



## Evaluating Measles Incidence Rates Using Machine Learning and Time Series Methods in the Center of Iran, 1997-2020

Javad Nazari<sup>1</sup>, Parnia-Sadat Fathi<sup>2</sup>, Nahid Sharahi<sup>3</sup>, Majid Taheri<sup>4</sup>, \*Payam Amini<sup>5</sup>, \*Amir Almasi-Hashiani<sup>6,7</sup>

1. Department of Pediatric, School of Medicine, Arak University of Medical Sciences, Arak, Iran
2. Student Research Committee, School of Medicine, Arak University of Medical Sciences, Arak, Iran
3. Health Department, Arak University of Medical Sciences, Arak, Iran
4. Medical Ethics and Law Research Center, Shahid Beheshti University of Medical Sciences, Tehran, Iran
5. Department of Biostatistics, School of Health, Iran University of Medical Sciences, Tehran, Iran
6. Department of Epidemiology, School of Health, Arak University of Medical Sciences, Arak, Iran
7. Traditional and Complementary Medicine Research Center (TCMRC), Arak University of Medical Sciences, Arak, Iran

\*Corresponding Authors: Email: amiralmasi2007@gmail.com, payam.amini87@gmail.com

(Received 12 Feb 2021; accepted 13 Apr 2021)

### Abstract

**Background:** Measles is a feverish condition labeled among the most infectious viral illnesses in the globe. Despite the presence of a secure, accessible, affordable and efficient vaccine, measles continues to be a worldwide concern.

**Methods:** This epidemiologic study used machine learning and time series methods to assess factors that placed people at a higher risk of measles. The study contained the measles incidence in Markazi Province, the center of Iran, from Apr 1997 to Feb 2020. In addition to machine learning, zero-inflated negative binomial regression for time series was utilized to assess development of measles over time.

**Results:** The incidence of measles was 14.5% over the recent 24 years and a constant trend of almost zero cases were observed from 2002 to 2020. The order of independent variable importance were recent years, age, vaccination, rhinorrhea, male sex, contact with measles patients, cough, conjunctivitis, ethnic, and fever. Only 7 new cases were forecasted for the next two years. Bagging and random forest were the most accurate classification methods.

**Conclusion:** Even if the numbers of new cases were almost zero during recent years, age and contact were responsible for non-occurrence of measles. October and May are prone to have new cases for 2021 and 2022.

**Keywords:** Measles; Machine learning; Time series; Infection

## Introduction

Measles is among the most infectious disorders of humans that may cause severe illness and adverse symptoms (1). This disease is caused by measles virus and includes several symptoms such as fever (may be as high as 105 °F [40.5 °C]), malaise,

cough, coryza (nasal mucous membrane inflammation), conjunctivitis, Koplik spots (enanthem, or a rash on the mucous membranes), and maculopapular rash (exanthema, or a skin rash) (1, 2).



Copyright © 2022 Nazari et al. Published by Tehran University of Medical Sciences.

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International license.

(<https://creativecommons.org/licenses/by-nc/4.0/>). Non-commercial uses of the work are permitted, provided the original work is properly cited

Measles can quickly spread by sick people's coughs and sneezes and can even transmit by close interaction with mouth or nasal secretions (2). A significantly higher basic reproduction number for measles in comparison to other spreading viruses such as Influenza is reported (3). Measles was an instance of the relationship between demographic factors and population patterns of the epidemics. Case-fatality levels tend to be significant in tropical areas such as Asia and sub-Saharan Africa and grow to 20%-30% among disadvantaged groups like refugees (3).

Greater regulation in vaccines has decreased the rate of this infection in recent years worldwide (4). However, the United Nations (UN) has warned countries about the increase in the amount of measles reports worldwide by 48.4% in 2019 which is also growing steadily due to inadequate monitoring of the vaccinations, development in anti-vaccination campaigns, economic and political issues surrounding health-care programs (3). Compared to 2016, the frequency of measles increased by 167% in 2018, where America and Africa had the most and least rate respectively (5). The estimated annual measles occurrence rate in Iran was small, however in 2016, following a rise in the frequency of positive measles reports in 2015, the incidence rate of measles experienced a significant decrease in the next year (6). As the Eastern Mediterranean region witnessed the largest increase in measles cases in 2019, Iran was granted a measles exclusion status in October 2019 (7). Although the number of new cases has been diminished significantly in Iran during recent years, the central parts such as Markazi Province have the highest risk point in Iranian districts (8).

Removal of measles is a public health concern since travel-related infections will be inevitable as long as the virus continues to circulate in any part of the globe. Therefore, the goal of this study was to classify factors that placed people at a higher risk of measles using different classification methods and assess their impact on the series of measles monthly incidence frequencies using time series approaches.

## Methods

### *Data*

This case study was conducted on the measles incidence in Markazi Province, the center of Iran, from Apr 1997 to Feb 2020. The data were extracted from the database of the Vice-chancellor of Health Services, Arak University of Medical Sciences, Markazi, Iran.

The data contained the information about individuals' measles test results (positive/negative) based on measles-specific IgM antibody, gender (male/female), age (year), location (urban/rural), any contact with measles patients (yes/no), ethnic (Iranian/ non-Iranian), and some clinical signal such as rhinorrhea (yes/no), fever (yes/no), conjunctivitis (yes/no), cough (yes/no), and history of vaccination (yes/no).

The result of measles test was considered as the binary response variable and independent variables were utilized to classify the cases into two levels of outcome using classification approaches. Moreover, regarding the nature of monthly measles new cases over the study period and the excess zeros in the series, Zero-Inflated Negative Binomial (ZINB) regression for time series was utilized.

### *Statistical Analysis*

Logistic Regression (LR) links the probability of measles to the predictors using a logit function. Odds ratio is used to report the effect of each variable (9). Linear Discriminant Analysis (LDA) refers the dependent variable to linear predictors and addresses the problem by the conditional likelihood of the factors given the outcome class (10). Random Forest (RF) combines the regression tree and classification in which powerful and quick computations are achieved over large datasets (11). Artificial Neural Network (ANN) involves input, output, and secret layers, where there are multiple nodes in each row. By adding a degree of nonlinearity, an activation mechanism converts the data within each layer into the next one (12). In Bagging, the number of bootstrap samples is selected from the training set and the noisy obser-

vations are reduced by bootstrapping and even removed (13). In Naïve Bayes (NB), the prior likelihood of contributing to each category of the outcome is conditioned on the predictor variables. Subjects will be assigned to the category with the highest posterior probability (14). Support Vector Machine (SVM) locates a hyperplane in a P-dimensional space (the number of predictors) that separately classifies the binary outcome. The aim is to seek a plane with the maximum range, that is to say, the maximum gap between the dependent variable categories (15).

### Comparing the methods

Sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and overall accuracy were given to assess the discriminative quality of the computational models. We split the dataset into two training sets (70%) and testing sets (30%). We then validated the methods 500 times and listed the assessment criterion as the average of the iterations.

### Zero-Inflated Negative Binomial (ZINB) regression for time series

This model assumes a negative binomial distribution for the observation with excess zero and fits a two-part model. The first part assesses the impact of predictors on the counts observations such as a usual negative binomial regression and outputs the estimate of coefficients where a logarithm link function is used. The second part of the model uses a logit link function to evaluate the impact of

independent variables on non-occurrence of the outcome (16-18).

The independent variables were set as the proportion of a certain level to the sample space of the variables to assess their effect on the series of measles incidences over time. Moreover, for ZINB time series model, age is categorized into six levels to make the convergence of maximizing the likelihood function possible and to ease the interpretation (19).

### Software

All analyses were implemented in R-version 3.6.3 using the randomForest, 1071, rpart, ZIM, and rminer packages.

### Ethics approval

This study was approved by the Ethics Committee of Arak University of Medical Sciences. In this study, we used the existing registered data from the Health deputy of Arak University of Medical Sciences.

## Results

### Data description

The data included the information of 2114 cases in which 306 (14.5%) experienced measles from 1997 to 2020. The majority of reports with positive and negative measles were in 1999 and 2003 respectively and a constant trend of almost zero cases was observed from 2002 to 2020 (Fig. 1).

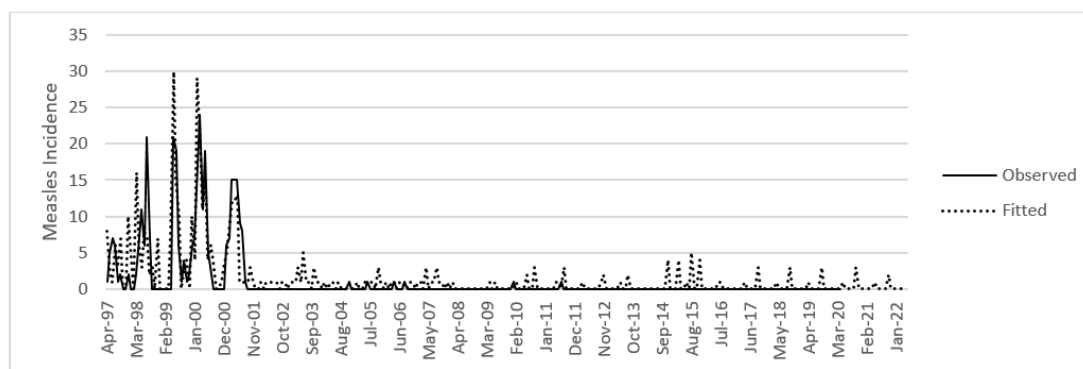


Fig. 1: The trend of Measles from 1997 to 2020 and forecasting the next two years in Markazi province, Iran

Characteristics of patients as well as their unadjusted association with the outcome are shown in Table 1. The mean (standard deviation) of cases

with and without Measles was 15.71 (7.18) and 12.55 (8.49) yr respectively.

**Table 1:** The distribution of variables between the categories of measles and the results of Chi-square and independent samples *t*-test test

Variable	Measles Mean ( $\pm$ SD) Or n (%)		P-value
	No	Yes	
Time (yr)	1382.6 ( $\pm$ 4.7)	1378.34 ( $\pm$ 1.6)	<0.001
Age	12.55 ( $\pm$ 8.49)	15.71 ( $\pm$ 7.18)	<0.001
Sex			0.129
Female	787 (86.9)	119 (13.1)	
Male	1021 (84.5)	187 (15.5)	
Ethnic			<0.001
Non-Iranian	75 (72.8)	28 (27.2)	
Iranian	1733 (86.2)	278 (13.8)	
Location			0.005
Rural	888 (83.4)	177 (16.6)	
Urban	920 (87.7)	129 (12.3)	
Vaccination			<0.001
No	679 (78.9)	182 (21.1)	
Yes	1129 (90.1)	124 (9.9)	
Contact			<0.001
No	1064 (88.2)	143 (11.8)	
Yes	744 (82.0)	163 (18.0)	
Fever			<0.001
No	661 (92.3)	55 (7.7)	
Yes	1147 (82.0)	251 (18.0)	
Cough			<0.001
No	913 (90.2)	99 (9.8)	
Yes	895 (81.2)	207 (18.8)	
Rhinorrhea			<0.001
No	902 (92.8)	70 (7.2)	
Yes	906 (79.3)	236 (20.7)	
Conjunctivitis			<0.001
No	942 (92.5)	76 (7.5)	
Yes	866 (79.0)	230 (21.0)	

**Performance of the models in predicting Measles**

Regarding the results in Table 2, except for SVM, LR, and NB, the other four approaches demonstrated sensitivities higher than 0.50. Bagging and LR approaches sowed the highest specificity and

PPV among the methods respectively. In contrast to other models, NB exposed the lowest NPV of 78%. Average on both train and test datasets, the Bagging and RF methods resulted in higher accuracies and were introduced as the best classifiers among different approaches.

**Table 2:** The result of different classification methods classifying measles using independent variables followed by 500 repetition of cross validation

<i>Method</i>	<i>Set</i>	<i>Sensitivity</i>	<i>Specificity</i>	<i>Positive predictive value</i>	<i>Negative predictive value</i>	<i>Total accuracy</i>
Bagging	Train	0.94 ± 0.01	0.98 ± 0.01	0.87 ± 0.01	0.99 ± 0.004	0.97 ± 0.01
	Test	0.57 ± 0.05	0.91 ± 0.01	0.50 ± 0.05	0.94 ± 0.01	0.87 ± 0.01
RF	Train	0.86 ± 0.02	0.94 ± 0.01	0.65 ± 0.02	0.98 ± 0.01	0.94 ± 0.01
	Test	0.70 ± 0.06	0.94 ± 0.01	0.50 ± 0.04	0.90 ± 0.01	0.90 ± 0.01
LR	Train	0.35 ± 0.03	0.97 ± 0.01	0.86 ± 0.02	0.84 ± 0.01	0.83 ± 0.01
	Test	0.37 ± 0.08	0.97 ± 0.02	0.93 ± 0.08	0.83 ± 0.02	0.82 ± 0.02
LDA	Train	0.65 ± 0.04	0.87 ± 0.01	0.16 ± 0.02	0.98 ± 0.01	0.86 ± 0.01
	Test	0.60 ± 0.12	0.87 ± 0.01	0.15 ± 0.02	0.98 ± 0.01	0.86 ± 0.01
ANN	Train	0.88 ± 0.02	0.66 ± 0.02	0.21 ± 0.05	0.98 ± 0.06	0.87 ± 0.01
	Test	0.86 ± 0.03	0.63 ± 0.02	0.17 ± 0.03	0.98 ± 0.05	0.85 ± 0.03
Naïve bayes	Train	0.34 ± 0.01	0.94 ± 0.01	0.68 ± 0.02	0.78 ± 0.01	0.76 ± 0.01
	Test	0.34 ± 0.03	0.94 ± 0.01	0.68 ± 0.04	0.78 ± 0.01	0.76 ± 0.01
SVM	Train	0.50 ± 0.02	0.54 ± 0.12	0.15 ± 0.04	0.86 ± 0.03	0.55 ± 0.12
	Test	0.48 ± 0.10	0.53 ± 0.10	0.15 ± 0.02	0.86 ± 0.02	0.52 ± 0.09

RF: Random Forest; LSSVM: Least-squares support-vector machine; LDA: Linear Discernment Analysis; NB: Naive Bayes; LR: Logistic Regression; ANN: Artificial Neural Network; SVM: support-vector machine

### *The association of Measles and independent variables*

Regarding the results shown in Tables 3 and 4, the classification approaches found almost the same

results. Recent years were associated with less number of new cases and time was the most significant variable predicting measles in Markazi Province.

**Table 3:** The importance of independent variables resulted by perfumed methods

<i>Independent Variable</i>	<i>Im- portance</i>	<i>ANN</i> Normalized Im- portance(%)	<i>Bagging</i> Im- portance	<i>SVM</i> Im- portance	<i>NB</i> Quality esti- mate of the attribute (%)	<i>LDA</i> Standard- ized Coeffi- cients	<i>RF</i> Mean de- crease ac- curacy
Rhinorrhea		11.5	43.42	0.05	0.8	0.08	9.69
Age		34	139.95	0.22	0.9	0.05	17.87
Fever		7.6	18.76	0.05	-0.1	0.71	13.11
Ethnic (Iran)		13.6	21.45	0.22	-0.2	-0.09	7.08
Gender (Male)		6.3	35.11	0.02	0.5	0.10	1.24
Conjunctivitis		7	39.10	0.02	-0.7	-0.10	10.58
Contact		7.8	38.25	0.03	-0.6	-0.09	7.02
Cough		9.5	36.49	0.03	-0.6	-0.44	11.04
Urban		12.1	36.48	0.04	-0.6	-0.13	4.89
Vaccine		13.1	47.12	0.13	-0.8	-0.31	20.08
Time (Year)		100	169.08	0.19	-1.9	-0.99	70.40

RF: Random Forest; LDA: Linear Discernment Analysis; CM: Core Model; ANN: Artificial Neural Network; BMI: Body Mass Index

**Table 4:** Logistic regression analyses for relationship between demographic/clinical factors and measles

Variable	Prevalence of Measles, Mean ( $\pm$ SD) Or n (%)	Unadjusted		Adjusted	
		OR (95% CI)	P-value	OR (95% CI)	P-value
Time (Year)	1999.34 ( $\pm$ 1.6)	0.97 (0.96-0.98)	0.021	0.97 (0.96-0.98)	0.023
Age (Year)	15.70 ( $\pm$ 7.18)	1.02 (1.01-1.03)	0.015	1.00 (0.99-0.1.01)	0.063
Sex			0.134		0.251
Female	119 (13.1%)	1		1	
Male	187 (15.5%)	1.02 (0.99–1.05)		1.01 (0.98-1.04)	
Ethnic			<0.001		<0.001
Non-Iranian	28 (27.2%)	1		1	
Iranian	278 (13.8%)	0.87 (0.82–0.93)		0.91 (0.85-0.97)	
Location			0.005		0.009
Urban	129 (12.3%)	1		1	
Rural	177 (16.6%)	1.04 (1.01–1.07)		1.03 (1.01–1.06)	
Vaccination			<0.001		<0.001
No	182 (21.1%)	1		1	
Yes	124 (9.9%)	0.89 (0.86–0.92)		0.92 (0.89-0.95)	
Contact			<0.001		0.127
No	143 (11.8%)	1		1	
Yes	163 (18.0%)	1.06 (1.03–1.09)		1.01 (0.99–1.03)	
Fever			<0.001		<0.001
No	55 (7.7%)	1		1	
Yes	251 (18.0%)	1.11 (1.07–1.14)		1.21 (1.14–1.28)	
Cough			<0.001		<0.001
No	99 (9.8%)	1		1	
Yes	207 (18.8%)	1.09 (1.06–1.12)		1.11 (1.06–1.16)	
Rhinorrhea			<0.001		0.083
No	70 (7.2%)	1		1	
Yes	236 (20.7%)	1.14 (1.11–1.17)		1.04 (0.99–1.09)	
Conjunctivitis			<0.001		0.079
No	76 (7.5%)	1		1	
Yes	230 (21.0%)	1.14 (1.11–1.18)		1.04 (0.99–1.09)	

BMI: Body Mass Index; OR: Odds Ratio; CI: Confidence Interval

Age was the second most important factor associated with measles. The unadjusted odds ratio showed that one-year increase in age is associated with 0.02 (95% confidence interval (CI): 0.01-0.03) likelihood of measles. Vaccination and rhinorrhea were the third and fourth most important affecting variables so vaccinated cases were 0.08 (95% CI: 0.89-0.95) less prone to experience measles and those with rhinorrhea were 1.04 (95% CI: 0.99-1.09) less times in risk comparing to those without this sign. Adjusted for other variables, the methods revealed that male cases are more at risk than females so the odds of measles were almost 0.01 more among men. Moreover, any contact

with measles patients increased the odds by about 65% and 1% in unadjusted and adjusted perspectives respectively. The methods also showed that cough and conjunctivitis have same amount of importance for predicting measles. Ethnic and fever had less influence in comparison to other variables.

Based on the results in Table 5, age is an effective variable both on the frequency of measles and zero inflation part of the model. The frequency of measles patients increases by 2.41 individuals in average as one-level increase in age (95% CI: 1.02 – 5.77).



**Table 5:** The results of Zero-Inflated Negative Binomial time series regression assessing the impact of independent variables on the series of measles

Variable	ZINB Regression		Zero-inflated part	
	Exponential (Beta)	95% CI	OR	95% CI
Age (Categorical)	2.41	1.02 - 5.77	0.04	0.01 - 0.23
Sex (Male proportion)	0.97	0.19 - 4.88	1.45	0.08 - 24.88
Ethnic (Iranian proportion)	0.19	0.03 - 1.05	0.81	0.02 - 26.78
Location (Urban proportion)	0.35	0.08 - 1.55	0.31	0.02 - 5.05
Vaccination (Yes proportion)	2.48	0.60 - 10.30	1.05	0.10 - 11.12
Contact (Yes proportion)	9.50	6.42 - 14.06	0.10	0.07 - 0.16
Fever (Yes proportion)	7.14	4.83 - 10.57	0.08	0.06 - 0.13
Cough (Yes proportion)	0.09	0.01 - 9.33	0.02	0.01 - 1.08
Rhinorrhea (Yes proportion)	0.10	0.01 - 1.60	0.42	0.04 - 5.12
Conjunctivitis (Yes proportion)	6.47	4.37 - 9.57	1.22	0.12 - 12.60

95% CI: 95% Confidence Interval; ZINB: Zero-inflated negative binomial

## Discussion

Our data showed a significant change point in 2003 when the health policy-makers started one of the major immunization campaigns against measles and rubella in three phases including catch-up, keep-up, and follow-up. This operation yielded an impressive reduction in measles incidence rate (7). In other words, the large rate of immunization coverage in all Iranian cities and villages is among the most significant reasons for the reduced incidence of measles based on the WHO risk evaluation method (20). However, the WHO reported peaks in places with large total vaccination coverage over recent months, such as the United States of America, Thailand and Tunisia, as the infection has quickly spread among many groups of unvaccinated people (21). Since vaccination coverage in many countries is suboptimal, measles seems to be spread across countries. To reach the elimination target, several governments have to bring incremental progress in the scope of their regular childhood immunization systems and close immunity gaps between different age groups skipped out on vaccination opportunities (22).

The individuals between 9 and 19 yr old had the higher rate of measles in our observations. While measles is typically a childhood illness, infection occurs in individuals of any age. The age might be

confounded with vaccination so that unvaccinated, partly vaccinated, or weakened immunity cases of any age are in danger. Particularly, unvaccinated youths are at the greatest risk. Depending on the local immunization procedures, age-specific attack levels could be higher in vulnerable babies younger than 12 months, school-age children, or young adults (23).

We found out that men are more at risk compared to women. That might be due to lack of balanced job development between the sexes so that men are busier in social jobs which need more contact and communication. Southwest of Tehran, Markazi Province is flooded with Afghan immigrants, mostly men, where the plurality of cases come from rural locations and refugee communities with higher incidence areas (8). At the other side, long-term political instability, religious conflicts, violence, and inadequate healthcare services have impacted Iran's neighboring countries and the emergence of almost 6 percent of Iran's population has placed increased strain at Iran's health care framework (24). It has been reported that aspects like neighboring the capital, situated on the key road to the western regions of the country, a substantial majority of conventional service and manufacturing systems that are ideal for workers contribute to Markazi province being able to handle a significant number of non-Iranian citizens

(24). This fact is a potential and considerable factor for increasing the risk of any contact with infected cases (25) which is responsible for the occurrence of measles based on zero-inflated analysis in our data. The rise in the perceived vulnerability of higher-risk communities was largely attributed to inadequate standard of monitoring, focus due to weak execution of the strategy, and the inclusion of vulnerable population groups (8).

The results of our study were achieved using classification and time series methods. The manner under which the predictors influence the result is important for deciding the correct technique of classification, in addition to the distribution of the result groups. It's recommended to replicate the cross-validation phase to verify the findings. In our study, the cross-validation technique was carried out over hundreds of repetitions to estimate the measles outcome. Machine learning and stochastic process approaches are widely used in different aspects such as RF for minimizing false negatives of measles prediction model (26) and zero-inflated models to investigate the transmission of measles (27).

Several limitations of this historical cohort study should be noted. Many other potential predictors have not been recorded by the center and could have significant impact on the classification and forecasting procedures in our study.

Although forecasting analysis confirmed the continuing procedure of constant low number of new cases for the next two years in addition to this fact that the number of new measles cases had dramatically decreased over the recent decade in Iran which yielded this country to receive the elimination certificate, there are many European and Americas regions who have lost their certificate in 2018 and 2019 (7). This is an alarm for the health policy makers to continue the restrictions on measles infection such as preventing contact with suspicious cases and strict oversight of immigration systems from neighbor countries which are the most deterministic factors to continue the measles elimination strategy.

## Conclusion

Even if the numbers of new cases were almost zero in the recent years, age and contact were responsible for non-occurrence of measles. October and May are prone to have new cases for 2021 and 2022.

## Journalism Ethics considerations

Ethical issues (Including plagiarism, informed consent, misconduct, data fabrication and/or falsification, double publication and/or submission, redundancy, etc.) have been completely observed by the authors.

## Acknowledgements

We appreciate the scientific support of Vice-chancellor of Research, Arak University of Medical Sciences, Markazi, Iran.

## Funding

This study was funded by Vice-chancellor of Research, Arak University of Medical Sciences, Markazi, Iran. The funder has no role in data analysis, interpretation and manuscript drafting.

## Conflict of interest

The authors declare that there is no conflict of interest.

## References

1. World Health Organization (2012). Global measles and rubella strategic plan: 2012.
2. Bester JC (2016). Measles and measles vaccination: a review. *JAMA pediatrics*, 170(12):1209-1215.
3. Congera P, Maraolo AE, Parente S, et al (2019). Measles in pregnant women: a systematic review of clinical outcomes and a meta-analysis of antibodies seroprevalence. *J Infect*, 80(2):152-160.



4. Li S, Ma C, Hao L, Su Q, et al (2017). Demographic transition and the dynamics of measles in six provinces in China: A modeling study. *PLoS Med*, 14(4):e1002255.
5. Patel MK, Dumolard L, Nedelec Y, et al (2019). Progress Toward Regional Measles Elimination—Worldwide, 2000–2018. *MMWR Morb Mortal Wkly Rep*, 68(48):1105-1111.
6. Mohammadbeigi A, Zahraei SM, Sabouri A, et al (2019). The spatial analysis of annual measles incidence and transition threat assessment in Iran in 2016. *Med J Islam Repub Iran*, 33:130.
7. Namaki S, Gouya MM, Zahraei SM, et al (2020). The elimination of measles in Iran. *Lancet Glob Health*, 8(2):e173-e174.
8. Mohammadbeigi A, Zahraei SM, Asgarian A, et al (2018). Estimation of measles risk using the World Health Organization Measles Programmatic Risk Assessment Tool, Iran. *Helvion*, 4:e00886.
9. Agresti A, Kateri M (2011). *Categorical data analysis*. ed. Springer.
10. Izenman AJ (2013). Linear discriminant analysis. In: *Modern multivariate statistical techniques*. Ed(s): Springer, pp. 237-280.
11. Cutler A, Cutler DR, Stevens JR (2012). Random forests. In: *Ensemble machine learning*. Ed(s): Springer, pp. 157-175.
12. Maroufizadeh S, Amini P, Hosseini M, et al (2018). Determinants of Cesarean Section among Primiparas: A Comparison of Classification Methods. *Iran J Public Health*, 47(12): 1913–1922.
13. Xiao T, Zhu J, Liu T (2013). Bagging and boosting statistical machine translation systems. *Artificial Intelligence*, 195:496-527.
14. Tapak L, Shirmohammadi-Khorram N, et al (2019). Prediction of survival and metastasis in breast cancer patients using machine learning classifiers. *Clinical Epidemiology and Global Health*, 7:293-299.
15. Amini P, Ahmadiania H, Poorolajal J, et al (2016). Evaluating the High Risk Groups for Suicide: A Comparison of Logistic Regression, Support Vector Machine, Decision Tree and Artificial Neural Network. *Iran J Public Health*, 45(9): 1179–1187.
16. Yang M, Zamba GK, Cavanaugh JE (2013). Markov regression models for count time series with excess zeros: A partial likelihood approach. *Statistical Methodology*, 14:26-38.
17. Yang M, Cavanaugh JE, Zamba GK (2015). State-space models for count time series with excess zeros. *Statistical Modelling*, 15:70-90.
18. Yang M, Zamba G, Cavanaugh J (2014). ZIM: Zero-inflated models for count time series with excess zeros. *R package version*, 1.
19. Xiong Y, Wang D, Lin W, et al (2014). Age-related changes in serological susceptibility patterns to measles: results from a seroepidemiological study in Dongguan, China. *Hum Vaccin Immunother*, 10(4):1097-03.
20. Zahraei SM, Eshtrati B, Gouya MM, et al (2014). Is there still an immunity gap in high-level national immunization coverage, Iran? *Arch Iran Med*, 17(10):698-701.
21. World Health Organization (2019). New measles surveillance data for 2019. Retrieved August, 24:2019.
22. Prevention ECfD, Control (2018) Monthly measles and rubella monitoring report, September 2018. (ed)^(eds), ECDC Stockholm,
23. Hughes SL, Bolotin S, Khan S, et al (2020). The effect of time since measles vaccination and age at first dose on measles vaccine effectiveness—A systematic review. *Vaccine*, 38(3):460-469.
24. Soleimanpour S, Hamedei Asl D, Tadayon K, et al (2014). Extensive genetic diversity among clinical isolates of Mycobacterium tuberculosis in central province of Iran. *Tuberc Res Treat*, 2014:195287.
25. Velayati AA, Farnia P, Mirsaedi M, et al (2006). The most prevalent Mycobacterium tuberculosis superfamilies among Iranian and Afghan TB cases. *Scand J Infect Dis*, 38(6-7):463-8.
26. Ahmad WMTW, Ab Ghani NL, Drus SM (2019). Minimizing False Negatives of Measles Prediction Model: An Experimentation of Feature Selection Based On Domain Knowledge and Random Forest Classifier. *International Journal of Engineering and Advanced Technology*, 9:3411-3414.
27. Mizumoto K, Kobayashi T, Chowell G (2018). Transmission potential of modified measles during an outbreak, Japan, March–May 2018. *Eurosurveillance*, 23(24): 1800239.