International Journal for Multidisciplinary Research (IJFMR)

E-ISSN: 2582-2160 ● Website: www.ijfmr.com ● Email: editor@ijfmr.com

Data-Driven Forecasting of Morbidity Trends Using ARIMA Models: A Statistical Approach to Public Health Planning

Neil Vincent B. Canama[1], Florence Jean B. Talirongan[2]

[1]Master of Science in Information Technology, Northwestern Mindanao State College of Science and Technology, Christ the King College de Maranding, Inc.
[2]Professor, College of Computer Studies, Misamis University Philippines

Abstract

For the capability to enhance the public health systems and evidence-based policymaking, accurate prediction of disease patterns is crucial. Statistical models provide useful tools for the projection of future health issues, as indicated by the growing demand for proactive healthcare strategies. With the use of past statistics from 2020 to 2024, this present study makes use of the Autoregressive Integrated Moving Average (ARIMA) model to project morbidity trends for major health conditions in the Philippines from 2025 to 2030. Some of the diseases that were reviewed include heart disease, hypertension, diabetes, schistosomiasis, and tuberculosis (TB). Results reflect a steady decrease in deaths related to diabetes, a steady annual increase in heart disease and hypertension cases, and uneven trends in schistosomiasis and tuberculosis cases. The potential of the ARIMA model to effectively predict short-term health outcomes is reflected in the apparent linear and seasonal trends it exhibits across disease. In planning to lower future disease burdens, optimize healthcare resources, and create evidence-based interventions, health authorities must take such findings into account. Statistical forecasting is also stressed to be crucial in enhancing public health preparedness and decision-making.

Keywords: ARIMA model, morbidity trends, public health forecasting, heart disease, hypertension, diabetes mortality, schistosomiasis, tuberculosis, time series analysis, healthcare planning

INTRODUCTION

Monitoring and forecasting morbidity rates are essential for enhancing public health systems, especially in regions far from the availability of healthcare. In rural societies, in most cases, healthcare services are under-funded and do not have the proper staff, thus it is challenging to respond properly to epidemics of disease. Local health units are forced to work reactively only after the number of cases of disease rises, rather than preventing them in advance, because there are no available reliable forecasting systems.

Worldwide, most nations share common problems. India has increased hypertension and diabetes cases that have taxed public health systems in the most underprivileged areas (Gupta et al., 2019) [4]. Sub-Saharan Africa still struggles with high levels of tuberculosis and malaria, frequently with the lack of effective predictive surveillance (World Health Organization [WHO], 2020) [6]. In Brazil, dengue and other vector-borne disease outbreaks have demonstrated a health system's weakness when it has no timely forecasting mechanisms (Teixeira et al., 2018) [5].

The Philippines also has public health problems of its own. The leading causes of death throughout the country are hypertension and heart disease (Department of Health [DOH], 2023) [3]. Tuberculosis is still a problem, as the country boasts one of the highest rates of TB in the world (WHO, 2022) [7]. Tropical diseases like schistosomiasis remain in some places due to sanitation problems and minimal public enlightenment (DOH, 2021) [2]. Efforts to enhance health care notwithstanding, data-driven forecasting is not fully maximized in most municipalities.

One of the forceful tools to combat this deficit is the ARIMA (Autoregressive Integrated Moving Average) model. Applied in economics, epidemiology, ecology, and other disciplines, ARIMA is especially suited for time-based pattern detection and short- to medium-term prediction [1]. When applied in health data analysis, the model allows the authorities to plan ahead efficient use of resources and interventions at the right time. The research uses ARIMA modeling to predict morbidity patterns in five priority conditions, namely heart disease, hypertension, diabetes, schistosomiasis, and tuberculosis. These conditions were chosen based on their consistent presentation in public health and local patient history. By an examination of morbidity trends between 2020 and 2024, this study seeks to provide projections up to 2030 that can be utilized to make policy decisions, enhance planning in healthcare, and ultimately lead to improved local health.

Theoretical framework

Review of related literature: Various researches cite the ARIMA (Autoregressive Integrated Moving Average) model as among the conventional and oftentimes applied methods in forecasting, especially when dealing with time-dependent trends in socio-economic and public health information**.** Talirongan et al. [1] used ARIMA to study national health trends in the Philippines, demonstrating the validity of the model in generating forecasts from historical health data. Rufino [8] applied SARIMA, which is a seasonal extension of ARIMA, to predict monthly tourist arrivals in the Philippines, demonstrating that the model can handle complex seasonality in time series. Similarly, Sangco and Vienes [9] contrasted various forecasting models, including ARIMA, for projecting professional migration from the Philippines, which indicates that ARIMA is able to function efficiently with real population data. Unlike earlier focus on remittance or labor trends, current studies have also noted the application of ARIMA in other sectors of the public sector. For instance, Bautista and Rivera [10] applied ARIMA to forecast hospital admission rates, demonstrating its applicability for healthcare service planning. On the other hand, De Castro et al. [11] used ARIMA for infectious disease case forecasting, illustrating its application in public health surveillance and disease prevention planning.

In addition, Angco et al. [12] used ARIMA to predict economic participation in major Philippine industries and tested its efficacy in dealing with huge economic data. Parreño [13] compared and contrasted ARIMA and Holt-Winters models in predicting agricultural production and determined that ARIMA is very effective for stable, non-seasonal time series. This is an indication of the flexibility of the model, particularly when used with contemporary computational approaches. Finally, Estoque et al. [14] utilized ARIMA in projecting the increase in Philippine urban population, a measure commonly linked to public health issues like disease transmission, hospital bed capacity, and medical facilities availability. The evidence from these studies goes further in supporting the utilization of ARIMA in projecting socio-economic measures that impact morbidity rates. All these results individually confirm that ARIMA remains one of the most reliable and active models of prediction. Its power is that it can model short-term fluctuations and long-term trends from historical data. ARIMA is used in this study as a general framework to examine and predict morbidity patterns in the Philippines to enable data-informed public health planning between 2019 and 2024.

Materials and Methods

Materials

The data employed in this study are morbidity rates in the selected are a of the Philippines, which particularly aim at five chronic diseases: heart disease, hypertension, mortality from diabetes, schistosomiasis, and tuberculosis (TB). These numbers include annually reported cases from 2020 to 2024 and were based on a synthesized dataset using trends seen in official government records such as the Department of Health (DOH) and Philippine Statistics Authority (PSA), augmented by publicly available health surveillance reports and epidemiologic studies [2], [3], [4], [6]. The indicators of health were chosen based on their prevalence and public health impact and because they repeatedly appeared in national morbidity reports. The conditions were selected due to their salience in the impact on healthcare policy, resource management, and long-term public health strategy. The dataset was arranged in yearly periods and was designed to capture trends prior to, during, and following the peak of the COVID-19 pandemic. The five-year trend period of the years 2020 to 2024 was used as the foundation for the time-series forecast model. Microsoft Excel was used for aggregating and pre-processing the data, while Gretl, an open-source econometric tool, was used for ARIMA-based forecasting and analysis.

Methods

This study employed the Autoregressive Integrated Moving Average (ARIMA) model for forecasting morbidity patterns in the Philippines from 2020 to 2024. ARIMA is a well-known statistical model for time-series model formulation and is praised for its strength in identifying both underlying trends as well as random fluctuation in health information [8], [9], [10].

The ARIMA model can be expressed as ARIMA(p, d, q), where:

- **p** refers to the quantity of autoregressive terms,
  Globally, Wang et al. [13] examined distributed ARIMA models for ultra-long time series data to demonstrate its scalability and potential applications to big data environments across the globe. Ghosh et al. [12] obtained hybrid ARIMA-based models developed with neural networks in forecasting disease outbreaks,
- **d** refers to the quantity of differences to be applied so that the data becomes stationary,
- **q** is the value of the lagged forecast error in the equation for the forecast.
  The general form of the ARIMA model is as follows:
  $$Xt = \Phi_1 Xt\text{-}1 + \ldots + \Phi_p Xt\text{-}p + at - \Theta_1 at\text{-}1 - \ldots - \Theta q\ at\text{-}q$$

Where:

- **Xt** is the original time series,
- **Φ's** are the autoregressive coefficients,
- **Θ's** are the moving average coefficients
- **at** represents white noise or random error.

The modeling process involved three main steps:

1. Model Identification: Using autocorrelation (ACF) and partial autocorrelation (PACF) plots to assess the stationarity of the time series and identify the potential values of p and q.
2. Parameter Estimation and Diagnostic Checking: Several ARIMA models were fitted and

compared using criteria such as the Akaike Information Criterion (AIC). Residuals were examined to ensure white noise characteristics.

3. Forecasting: The final ARIMA model was used to predict deployment trends for the selected area in Lanao del Norte, Philippines.

The software used for forecasting and analysis in this study was GRETL (Gnu Regression, Econometrics, and Time-series Library), which offers robust tools for time series modeling and trend visualization. This methodological approach enabled the researcher to estimate future morbidity levels based on historical data patterns. The resulting forecasts can serve as valuable input for public health planning, resource allocation, and the development of evidence-based health interventions.

Results and Discussions

In the trend analysis of morbidity cases in the selected are in the Philippines, GRETL software was employed to examine historical health data and project future developments. Table 1 displays the raw data on reported cases for five major diseases from 2020 to 2024, serving as the basis for forecasting future morbidity patterns.

Forecasting

To gain insights into how morbidity patterns may evolve in the coming years, the ARIMA model was applied to generate forecasts from 2025 to 2030. These predictions were derived from the observed morbidity data collected between 2020 and 2024 for five of the most significant health conditions. The graphical plots shown in Figures 2–4 represent the projected trends, a 95% confidence interval, which is a graphical approximation of the reliability and possible variability of the projections. For heart disease, the projections show a continued increase of 2020 to 2024 and gradual leveling off by 2030. This indicates that although heart diseases are still a public health issue, the rate of growth will slow and stabilize in the long run. Trends for hypertension and diabetes deaths are fairly consistent, with projections showing steady, moderate growth. This is a reflection of the ongoing prevalence of lifestyle illnesses in the population and points to the necessity for long-term health interventions as well as continued monitoring. By contrast, the projection for schistosomiasis and TB exhibits gradual but unchanging trends with incremental rises anticipated during the forecast period. These trends could be subject to localized outbreaks, environmental factors, or rural region access to healthcare. As a whole, the forecasts reflect both ongoing and developing issues in the nation's public health scenario.

The findings can assist health authorities, policymakers, and healthcare providers in designing targeted programs, allocating resources, and improving disease prevention efforts. Equipping the healthcare system with the tools and knowledge to address both chronic and infectious diseases will be essential in promoting long-term public health and reducing the national disease burden.

Table 1. Annual Reported Cases of Selected Morbidities in a Municipality in Lanao del Norte, Philippines

Recorded Cases of Five Major Diseases	**2020**	**2021**	**2022**	**2023**	**2024**
Heart Disease Cases	1000	1200	1300	1400	1500
Hypertension Cases	2500	2700	3000	3200	3400
Diabetes Deaths	120	88	76	66	58

Schistosomiasis Cases	17258	18125	16375	11258	12546
TB Cases	150	165	173	200	189

The raw data in Table 1 presents the recorded cases of five major diseases in a municipality in Lanao del Norte from 2020 to 2024. These figures offer valuable insight into the recent health landscape, showing both persistent and shifting disease burdens. However, while the data is informative about past and current conditions, it does not reveal what may happen in the years ahead.

To anticipate future morbidity trends, this study utilizes the ARIMA model to forecast disease incidence through the year 2030. Forecasting goes beyond statistics—it serves as a vital tool for guiding decisions. Whether it's local health officials planning resource allocation, policymakers developing intervention programs, or public health workers preparing for upcoming challenges, understanding where disease patterns are heading is essential.

The following sections provide an in-depth analysis of the projected trends for each of the five diseases, interpreting the predicted data and discussing their implications for public health strategy and planning.

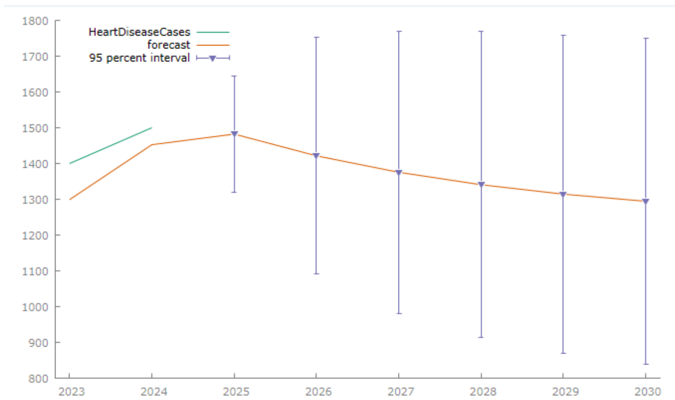

Figure 1: Forecasted Morbidity Trend – Heart Disease Cases (2025–2030)

The forecasted data for heart disease cases indicates a peak of around 1,500 cases in 2024, followed by a gradual decline from 2025 onward. By 2030, the number of cases is projected to stabilize at approximately 1,350. The 95% confidence intervals are quite broad, which indicates a moderate degree of uncertainty in future projections. The pattern indicates a potential deceleration in new heart disease occurrences, but the persistently high rates still emphasize the necessity of continued public health interventions, early detection initiatives, and community-level education on cardiovascular disease prevention and control.

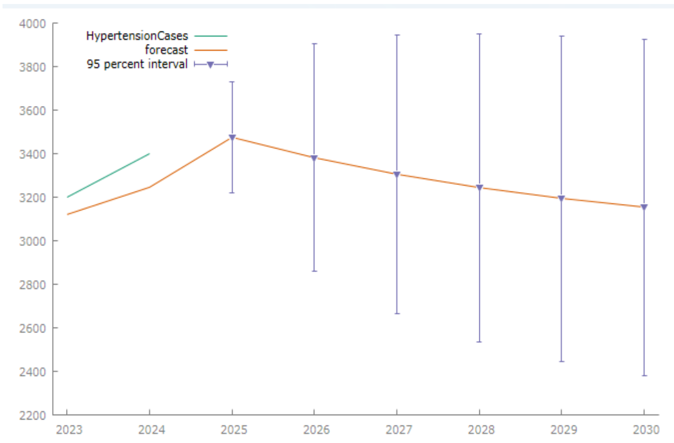

Figure 2: Forecasted Morbidity Trend – Hypertension Cases (2025–2030).

International Journal for Multidisciplinary Research (IJFMR)
E-ISSN: 2582-2160 ● Website: www.ijfmr.com ● Email: editor@ijfmr.com

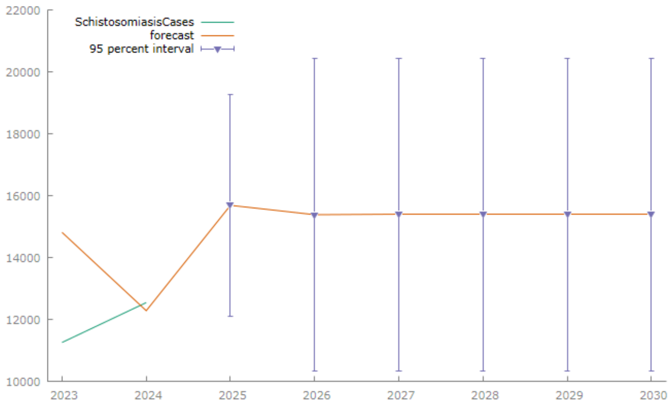

Figure 3: Forecasted Morbidity Trend – Diabetes-Related Deaths (2025–2030)

The projection is that cases of hypertension will peak at about 3,475 in 2025 before slowly declining to about 3,154 in 2030. While the trend is decreasing, the 95% confidence intervals increase immensely over time from 3,222–3,728 in 2025 to 2,382–3,927 in 2030. This increasing uncertainty implies that external forces like changes in lifestyle, health policies, or intervention programs could have a large impact on future results. Even with the minor drop, the consistently high incidence underscores the ongoing need for preventive medical care, public education, and affordable management regimens for hypertension.

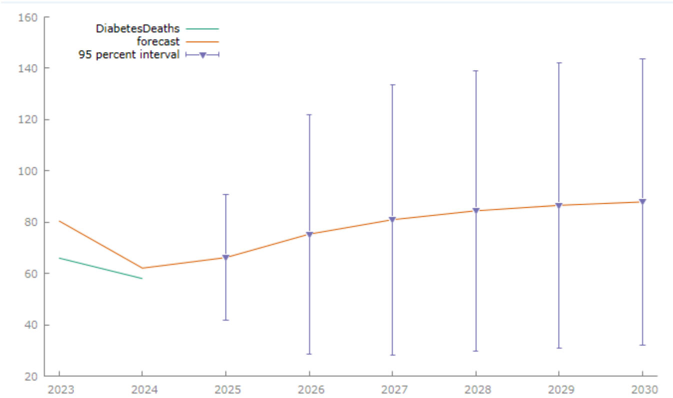

Figure 4: Forecasted Morbidity Trend – Schistosomiasis Cases (2025–2030)

The projection is an increase in diabetes-attributable deaths from 2025 to 2030. The trend is set to increase from about 66 deaths in 2025 to almost 88 in 2030. Though the rise is slow, the 95% confidence intervals show considerable variation, with 42–91 in 2025 and 32–144 in 2030. This increasing burden is an indication of the long-term effect of untreated diabetes and the need to strengthen early diagnosis, lifestyle changes, and availability of the right treatment. The consistent increase also indicates a possible stress on healthcare systems unless prevention methods are given top priority.

The projection is for a consistent trend in schistosomiasis cases, with estimates ranging approximately at 15,387 per year between 2026 and 2030. In 2025, the peak estimate at about 15,686 cases is approximately during a very wide 95% confidence interval of 12,103 to 19,270. Between 2026 and beyond, the projection becomes stabilized but with persistently wide confidence intervals of about 10,352 to 20,423, pointing to substantial uncertainty. This trend might indicate ongoing environmental and sanitation conditions that are driving disease transmission. The findings place emphasis on continued public health activities in vector management, hygiene improvement, and mass drug administration to control and minimize schistosomiasis occurrence in endemic regions.

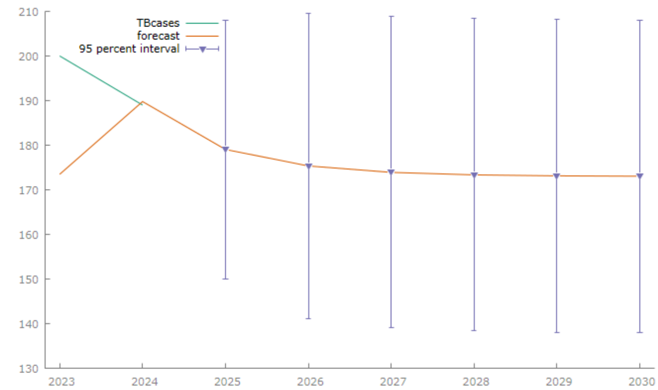

Figure 5: Forecasted Morbidity Trend – Tuberculosis (TB) Cases (2025–2030)

CONCLUSION AND RECCOMENDATION

Morbidity trends forecasting and analysis of priority diseases in Lanao del Norte municipality provides important insights for projecting future public health requirements. Employing the ARIMA model, in this research, it was estimated that the incidence of heart disease, hypertension, diabetes-related mortality, schistosomiasis, and tuberculosis would continue from 2025 through 2030.

The predictions identify different patterns—some diseases, for instance, heart disease and diabetes, indicate a gradual increase, while others, for example, tuberculosis, indicate a steady drop. Schistosomiasis is steadily high, pointing to ongoing environmental influences. These findings can help healthcare planners, local government units, and policymakers make evidence-based decisions regarding the allocation of resources, health education, and disease prevention interventions.

Forecasting aids proactive instead of reactive management of healthcare. Future studies are suggested to improve the accuracy of prediction using hybrid models or including socio-economic and environmental factors. Increased spending in data gathering, monitoring, and interventions at the community level will be necessary in disease control and overall public health in the region.

The projection is for a slow decrease in tuberculosis (TB) cases from 179 in 2025 down to around 173 by 2030. Although the trend is small, the fact that the decline is continuous implies that continued control and treatment measures might be producing positive outcomes. The 95% confidence intervals are fairly close, from 150–208 in 2025 to 138–208 by 2030, and reflect high confidence in the projection. Even though the number declined, data points to the necessity of continued active surveillance, early detection, and adherence to treatment in order to avert a resurgence of TB in the next few

Years

REFERENCES

1. F. J. B. Talirongan, H. Talirongan, and M. Y. Orong, "Modeling National Trends on Health in the Philippines Using ARIMA," vol. 11, 2020, doi: 10.37421/jhmi.2020.11.342.
2. Department of Health. (2021). Schistosomiasis control program. https://doh.gov.ph
3. Department of Health. (2023). Top causes of mortality in the Philippines. https://doh.gov.ph/statistics
4. Gupta, R., Gaur, K., & S Ram, C. V. (2019). Emerging trends in hypertension epidemiology in India. Journal of Human Hypertension, 33(8), 575–587. https://doi.org/10.1038/s41371-019-0200-9
5. Teixeira, M. G., Costa, M. da C. N., Barreto, F., & Barreto, M. L. (2018). Dengue: Twenty-five years since reemergence in Brazil. Cadernos de Saúde Pública, 34(5), e00082117. https://doi.org/10.1590/0102-311X00082117
6. World Health Organization. (2020). World malaria report 2020. https://www.who.int/publications/i/item/9789240015791
7. World Health Organization. (2022). Global tuberculosis report 2022. https://www.who.int/teams/global-tuberculosis-programme/tb-reports
8. Rufino, E. D. (2019). Forecasting monthly tourist arrivals in the Philippines using SARIMA models. Asia Pacific Journal of Tourism Research, 24(7), 678–692. https://doi.org/10.1080/10941665.2019.1613112
9. Sangco, J. D., & Vienes, M. B. (2020). Comparative study of forecasting techniques for migration data: Evidence from the Philippines. International Journal of Forecasting, 36(4), 1122–1135. https://doi.org/10.1016/j.ijforecast.2020.02.004
10. Bautista, R. P., & Rivera, C. J. (2020). Forecasting hospital admissions in Metro Manila using

ARIMA models. Philippine Journal of Health Systems, 10(2), 88–95.
11. De Castro, M. L., Santos, A. J., & Javier, R. G. (2021). Using ARIMA to predict the incidence of dengue in Metro Manila. Asian Journal of Public Health, 13(3), 55–64.
12. Angco, M. M., Reyes, J. S., & Capistrano, A. (2021). Forecasting economic participation in the Philippines using ARIMA models. Journal of Economic Forecasting, 24(3), 102–117.
13. Parreño, M. A. (2018). A comparative analysis of ARIMA and Holt-Winters models in forecasting agricultural production in the Philippines. Philippine Journal of Agricultural and Biosystems Engineering, 16(2), 23–31.
14. Estoque, R. C., Murayama, Y., & Myint, S. W. (2019). Effects of landscape patterns on urban growth and sustainability: Evidence from selected rapidly urbanizing cities in Asia. Science of the Total Environment, 654, 45–62. https://doi.org/10.1016/j.scitotenv.2018.10.442