships between variables such as age, sex, socioeconomic status, and health behaviors can be analyzed simultaneously. This process helps uncover the multidimensional determinants of disease. The

Journal of Public Health Indonesian
Volume.2 Issue.5, (January, 2026) Pages 72-78
E-ISSN: 3048-1139
DOI : https://doi.org/10.62872/zjyppe33
https://nawalaeducation.com/index.php/JHH

precision of biostatistical analysis provides a strong scientific basis for epidemiological interpretation. The resulting information has strategic value for disease control; thus, biostatistics serves as a key analytical instrument in understanding population-based disease risk (Sakarna et al., 2025).

Utilizing public health data requires the ability to handle data variability and uncertainty. A biostatistical approach provides a mathematical framework for systematically managing this variation. The resulting risk estimates are based on the principles of probability and statistical inference. This analysis allows for measuring the strength of the relationship between risk factor exposure and disease incidence. Interpretation of the analysis results provides a quantitative overview of the contribution of each factor. This approach also helps distinguish causal from noncausal relationships. Clarity about these relationships is crucial for designing appropriate health interventions. Therefore, biostatistics contributes significantly to prioritizing risk factors.

Literature shows that a biostatistical approach can improve the accuracy of identifying vulnerable groups. Grouping populations based on risk levels allows for more accurate segmentation. This process supports more focused health program planning. Population-based risk analysis reduces reliance on general assumptions that often do not reflect local conditions (Sidhu, 2024). A statistical approach allows for model adjustments to specific population characteristics. The flexibility of biostatistical methods enhances the relevance of analysis results. Findings can be tailored to the needs of policymakers, thus making risk analysis more contextual and applicable.

The reliability of biostatistical analysis results depends heavily on the quality of the data used. The literature emphasizes the importance of the validity and reliability of public health data (Soviadi et al., 2025). Data cleaning and bias control processes are integral to the analysis. Statistical approaches enable the systematic detection of data errors. Careful data handling improves the accuracy of risk estimates. Valid analysis strengthens trust in research results. This trust is crucial for stakeholder acceptance of the results. Therefore, data quality and analytical methods are inseparable.

The biostatistical approach also contributes to the development of a conceptual framework for disease risk analysis. The literature shows that integrating epidemiological and statistical theories yields a more comprehensive understanding (Tanjung et al., 2023; Xue et al., 2024). The analysis focuses not only on numbers but also on the broader meaning of public health. This approach allows for a synthesis of empirical evidence and theoretical concepts. Understanding of disease risk becomes more systematic and structured (Chu et al., 2019). The analytical process enriches scientific discourse on the determinants of health. This contribution strengthens biostatistics' position as the foundation of public health analysis. Thus, the biostatistical approach plays a strategic role in the development of public health science.

Development and Utilization of Disease Risk Prediction Models to Support Public Health Preventive Efforts

The development of disease risk prediction models is one of the main applications of the biostatistical approach. These models are designed to estimate the probability of disease occurrence based on population characteristics. The literature shows that predictive models are capable of simultaneously integrating multiple risk factors (Mustopa, 2025). The modeling process simplifies data complexity without losing analytical meaning. The resulting risk estimates provide a prospective picture of health threats. This predictive information has significant value for health planning. This approach encourages a shift in focus from response to disease to prevention. Therefore, predictive models are a strategic tool in public health.

The use of risk prediction models allows for early identification of high-risk groups. This information can be used to direct preventive interventions more efficiently. The literature confirms that risk-based interventions tend to be more effective than universal approaches. The use of health resources can be optimized through data-driven prioritization. Predictive models help reduce resource waste on poorly

Journal of Public Health Indonesian
Volume.2 Issue.5, (January, 2026) Pages 72-78
E-ISSN: 3048-1139
DOI : https://doi.org/10.62872/zjyppe33
https://nawalaeducation.com/index.php/JHH

targeted interventions (Karo-Karo et al., 2026; Indriana et al., 2025). This efficiency is crucial in health systems with limited budgets. Predictive approaches also support the sustainability of health programs. Thus, the use of predictive models strengthens preventive strategies.

The quality of a predictive model is greatly influenced by the biostatistical methods used. The literature highlights the importance of selecting appropriate variables and analysis techniques. Models that are too simple risk overlooking important factors, while models that are too complex are difficult to interpret. The balance between accuracy and interpretability is a key concern. Evaluating model performance through internal and external validation improves prediction reliability. The evaluation process ensures that the model can be applied to diverse populations. Model reliability strengthens user confidence in the prediction results. Therefore, model development requires careful methodological consideration.

Risk prediction models also serve as scientific communication tools. Prediction results presented quantitatively facilitate understanding for policymakers. Risk information can be translated into relevant indicators for program planning. Literature shows that visualizing prediction results improves the usability of information (Lumingkewas & Mokodaser, 2025). Clear presentation helps bridge the gap between scientific analysis and public policy. Predictive models serve as a means of translating knowledge from research to practice. This process strengthens the relationship between academics and decision-makers. Thus, predictive models serve as a bridge between science and policy.

The development of risk prediction models is inseparable from the social and demographic dynamics of the population. The literature emphasizes the need for regular model updates. Changing disease patterns and risk factors affect prediction accuracy. Adaptive models are better able to reflect actual health conditions. A biostatistical approach allows model adjustments to these changes. This flexibility enhances the long-term relevance of predictive models. Models that are responsive to change support the sustainability of preventive efforts. Therefore, predictive model development is an ongoing process.

Methodological Challenges and Policy Implications in Applying Biostatistical Approaches to Public Health Data

The application of biostatistical approaches faces various complex methodological challenges. The literature shows that the quality of public health data often varies. Incomplete and inconsistent data can affect analysis results. Statistical approaches require certain assumptions that are not always met by field data. Violation of these assumptions can potentially reduce the validity of the results. Addressing these issues requires adequate methodological expertise. These challenges require careful interpretation of analysis results. Therefore, methodological quality is a primary concern in the application of biostatistics.

Limited human resources also pose a significant challenge. Astuti et al. (2024) highlighted the shortage of biostatistics experts in many health systems. Limited analytical skills hinder optimal data utilization. Training and capacity building are urgent needs. Analyst competence impacts the quality of modeling and interpretation of results. Analytical errors can lead to inappropriate policies. Investing in human resource development is a strategic step. Thus, institutional capacity influences the successful implementation of biostatistical approaches (Andriani et al., 2025).

The policy implications of biostatistics applications depend heavily on the translation of analytical results. The literature indicates a gap between research findings and policy implementation. Technical statistical language is often difficult for policymakers to understand. Knowledge translation is a challenge. Simplifying without losing scientific meaning requires strong communication skills. Failure to translate analytical results can diminish policy impact. Therefore, an interdisciplinary approach is increasingly important. Integration of statistical analysis and policy strengthens the effectiveness of health interventions.

Journal of Public Health Indonesian
Volume.2 Issue.5, (January, 2026) Pages 72-78
E-ISSN: 3048-1139
DOI : https://doi.org/10.62872/zjyppe33
https://nawalaeducation.com/index.php/JHH

The application of a biostatistical approach also has ethical and social implications. The literature raises the issue of using sensitive public health data. Protecting data privacy and confidentiality is crucial. Population-based risk analysis must address the potential stigmatization of specific groups. Risk-based policies need to be designed in an inclusive and equitable manner. An ethical approach strengthens the legitimacy of health data use. Public trust in the health system is influenced by data analysis practices. Therefore, ethical considerations must be integrated into the application of biostatistics.

Despite facing various challenges, the biostatistical approach still has positive implications for health policy. The literature shows that data-driven policies tend to be more effective and accountable. Risk analysis provides a rational basis for resource allocation. Preventive policies can be designed in a more targeted manner. This approach supports transparent decision-making. Methodological challenges can be addressed through strengthening systems and capacities. The integration of biostatistical analysis into health policy improves the quality of governance. Thus, the biostatistical approach plays a strategic role in public health policy development.

CONCLUSIONS

Biostatistical approaches have proven to play a strategic role in identifying and analyzing disease risks based on public health data. Statistical analysis enables a more comprehensive understanding of the multifactorial relationships that influence disease incidence at the population level. Utilizing public health data through biostatistical methods can transform raw data into scientifically valuable predictive information. Disease risk prediction models contribute significantly to supporting more targeted and efficient prevention efforts. Predictive approaches encourage a shift in the health paradigm from curative to preventive. The reliability of prediction results is greatly influenced by data quality and the accuracy of the analytical methods used. Methodological challenges and limited human resources remain obstacles to the optimal application of biostatistical approaches. Translating analysis results into health policy requires adequate scientific communication skills. Integration between biostatistical analysis and the policy process strengthens the effectiveness of public health interventions. Ethical and social considerations are crucial aspects in utilizing population-based health data. Strengthening institutional and methodological capacity is necessary to optimize the use of biostatistical approaches. Thus, biostatistical approaches are an essential foundation for developing evidence-based public health policies.

REFERENCES

Andriani, L., Silaen, P. Y., Humayrah, W., Qurniyawati, E., Murdani, A. P., Yuliani, C., ... & Sihombing, I. U. A. (2025). *Pengelolaan dan Pemanfaatan Data Kesehatan: Perspektif Epidemiologi*. Sada Kurnia Pustaka

Astuti, A. D., Nurjanah, A., Sativa, S. Z., Rangkuti, S. R., Nafisah, N., Fitri, N., ... & Nasution, I. S. (2024). Analisis Kebijakan Kesehatan Mendorong Partisipasi Masyarakat Dalam Program Pencegahan Penyakit. *JKEMS-Jurnal Kesehatan Masyarakat*, *2*(2), 96-104.

Chu, S. H., Huang, M., Kelly, R. S., Benedetti, E., Siddiqui, J. K., Zeleznik, O. A., ... & Consortium of Metabolomics Studies Statistics Working Group. (2019). Integration of metabolomic and other omics data in population-based study designs: an epidemiological perspective. *Metabolites*, *9*(6), 117.

Hayati, A. N., & Pawenang, E. T. (2021). Analisis spasial kesehatan lingkungan dan perilaku di masa pandemi untuk penentuan zona kerentanan dan risiko. *Indonesian Journal of Public Health and Nutrition*, *1*(2), 164-171.

Journal of Public Health Indonesian
Volume.2 Issue.5, (January, 2026) Pages 72-78
E-ISSN: 3048-1139
DOI : https://doi.org/10.62872/zjyppe33
https://nawalaeducation.com/index.php/JHH

Indriana, N. P. R. K., Dewi, I. A. U., & Darmayanti, P. A. R. (2025). Prediksi Stunting Berbasis Machine Learning melalui CERDIS: Cepat Resposif Deteksi Dini Stunting. *J-REMI: Jurnal Rekam Medik dan Informasi Kesehatan*, *7*(1), 60-72.

Izza, N. C., & Rizmayanti, A. I. (2024). Analisis Rekam Medis dengan Metode Data Mining untuk Memprediksi Faktor Risiko Stunting dalam Kesehatan Masyarakat. *Jurnal Manajemen Informasi dan Administrasi Kesehatan*, *7*(1), 1-9.

Karo-Karo, J., Syakir, A. R., Raihan, R., Sumanto, S., Budiawan, I., Pakpahan, R., & Christian, A. (2026). Prediksi Penyakit Jantung Menggunakan Algoritma Machine Learning Berdasarkan Indikator Kesehatan. *Jurnal Ilmiah Sistem Informasi*, *5*(1), 84-95.

Kogevinas, M., Chatzi, L., Popay, J., Baum, F., Greenland, S., VanderWeele, T. J., ... & Buehler, J. W. (2022). Epidemiological and biostatistical approaches. *Oxford Textbook of Global Public Health*.

Lestari, M. D. (2024). ANALISIS DATA MINING UNTUK PEMANTAUAN KESEHATAN MASYARAKAT. *Jurnal Dunia Data*, *1*(6).

Li, M. (2024). Exploring the Future of Biostatistics in Genomic Research: Opportunities and Challenges. *Genomics and Applied Biology*, *15*.

Lumingkewas, C., & Mokodaser, W. G. (2025). Integrasi XGBoost dan Visualisasi Gradio untuk Memprediksi Pendapatan Pembayar Asuransi: Studi Kasus Rumah Sakit Swasta di Manado. *Techno. com*, *24*(2).

Mustopa, R. (2025). PENGEMBANGAN MODEL PROMOSI KESEHATAN PENYAKIT TUBERKULOSIS BERBASIS M-Health UNTUK MENINGKATKAN PEMBERDAYAAN KADER. *Jurnal Kesehatan Manarang*, *11*(2).

Olowe, K. J., Edoh, N. L., Zouo, S. J. C., & Olamijuwon, J. (2024). Conceptual frameworks and innovative biostatistical approaches for advancing public health research initiatives. *International Journal of Scholarly Research in Medicine and Dentistry*, *3*(2), 11-21.

Sakarna, Y., Kalakota, S., Gottipati, A., Kaviandost, P., Chileveru, K., Bhola, R., & Singh, S. (2025). Applications of Biostatistics in Healthcare. *Public Health, Epidemiology, and Beyond. Saudi J Med*, *10*(5), 270-276.

Saraswati, D. (2024). Inovasi Pelayanan Kesehatan: Deteksi Dini Penyakit Jantung Koroner melalui Posbindu PTM. *Jurnal Kesehatan dan Kebidanan Nusantara*, *2*(1), 10-16.

Shafira, N. I., & Harits, J. M. (2025). PEMODELAN REGRESI BINOMIAL DENGAN INLA: STUDI KASUS FAKTOR RISIKO PENYAKIT DIABETES. *SINERGI: Jurnal Riset Ilmiah*, *2*(1), 137-146.

Sidhu, G. K. (2024). Understanding the Nexus of Epidemiology and Biostatistics Unveiling the Dynamics of Public Health. *Health Science Journal*, *18*(4), 1-3.

Soviadi, N. V., Anjarwati, N. N., Kep, M., An, S. K., Lesmono, W. D., Oresti, N. S., ... & Kep, M. (2025). *Biostatistik Terapan: Analisis Data Kesehatan untuk Pengambilan Keputusan Berbasis Bukti*. PT Bukuloka Literasi Bangsa.

Tanjung, N., Auliani, R., Rusli, M., Siregar, I. R., & Taher, M. (2023). Peran kesehatan lingkungan dalam pencegahan penyakit menular pada remaja di Jakarta: Integrasi ilmu lingkungan, epidemiologi, dan kebijakan kesehatan. *Jurnal Multidisiplin West Science*, *2*(09), 790-798.

Xue, D., Hajat, A., & Fohner, A. E. (2024). Conceptual frameworks for the integration of genetic and social epidemiology in complex diseases. *Global Epidemiology*, *8*, 100156.