



Epidemiological Prognosis of Pertussis Incidence in Bulgaria

Ralitsa Raycheva¹, Yordanka Stoilova², Ani Kevorkyan², Vanya Rangelova²

¹ Department of Social Medicine and Public Health, Faculty of Public Health, Medical University of Plovdiv, Bulgaria

² Department of Epidemiology and Disaster Medicine, Faculty of Public Health, Medical University of Plovdiv, Plovdiv, Bulgaria

Corresponding author: Vanya Rangelova, Department of Epidemiology and Disaster Medicine, Faculty of Public Health, Medical University of Plovdiv, 15A Vassil Aprilov Blvd., 4000 Plovdiv, Bulgaria; E-mail: vaniaran1238@gmail.com; Tel.: +359 883 403 683

Received: 2 Jan 2020 ♦ **Accepted:** 13 Mar 2020 ♦ **Published:** 30 Sep 2020

Citation: Raycheva R, Stoilova Y, Kevorkyan A, Rangelova V. Epidemiological prognosis of pertussis incidence in Bulgaria. *Folia Med (Plovdiv)* 2020;62(3):509-14. doi: 10.3897/folmed.62.e49812.

Abstract

Introduction: Epidemiological forecasting facilitates scientifically sound solutions to upcoming theoretical and practical issues, in the development of public health management, in particular of infectious diseases.

Aim: To critically analyze the most recent scientific advances in the biosocial nature and methodology of epidemiological forecasting to present a real-life example of pertussis, a disease with shifting epidemiology.

Materials and methods: For the prediction of pertussis morbidity the autoregressive integrated moving average (ARIMA) the model was established by utilizing the method of time series analysis to construct a model of overall morbidity using Time series modeller in SPSS v.25. To model pertussis morbidity we obtained official data from the Ministry of Health and the National Center for Infectious and Parasitic Diseases, since the beginning of disease registration from 1903 until 2018. We also analyzed the shifting epidemiology of pertussis.

Results: The proper identification procedures we applied indicated ARIMA (3,0,0) model to best fit our original time series of the annual whooping cough morbidity for the 1921-2018 period. The model predicts better morbidity in a one-step forecast. The incidence rate is expected to be stable at about 1.35 per 100,000 in the next three years, which is close to the 2016 level and lower than those in 2017-2018.

Conclusion: The ARIMA (3,0,0) model in our study is an adequate tool for presenting the pertussis morbidity trend and is suitable to forecast near-future disease dynamics, with acceptable error tolerance.

Keywords

ARIMA model, pertussis (whooping cough), short-term epidemiological forecasting

INTRODUCTION

In the 1940s, one of the founders of epidemiology of infectious diseases, LV Gromashevsky, said: "The scientific thought has always sought not only to accumulate facts but also to systematize them and to explain them in a single theoretical system".¹ An epidemiological prognosis provides an opportunity for scientifically sound solutions to the-

oretical and practical issues in the development of science and public health management, in particular of infectious diseases at future stages. The first attempts for a discussion in Bulgaria date from the late 1960s² and the early 1970s. In 1976, the book "Epidemiological prognosis", based on the biosocial nature of the epidemic process of infectious diseases, was published by Emil Kamenov, professor of epidemiology.³

Science and practice in Bulgaria are in debt to the basic methods of epidemiological prognosis and mathematical modeling of infectious diseases. The other descriptive, analytical and experimental methods are broadly in use. At present, the molecular epidemiology and a variety of mathematical techniques are reported in the literature dealing with epidemiologic prognosis to be in the focus of interest of scientists.⁴⁻⁶ Prognostic studies allow better understanding of the natural history of an epidemic process, facilitate the selection of appropriate optimal solutions for good medical practice, as well as more accurately predict the outcome of the disease.⁷ The evaluation of the results of an appropriate statistical analysis allows us to draw conclusions that can influence the epidemic process.

AIM

The aim of the present study was to analyze the most recent scientific advances in the biosocial nature and methodology of epidemiological prognosis in order to present a real-life example of pertussis, a disease with shifting epidemiology.

MATERIALS AND METHODS

We obtained official data from the Ministry of Health and the National Center for Infectious and Parasitic Diseases, Sofia, for pertussis morbidity since the beginning of its registration in 1903 until 2018.

In the article we analyzed the shifting epidemiology of pertussis by means of a chronological overview. For prediction of pertussis morbidity, the autoregressive integrated moving average (ARIMA) model was established by utilizing the method of time series analysis to construct a time series model of overall morbidity using Time series modeller in SPSS v.25. The ARIMA modeling was based on the yearly morbidity pertussis rates (per 100,000 population) from 1903 to 2018 (the whole period) and from 1921 to 2018 (due to missing data from 1908 to 1920) in Bulgaria. The decomposition of pertussis morbidity time series into a stochastic trend and a stationary component was reasoned based on the model defined. In addition, we assessed the importance of the short-term and immediate effect of one shock.⁴

RESULTS

We selected the Auto-Regressive Integrated Moving Average (ARIMA), a forecast model pointed out by many researchers as having high accuracy for short-term forecasting. It was proposed by Box and Jenkins in the early 1970s. It includes three components: 'AR' – autoregression, 'I' – integration, 'MA' – moving average. ARIMA is an epidemiological method for predicting future values based on previously observed values and is one of the most effective forecasting for time series, i.e. time series have been increasingly used in recent years for epidemiological research for decision-making purposes. Extracting meaningful statistics and data characteristics is tantamount to dynamic time series analysis.

We built four models (Table 1). On the one hand, we examined the extent to which missing data from the period 1908-1920 influenced the model in terms of its best approximation to the observed data (Model 1 – ARIMA (2,1,1) vs. Model 2 – ARIMA (3,0,0)), and on the other, we tested the accuracy of prediction by a single-step prediction within the sample (Model 1A – ARIMA (2,1,2) and Model 2A – ARIMA (3,0,0)). In addition, we performed three years extended out-of-sample short-term forecast (Model 1 – ARIMA (2,1,2) and Model 2 – ARIMA (3,0,0)) (Fig. 1).

Both models 1 and 2 have a similar coefficient of determination (R^2) of 64% and therefore both have a good approximation to the observed values (Table 1). The one-step

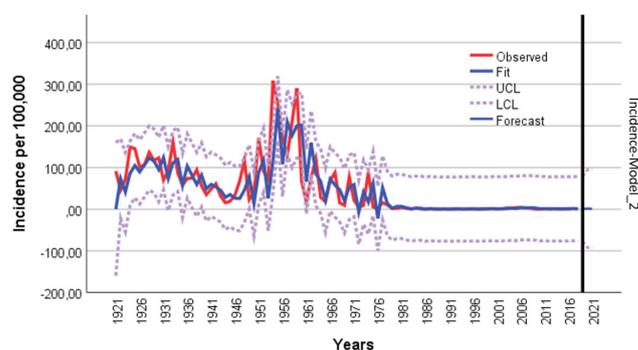


Figure 1. Time series and ARIMA model (3,0,0) 3 years ahead forecast.

Table 1. Comparison of the ARIMA models, one-step and short-term forecast for pertussis morbidity

Model	Time period	ARIMA	R2	Ljung-Box Q		Forecast			
				Q	Sig.	2018	2019	2020	2021
1	1903-2018	2,1,2	0.641	16.50	0.350		0.77	1.10	1.12
1A	1903-2017	2,1,2	0.639	16.42	0.355	0.79	-	-	-
2	1921-2018	3,0,0	0.645	24,21	0.062		1.35	1.31	1.38
2A	1921-2017	3,0,0	0.643	23.99	0.065	0.93	-	-	-

forecast for Model 1A differs from the real data for 2018 (1.62 per 100 000) by 0.83 per 100,000, while in Model 2A the difference is 0.69 per 100 000.

Model 1 approximates a period of 115 years, with linear time series ($d=1$) and eliminated by subtraction from the data. There is a correlation between the observations at dY and the correlation coefficient is ϕ_2 ($p=2$). There is also a relationship between the current value and the random deviation at lag 2 ($q=2$) (i.e. two steps back in the row), with the correlation coefficient θ_3 representing the magnitude of this relationship. Model 1 represents the approximation of a more complex time series having simultaneously three components in the ARIMA model, i.e. a mixed model – with a trend: relationship between each current and several previous row values; and dependence also between the deviations with different lags. Due to data gap Model 2 approximates a period of 97 years with the time series being stationary ($d=0$) and missing the moving average component ($q=0$). The number of autoregressive components (p) is 3, therefore there is a correlation between the observations in $d3Y$ and the correlation coefficient is ϕ_3 . Model 2A is closer to Model 2 in forecast than Model 1A to 1.

The benefits of these models make them optimally suited for pertussis epidemiological modelling, a disease with shifting epidemiology. Additionally, for the period 2012-2018 we have analyzed the shifting epidemiology of pertussis by focusing on the age distribution of cases. The registered cases of pertussis are 606 (average 86 cases per year) and they vary between 35 (2015) and 116 (2017). Children under 4 years of age are with the highest proportion – 60.4%, followed by children between 5-9 years – 13.86% (Fig. 2). Almost 15% of all pertussis cases are in children above 10 years of age (10-14 yrs. – 10.07% and 15-19 yrs. – 4.95%).

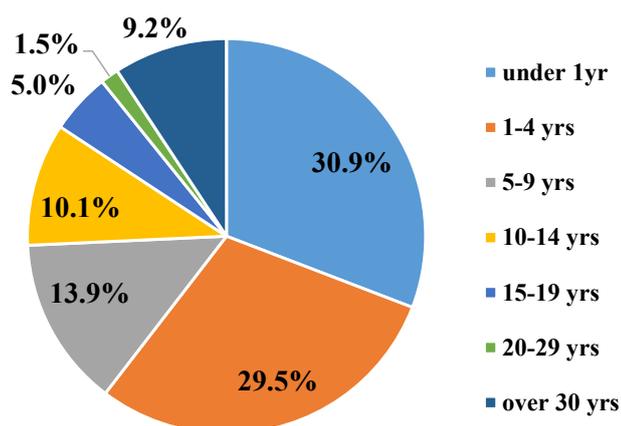


Figure 2. Distribution of the pertussis cases ($n=606$) by age groups for the period 2012-2018.

DISCUSSION

The biosocial nature of infectious diseases is based on a branch of epidemiology – molecular epidemiology (the

fusion of molecular biology with the epidemic process to further characterize the epidemiological discipline). Molecular techniques make the study more comprehensible and indicate: reliably to the highest degree the causes of the occurrence and spread of infectious diseases; unknown aspects of infectious diseases; solutions to more effective prevention and control strategies at an individual and population levels; ideas for future directions for research in this field; necessity to train competent professionals.⁸

Looking back, we have found that in Bulgaria the epidemiological prognosis is based on the long-standing and currently complex biosocial nature of the epidemic process.

Unfortunately, in our country, for a long period, the mathematical and statistical modelling methods are insufficiently developed and slightly used, even underestimated. In the 1970's they are defined as vehemence.³ At the same time Box and Jenkins propose Auto-Regressive Integrated Moving Average.⁵ Models such as Exponential Smoothing (ETS Model / Error, Trend, Seasonality), Auto-Regressive and Moving Average Model (ARMA Model), Neural Network and some others have been the subject of intense methodological developments in recent years.⁹

Looking ahead the following categories are defined according to the literature: epidemiological prognosis model; the modeling process and its logical sequence; the available and evolving science-based and refined accurate prognostic models such as ARIMA, etc. In this regard, prognostic studies require the selection of variables or risk factors that assess their impact on the outcome of the epidemic process. Many variables have a confounding effect on other parameters and this should be minimized. The conditions under which the prognostic study is conducted should correlate with the local practices of the institutions in which the finding might apply. The evaluation of the prognostic studies is based on determining the internal validity of the design and assessing the impact of systemic errors or biases.⁶

The ARIMA (p, d, q) models are applied when the time series are stationary, i.e. characterizes the change of a random process in which there is no dynamics (there is no trend of change). Most epidemiological time series are non-stationary and before the application of ARIMA models, they need to be transformed into stationary ones.

The proper identification procedures, we applied, indicated ARIMA (3,0,0) model to best fit our original time series of the annual whooping cough morbidity for the 1921-2018 period. In trend-stationary models, characterized by given common properties, the time series returns to the trend, in particular, at stationarity – to the average.⁴ As noted Model 2A better predicts morbidity in one-step forecast. The incidence rate is expected to be stable at about 1.35 per 100,000 in the next three years, which is close to 2016 levels and lower than in 2017-2018.

The accurate simulation and prognosis of the cases of infectious diseases indicates the resources and the methods for planning of future epidemics. However, in some diseases, the cases cannot be easily identified or managed. Such an example is whooping cough, which is similar to other

respiratory diseases which have a mild or atypical course. In addition, there is a shift in morbidity, which is affected and limited by several reasons. First is the shifting trends in morbidity – in countries with mass vaccination the morbidity has decreased by 100 times since 1953^{10,11} and in Bulgaria the reduction is by 150 times.¹² Second is the periodic changes (the incidence has increased in the last decade in the EU). Globally pertussis is endemic, it never completely disappears, making it an important public health problem.¹³⁻¹⁵ For an example, with 86% global immunization coverage in 2016, 131,991 have been infected and 89,000 have died, 95% of them in developing countries.¹⁴ Third is the accidental disorders due to lack of or inaccurate diagnosis, lack of awareness, etc. Some effective and accurate models also need to be identified for the prediction of the disease based on historical data.¹²

Analysis of the shifting epidemiology of whooping cough has revealed some important facts:

1. Despite the high vaccination coverage among young children worldwide, many countries have reported an increase in the cases of whooping cough in adolescents and adults.^{16,17} Our study identified 91 (15.02%) of a total of 606 cases in children over 10 years of age which is in favour of a booster dose of the vaccine, for an example at 12 years.

2. The age distribution of the laboratory-confirmed 4341 cases of whooping cough in England for 2017 reveals: 81% of the cases are in children over 14 years of age, 10% – 0-9 years of age, 9% – 10-14 years.^{18,19} For 2017 our results show a much lower proportion of those over 14 years old – 17.23%, but extremely high proportions – 74.14% in children 0-9 years and 8.62% for 10-14 years old.

3. Natural infection and vaccination do not provide life-long protection against whooping cough. The change during the years in the estimates of the duration of protection after natural infection is as follows: for life in 1951, for 20 years in 1995, for 7-10 years in 1997 and for 3.5-12 years in 2002. The duration of protection after vaccination with whole-cell vaccines is estimated to be 4-12 years, and with acellular vaccines 5-6 years.^{14,20,21}

CONCLUSION

The available articles and reports, studying pertussis infection in Bulgaria are limited and in their majority are descriptive, with clinical focus on the severity of illness and complications. To our knowledge the presented ARIMA model is the first in the country to outline the dynamics of the epidemic process over a period of hundred and fifteen years. The computational model accurately analyzes the epidemiological picture of the increasing impact of pertussis on public health over the coming years. The ARIMA (3,0,0) model in our study is an adequate tool for presenting the pertussis morbidity trend and is suitable to forecast near future disease dynamics, with acceptable error tolerance. The application of short-term forecasting models is aiding decision-making for rational allocation of health re-

sources and favours successful planning of prevention and control of infectious disease.

A key role for the reduction of health and economic losses according to the status of the disease in Bulgaria and according to foreign researchers can have:

1. Improvement of the available diagnostic tests – culture, PCR, serology.
2. Awareness of healthcare professionals in order to report cases.
3. Significance of whole-cell and acellular vaccine-theory and practice.
4. Emergence of new pathogenic pertussis strains with prevalence in vaccinated patients.
5. Net population status.
6. Booster dose at age 12.
7. Improving vaccine efficacy and duration of protection.

Acknowledgments

This paper was supported by Ministry of Education and Science, Republic of Bulgaria (National Program ‘Young scientists and postdoctoral students’).

REFERENCES

1. Gromashevsky LV. [General Epidemiology.] 4th edition. Moscow; 1965 [In Russian].
2. Bratovanov D. Epidemiological prognosis. Methods and basic principles of epidemiologic prognosis. *Folia Med (Plovdiv)* 1969;11(3): 170–83.
3. Kamenov E. [Epidemiologic prognosis.] Sofia: Medicina y fizikultura; 1976 [In Bulgarian]
4. Hamilton JD. *Time Series Analysis*. Princeton NJ; 1994.
5. Box G, Jenkins GM, Reinsel GC. *Time Series Analysis: Forecasting and Control*. 3rd ed. Prentice Hall; 1994.
6. Mak K, Kum CK. How to appraise a prognostic study. *World J Surg* 2005; 29(5): 567–9.
7. Giesecke J. *Modern Infectious Disease Epidemiology*. 3rd edition. London: Edward Arnold; 1994.
8. Eybpoosh S, Haghdoost AA, Mostafavi E, et al. Molecular epidemiology of infectious diseases. *Electronic physician* 2017;9(8): 5149–58.
9. Wang Y, Xu C, Wang Z, et al. Time series modeling of pertussis incidence in China from 2004 to 2018 with a novel wavelet based SARI-MA-NAR hybrid model. *PLoS One* 2018;13(12):e0208404.
10. Gonfiantini MV, Carloni E, Gesualdo F, et al. Epidemiology of pertussis in Italy: disease trends over the last century. *Euro Surveill* 2014;19(40):20921.
11. Chen Z, Zhang J, Cao L, et al. Seroprevalence of pertussis among adults in China where whole cell vaccines have been used for 50 years. *J Infect* 2016;73(1):38–44.
12. Stoilova Y. [Study on the dynamics of the epidemic processes as a result of the prophylaxis of infectious diseases in Bulgaria for the period 1977-2017]. [Post-doctoral dissertation.] Plovdiv: Medical University; 2009 [In Bulgarian].

13. Clarke MF, Rasiah K, Copland J, et al. The pertussis epidemic: informing strategies for prevention of severe disease. *Epidemiol Infect* 2013;141(3):463–71.
14. World Health Organization. Pertussis. [document on the internet]. Geneva, Switzerland; 2019. Available from: https://www.who.int/immunization/monitoring_surveillance/burden/vpd/surveillance_type/passive/pertussis/en/
15. Zeng Q, Li D, Huang G, et al. Time series analysis of temporal trends in the pertussis incidence in Mainland China from 2005 to 2016. *Scientific reports* 2016;6:2367.
16. Wang K, Fry NK, Campbell H. Whooping cough in school age children presenting with persistent cough in UK primary care after introduction of the preschool pertussis booster vaccination: prospective cohort study. *BMJ* 2014;24;348:g3668.
17. Torres RS, Santos TZ, Torres RA, et al. Resurgence of pertussis at the age of vaccination: clinical, epidemiological and molecular aspects. *J Pediatr (Rio J)* 2015; 91:333–8.
18. Public Health England (PHE). Laboratory confirmed cases of pertussis reported to the enhanced pertussis surveillance programme in England: annual report for 2017. Available from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/703519/hpr1518_prtsss_ANN.pdf
19. Childhood vaccination coverage statistics, England 2016-17, National statistics. Available from: <https://files.digital.nhs.uk/pdf/d/3/nhs-immu-stat-eng-2016-17-rep.pdf>
20. World Health Organization (WHO). Pertussis vaccines: WHO position paper. *Weekly Epidemiol Rec* 2015;90:433–60.
21. Wendelboe AM, Van Rie A, Salmaso S, et al. Duration of immunity against pertussis after natural infection or vaccination. *Pediatr Infect Dis J* 200; 24(5 Suppl): S58–61.

Эпидемиологический прогноз заболеваемости коклюшем в Болгарии

Ралица Райчева¹, Йорданка Стоилова², Ани Кеворкян², Ваня Рангелова²

¹ Кафедра „Социальная медицина и общественное здравоохранение“, Факультет общественного здравоохранения, Медицинский университет – Пловдив, Пловдив, Болгария

² Кафедра „Эпидемиология и медицина чрезвычайных ситуаций“, Факультет общественного здравоохранения, Медицинский университет – Пловдив, Пловдив, Болгария

Адрес для корреспонденции: Ваня Рангелова, Кафедра „Эпидемиология и медицина чрезвычайных ситуаций“, Факультет общественного здравоохранения, Медицинский университет – Пловдив, бул. „Васил Априлов“ № 15А, 4002 Пловдив, Болгария; E-mail: vaniaran1238@gmail.com; Tel.: +359 883 403 683

Дата получения: 2 января 2020 ♦ **Дата приемки:** 13 марта 2020 ♦ **Дата публикации:** 30 сентября 2020

Образец цитирования: Raycheva R, Stoilova Y, Kevorkyan A, Rangelova V. Epidemiological prognosis of pertussis incidence in Bulgaria. *Folia Med (Plovdiv)* 2020;62(3):509-14. doi: 10.3897/folmed.62.e49812.

Резюме

Введение: Эпидемиологический прогноз поддерживает научно обоснованные решения предстоящих теоретических и практических проблем в развитии менеджмента общественного здравоохранения и, в частности, инфекционных заболеваний.

Цель: Критически проанализировать последние научные достижения биологического характера и методологию эпидемиологических прогнозов, чтобы представить реальные примеры коклюша, заболевания с переменной эпидемиологией.

Материалы и методы: Для прогнозирования заболеваемости коклюшем была разработана интегрированная модель авторегрессии скользящего среднего (ARIMA) с использованием метода анализа временных рядов для построения общей модели заболеваемости с использованием модели временных рядов в SPSS v.25. Для построения модели коклюша мы использовали официальные данные Министерства здравоохранения и Национального центра по инфекционным и паразитарным болезням с начала регистрации заболевания с 1903 по 2018 год. Мы также проанализировали изменение эпидемиологии коклюша.

Результаты: Правильные процедуры идентификации, которые мы использовали, показали, что модель ARIMA (3,0,0) наилучшим образом соответствует нашим исходным временным моделям ежегодной заболеваемости коклюшем за период 1921–2018 гг. Модель прогнозирует лучшую заболеваемость с помощью одношагового прогноза. Ожидается, что заболеваемость будет стабильной на уровне 1,35 на 100 000 в течение следующих трёх лет, что близко к уровню 2016 года и ниже, чем в 2017–2018 годах.

Заключение: Модель ARIMA (3,0,0) из нашего исследования является адекватным инструментом для представления тенденции развития коклюша и подходит для прогнозирования динамики заболевания в ближайшем будущем с приемлемой допустимой погрешностью.

Ключевые слова

ARIMA модель, коклюш, краткосрочный эпидемиологический прогноз
