

## Predictive evaluation of pediatric patients based on their typhoid fever status using linear discriminant model

James A. Ndako<sup>a,\*</sup>, Joseph A. Olisa<sup>b</sup>, Ilochi C. Ifeanyichukwu<sup>c</sup>, Charles E. Okolie<sup>a</sup>, Stephen K.S. Ojo<sup>d</sup>, Segun L. Jegede<sup>e</sup>

<sup>a</sup> Department of Microbiology, Landmark University Omu-Aran, Nigeria

<sup>b</sup> Department of Medical Services, Landmark University Medical Center, Omu-Aran Nigeria

<sup>c</sup> Department of Medical Laboratory Services, Landmark University Medical Center, Omu-Aran Nigeria

<sup>d</sup> Department of Microbiology, Federal University Oye-Ekiti, Nigeria

<sup>e</sup> Department of Research, Statistics and documentation (VCO), Landmark University Omuaran, Nigeria

### ARTICLE INFO

#### Keywords:

Typhoid fever  
Prevalence rate  
Pediatric patients  
Linear discriminant  
Hematological parameters

### ABSTRACT

Epidemiologic studies have established a relationship between pediatric patients and typhoid fever infection. This study was carried out to ascertain if specific hematological measurements of the pediatric patients discriminate between their positive and negative status to typhoid infection and to produce a rule for classifying other pediatric patients. Discriminant analysis was applied to predict the probability of a specific categorical outcome based on several explanatory variables (predictors). This study analyzed the differentiation between two hundred pediatric patients attending Landmark University Medical Centre based on their typhoid fever status. The hematological parameters considered were Packed Cell Volume, White Blood Cell count; Neutrophil, Erythrocyte level, Hemoglobin and Platelet count, Assay of samples were performed using standard procedures. Fisher's Linear Discriminant Method was used for classification of variables in this study. With the use of the Fisher's Linear Discrimination method for classification of the obtained data, a minimum value of  $-0.0067$  was obtained implying that any new pediatric patient with a discriminant score above  $-0.0067$  would be diagnosed to be typhoid negative; otherwise, they would be classified as typhoid positive pediatric patients. The efficiency of this method of classification was tested using two approaches; Retribution estimate approach and leaving-one out approach which showed a prevalence rate of typhoid positive patients at 75.8% and 74.7% respectively. This data analysis hypothesizes that typhoid fever is highly endemic amongst our study subjects. A point-of-care diagnosis with a strong positive predictive value, which improves pediatric enteric fever diagnosis, is strongly advocated.

### Background

Discriminant analysis focuses on associating multiple independent variables with a categorical dependent variable through the formation of a composite of independent variables. This type of multivariate analysis can determine the extent to which any of the composite variables discriminates between two or more pre-existing groups of subjects and can also derive a classification rule for predicting the group membership of new observations [1].

*Salmonella enterica* serovars: Paratyphi A, B and C is the organism responsible for typhoid and paratyphoid fever respectively which is a major bacterial infection globally [2]. Typhoid fever affects over 30 million individuals and causes approximately 500,000 deaths annually [3]. Typhoid and paratyphoid fever can be distinguished based on their

clinical features. In various parts of the world, an increasing prevalence rate has been recorded and this is predominantly observed in Asia, South-eastern, Africa and South and Central America [4]. The majority of existing epidemiological evidence comes from studies in adult populations. Even though young children in endemic areas suffer a significant burden of illness [5]. There are conflicting views on young children's illness rates – particularly infants [6,7].

The reason for this increased incidence rate is not specifically known but this is due to poor sanitary measures and personal hygiene in most areas [8]. Typhoid fever can be transmitted by ingesting food or water which has been contaminated with the causative organism (*Salmonella typhi*) [9]. This organism is able to remain latent in the tissues of the infected host for a long time as it has the ability to evade responses from the immune system at eliminating the organism [10].

\* Corresponding author.

E-mail address: [ndako.james@lmu.edu.ng](mailto:ndako.james@lmu.edu.ng) (J.A. Ndako).

<https://doi.org/10.1016/j.mehy.2020.110264>

Received 16 May 2020; Received in revised form 16 July 2020; Accepted 6 September 2020

Available online 22 September 2020

0306-9877/ © 2020 Elsevier Ltd. All rights reserved.

Untreated and unsafe sources of water or food are easy means of harboring typhoid fever [11]. However, treatment of typhoid fever is a greater source of concern as the causative organism has become resistant to various antibiotics such as co-trimoxazole, ampicillin and chloramphenicol and recently, nalidixic acid due to constant usage of these antibiotics [12]. Prioritization of vaccination against typhoid fever can be a greatly efficient means of reducing the prevalence rate of this infection.

Improved sanitary measures, treatment of water before drinking can also aid in curbing the spread of this infection [13]. Based on several literatures, a significant relationship has been observed between pediatric patients and their typhoid fever status based on hematological parameters such as Platelet count, Monocyte and Lymphocyte level, Hemoglobin concentration and several others [14,15].

Several literatures have established a relationship between pediatric patients and typhoid fever infection and this study further analyzed the differentiation between two hundred pediatric patients based on their typhoid fever status by employing six hematological parameters for analysis and observable hematological changes.

**Materials and method**

*Study area*

This research was carried out at Landmark University Omuaran is a city located in North Central Nigeria, with an Average relative humidity (%) (at 15:00 LST). Omuaran is situated in Kwara–South Senatorial zone of Nigeria.

*Study design and study population*

The study was conducted at the Medical Laboratory department of the Landmark University Medical Center Omuaran –Nigeria, which is a tertiary health institution of learning. The data set used in this research was obtained from the study population which comprised of randomly selected volunteer subjects attending the outpatient department of the Medical facility. The sample size used in this research work was obtained from Two–Hundred (200) volunteers of different age groups and gender, the age range of the study subjects were infants and children aged ≤ 5 years.

*Ethical permit and consent*

A proposal of the project was submitted to the Ethical Review Board of the Landmark University Medical Center. Where ethical approval with referenced number LMC/2019/02/29 was obtained. This Research was conducted following standards for reporting Diagnostic accuracy Guidelines and in Compliance with ethical standards.

*Informed consent*

Informed consent was obtained from all individual Participants or their next of kin/guardians as appropriate. Well-structured questionnaires were used to collect demographic data and other pertinent information, Response to these questionnaires was done through the assistance of the respective volunteer parents/guardian of our study subjects.

*Sample collection/analysis*

Blood samples were collected from the subjects, the serum sample obtained was used for Widal test while samples of whole blood (plasma) were used to assay for full blood count and the determination of the various hematological parameters. A confirmatory test was further carried out by using the Kirby-Bauer disk diffusion method for antibiotic sensitivity test to confirm for typhoid fever infection according to

the method of Keddy et al, [16]. The hematological parameters considered were Packed Cell Volume, White Blood Cell count, Neutrophil and Erythrocyte level, Hemoglobin and Platelets.

*The hypothesis*

Typhoid fever carries a significant public health importance as Children are worst affected. Clinical diagnosis is usually unreliable and diagnostic tests are often not available in disease-endemic countries. Consequently, the incidence and relative importance of etiologic agents of the febrile illness remain unknown in many parts of the world.

*Evaluating the hypothesis using Fisher’s linear discriminant model*

Fisher’s Linear Discriminant model was used for classification in this study. This is a means of classifying multivariate data which is concerned with deriving rules and procedures for assigning individuals to one of a set of a prior specific group in some optimal fashion based on a set of measurements,  $x_1, x_2, \dots, x_m$ , taken on each individual. Fisher suggested the variables are transformed linearly as,  $z = a_1x_1 + a_2x_2 + \dots + a_mx_m$ . This is done such that the separation between the group means on the transformed scale on the n-groups, say  $n = 2$ , on the transformed scale  $\bar{z}_1$  and  $\bar{z}_2$  would be maximized relative to the within group variation on the scale. The transformation coefficients are estimated as:

$$a^T = S^{-1}(\bar{x}_1 - \bar{x}_2)$$

where  $a^T = [a_1, a_2, \dots, a_m]$ ,  $\bar{x}_1$  and  $\bar{x}_2$  are the group mean vectors and S is the pooled within group covariance matrix which is estimated from

$$z_c = \frac{\bar{z}_1 + \bar{z}_2}{2}$$

The classification is such that, an individual is assigned to group 1 if  $\bar{z}_1 - z_c > 0$  and group 2 otherwise, given that  $\bar{z}_1$  is the larger of the two means.

**Result and discussion**

The data for the analysis is described using statistics in Table 1. From Table 2, the Box’s test for equality of the covariance matrix rejects the null hypothesis at p-value < 0.05. That is, the covariance matrices of each status are not the same. To avoid much complication, the linear discriminant analysis may still be preferable. This conclusion is reached by assuming normality as discussed by Hands [17] through reliance on the robustness of fisher approach to deal with any minor departure from assumption. The eigenvalue (0.453) gives the ratio between the group to the within sum of squares of discriminant scores of the two categories of patients having either a positive or negative typhoid status. This criterion is maximized in the discriminant function analysis.

The canonical correlation (cc), known as the correlation between the group membership and the discriminant function scores is 0.5585

**Table 1**  
Descriptive Statistics of the data used for classification.

Status	Statistics	PCV	WBC	NEUT	ERY	HB	PLT
Positive	Mean	37.242	6.148	63.566	216.657	12.295	11.354
	SD	5.008	2.062	16.399	80.643	1.664	5.939
Negative	Mean	40.653	6.726	62.257	255.069	13.498	5.822
	SD	3.872	2.907	15.446	50.016	1.317	7.804
Total	Mean	38.965	6.440	62.905	236.055	12.903	8.560
	SD	4.776	2.534	15.899	69.501	1.612	7.461

SD: Standard Deviation, PCV: Packed Cell Volume; WBC: White Blood Cell; NEUT: Neutrophil; LYMP: Lymphocytes; MONO: Monocytes; ESR: Erythrocyte Sedimentation Rate; HB: Haemoglobin concentration; PLT: Platelet count.

**Table 2**  
Statistics Check for Good Classification.

Box's M	Eigenvalue	Wilks' Lambda
value = 62.2897	Value = 0.453	value = 0.6881
Approximate = 2.8705	% of Variance = 100	Chi-square = 72.9005
df1 = 21	Cumulative % = 100	df = 6
df2 = 144069.45	Canonical	
pvalue = $1.1616 \times 10^{-5}$	Correlation = 0.5585	pvalue = $1.0377 \times 10^{-13}$

**Table 3**  
Classification Function Coefficients.

Status	PCV	WBC	NEUT	ERY	HB	PLT	Constant
Positive	2.787	0.161	0.264	0.078	-2.258	0.364	-58.065
Negative	2.552	0.240	0.251	0.088	-0.996	0.253	-66.474

PCV: Packed Cell Volume; WBC: White Blood Cell; NEUT: Neutrophil; LYMP: Lymphocytes; MONO: Monocytes; ESR: Erythrocyte Sedimentation Rate; HB: Haemoglobin concentration; PLT: Platelet count. Function at Status Centroid for the positive status = -0.677; Function at Status Centroid for the negative status = 0.663.

which can be considered to be average and positive. ( $cc^n \times 100$ ), that is, the status differences can only explain 31.19% of the variance in the discriminant function scores. This can be observed by implication from the lambda value which presents the proportion of the total variance which is not explained by differences among the groups in the discriminant scores as 68.81%. The Chi-square test confirms that the hematological properties differ between the two pediatric patients categories (p-value < 0.05). It is important to note that failure to accept the hypothesis of equal mean vector would imply a little or no point in carrying out a linear discriminant analysis (Table 2).

The Fisher's linear discriminant function for typhoid classification from hematological properties was obtained by subtracting the coefficients given for each property under each status (Table 3) and it is presented as:

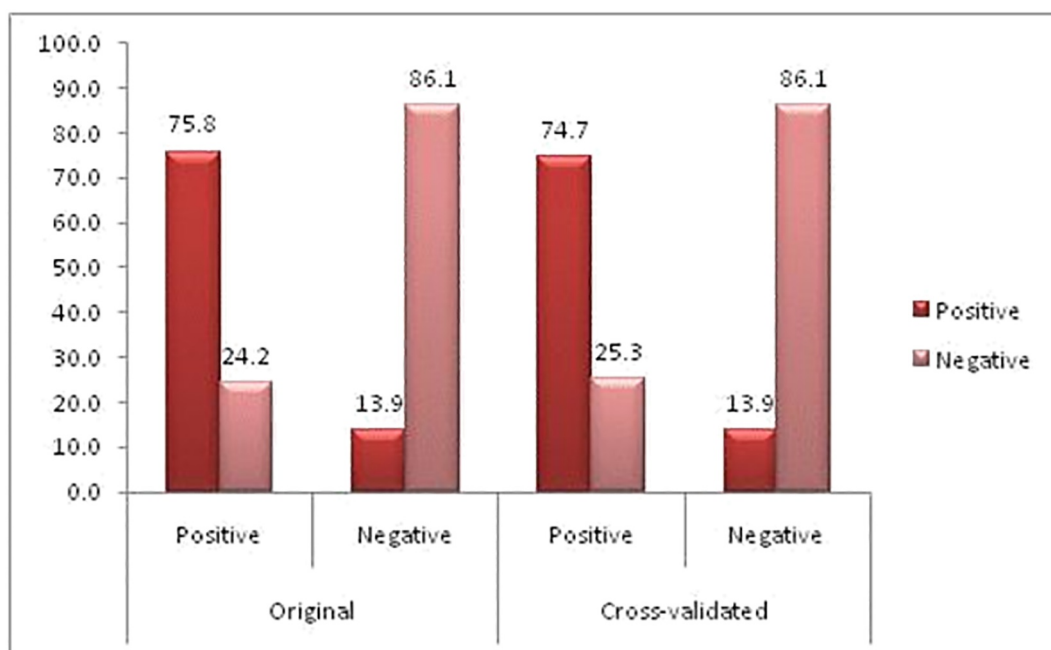
$$z = 0.235PCV - 0.080WBC + 0.012NEUT - 0.011ERY - 1.261HB + 0.110PLT$$

The sample mean of the discriminant function scores was obtained as  $\bar{z} = 8.409$  by taking the difference between the constant coefficients. These discriminant scores for each property is centered such that their sample mean is zero. To allocate the pediatric patient into either the positive or the negative typhoid group, the discriminant scores can be compared with the arithmetic mean of their status means. In this case, -0.0067 is the threshold against which a pediatric patient discriminant score is evaluated; this value was obtained from the function at status centroid for both categories of pediatric patients as  $\frac{1}{2} \times (-0.677 + 0.663)$ . Thus, any new pediatric patient with discriminant scores above -0.0067 would be diagnosed to be typhoid negative; otherwise, they would be classified as typhoid positive pediatric patients.

Two approaches were used to confirm how well the classification model performs. The first is by calculating the retribution estimate, the result of this method can be observed from the bars corresponding to "original" in Fig. 1. The retribution estimate approach shows that 75.8% and 86.1% were correctly classified correctly while 24.2% and 13.9% were wrongly classified to be typhoid positive and negative respectively. It was generally observed that based on the retribution estimate method, 81.0% of the pediatric patients were correctly classified. However, estimating the misclassification rates using this approach has been stated by [18], 1997) to be overly optimistic. Thus, another method known as "leaving one out" will be considered for classification. The result for the leaving one out method is reported by the cross-validated bars in Fig. 1. 74.7% and 86.1% were correctly classified correctly while 25.3% and 13.9% were wrongly classified to be typhoid positive and negative respectively. By comparing both methods, the efficiency of the classification dropped by 0.5%, that is, from 81.0% to 80.5%.

*Consequences of the hypothesis*

Typhoid fever has a high profile as a life-threatening disease for which treatment is becoming increasingly difficult because of the emergence of strains resistant to multiple antibiotics [19–22]. Unfortunately, the most widely used diagnostic test, the Widal test, performs very poorly as a screening test for febrile children especially in areas of low prevalence likely to be typical of much of rural Africa.



**Fig. 1.** Classification Result Chart. Correct classification rate (a) based on original is 81.0% (b) based on cross-validated is 80.5%.

Hence, predicting rate of typhoid fever infection among pediatric patients is one of the important issues which many researchers should develop interest in biomedical research especially in diagnosis, so as to further guide the physicians during treatment. The approach in this study, presents an efficient hybrid model for diagnosis of typhoid fever using an algorithm to optimize classification by linear discriminant approach (LDA). This model enhance the performance of linear discriminant analysis.

In this research, the discriminant analytic method highlights the probability of the classification of the patients under their diseased conditions. Two classification methods were used to check the sensitivity of the system, for a negative and positive classification method as applied; (75.8% and 86.1%) and (74.7% and 86.1%), correct classification were observed for the retribution estimate method and leaving-one out method respectively. The necessity for this predictive evaluation was further corroborated by the work of Vollaard *et al.*, [23]. This is to further validate diagnostic results obtained in such studies.

## Conclusion

The paucity of information on the predictive evaluation in diagnosis of typhoid fever contributes to ill-informed measures on priority interventions, particularly in vaccine regimen. The analytic parameters assayed in this study imply that typhoid fever is highly endemic amongst individuals in our area of study. This further suggests the need for prompt awareness among the populace and administration of typhoid fever vaccines to allow for increased immunity against this infection among this set of individuals at our location of study.

## Acknowledgements

Authors deeply acknowledge the Management of Landmark University Medical Center LMU-Omu-Aran for granting the approval to use the Medical Laboratory Facility. The efforts of all the Medical Laboratory Scientist in the Lab.is deeply appreciated. Miss Oludolapo Olatinsu is highly appreciated for the lay out of this work and for assisting at the interpretation of the Data.

## Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.mehy.2020.110264>.

## References

- [1] Tabachnick BG, Fidell LS. Using multivariate statistics. New York, NY, USA: HarperCollins; 1996.
- [2] Crump JA, Luby SP, Mintz ED. The global burden of typhoid fever. *Bull World Health Organiz* 2004;82:246–53.
- [3] Wain J. Typhoid fever. *Lancet* 2015;385(9973):1136–45.
- [4] Ochiai RL. Salmonella paratyphi A rates, Asia. *Emerg Infect Dis* 2005;11:1744–66.
- [5] Azmatullah A, Qamar FN, Thaver D, Zaidi AK, Bhutta ZA. Systematic review of the global epidemiology, clinical and laboratory profile of enteric fever. *J Glob Health* 2015;5(020407).
- [6] Dewan AM, Corner R, Hashizume M, Ongee ET. Typhoid fever and its association with environmental factors in the Dhaka Metropolitan Area of Bangladesh: a spatial and time-series approach. *PLoS Negl Trop Dis* 2013;7:e1998.
- [7] Date KA, Bentsi-Enchill A, Marks F, Fox K. Typhoid fever vaccination strategies. *Vaccine* 2015;33:C55–61.
- [8] Crump J, Mintz E. Global trends in typhoid and paratyphoid Fever. *Clin Infect Dis* 2010;50(2):241–6.
- [9] Tarrad, AL-Khafaji, Jawad K, AL-Yasari HF, AL-Taei MH, 2006. Prevalence of typhoid fever among pediatric patients at AL-Musaib District. *Med J Babylon* 3(1–2).
- [10] Buckle GC, Walker CL, Black R. Typhoid fever and paratyphoid fever: systematic review to estimate global morbidity and mortality for 2010. *J Global Health* 2012;2:10401.
- [11] Ali HA, Ahmed MS, Gangwani Jawahar L, Nadeem YJ, Ustadi Abdulla M, Hina SH. Hematological and biochemical changes in typhoid fever. *Pakist J Med Sci* 25(2);2009:166–171.
- [12] Khan K. Recent trends in typhoid research—a review. *Int J Biosci* 2012;2:110–20.
- [13] Ananth KR, Mohd IK, Sridhar DY. Prevalence of typhoid fever in pediatric patients of Abilabad. *Saudi J Med* 2018;3(4):186–90.
- [14] Anusuya B, Sumathi S. Haematological alterations due to typhoid fever in mayiladuthurai area, Nagapattinam. *Int J Res Pharmacol Pharmacother* 2015;4(2).
- [15] Emenuga V, et al. Some haematological and biochemical profile of typhoid fever. *Indian J Appl Res* 2014;4(3):330–2.
- [16] Keddy KH, Sooka A, Letsoalo M. Sensitivity and specificity of typhoid fever rapid antibody tests for laboratory diagnosis at two sub-Saharan African sites. *Bull World Health Organiz* 2011;89:640–7.
- [17] Hand DJ. Discriminant analysis, linear. Armitage P, Colton T, editors. *Encyclopedia of biostatistics*, vol. 2. Chichester: Wiley; 1998.
- [18] Hand DJ. Construction and assessment of classification rules. Chichester 1997.
- [19] Feng YC. The epidemiology of typhoid fever in the Dong Thap Province, Mekong Delta region of Vietnam. *Am J Trop Med Hygiene* 2000;62:644–8.
- [20] Akinyemi K, Smith S, Oyefolu A, Coker A. Multidrug resistance in Salmonella enterica serovar typhi isolated from patients with typhoid fever complications in Lagos, Nigeria. *Public Health* 2005;119:321–7.
- [21] Bhan M, Bahl R, Bhatnagar S. Typhoid and paratyphoid fever. *Lancet* 2005;366:749–62.
- [22] Chowta M, Chowta N. Study of clinical profile and antibiotic response in typhoid fever. *Indian J Med Microbiol* 2005;23:125–7.
- [23] Vollaard A, Soegianto AM, van Asten H, et al. Risk factors for typhoid and paratyphoid fever in Jakarta, Indonesia. *J Am Med Assoc* 2004;291:2607–15.