

Joint Binary Response Modeling for Childhood Comorbidity in Ethiopia

Tesfaye Abera Bokoro (✉ tesfabera@gmail.com)

Haramaya University <https://orcid.org/0000-0002-8835-8079>

Habtamu Kiros Gebresilasie

Haramaya University

Research article

Keywords: Diarrhea, ARI, bivariate binary response, comorbidity, joint modelling

Posted Date: November 24th, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-113070/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Joint Binary Response Modeling for Childhood Comorbidity in Ethiopia

Tesfaye Abera Bokoro^{1*} *Habtamu Kiros Gebresilasie*^{1†}

¹*Department of Statistics, Haramaya University, Dire Dawa, Ethiopia*

* tesfabera@gmail.com

† haymanot.kiros20@gmail.com

Abstract

Background: Childhood diarrhea and Acute Respiratory Infection (ARI) are some of the diseases that share common risk factors in tropical developing regions. The objective of the study was to identify risk factors of childhood diarrhea and ARI among children age under-five years old based on 2016 Ethiopian Demographic and Health Survey data.

Methods: A joint binary response model that accommodates the interdependence between the two diseases was employed.

Results: We found a common odds ratio value (4.30) greater than unity, describing a positive association between the two childhood diseases. Thereby, employing a joint model to assess the potential factors for diarrhea and acute respiratory infection (ARI) was reasonable. Moreover, it was identified that standard errors of the parameter estimates in the joint response model were smaller compared to the corresponding standard errors of the separate models.

Conclusion: In the joint model, explanatory variables such as residence, vaccination, mother's education, and antenatal care visits during pregnancy were found statistically significant risk factors for diarrhea, whereas residence, number of children ever born, vaccination, mother's education, and wealth index were statistically significant risk factors for childhood Acute Respiratory Infection. The two correlated dichotomous response variables that is, diarrhea and ARI were affected together significantly by the risk factors such as residence, vaccination, and mother's education.

Key words: Diarrhea; ARI; bivariate binary response; comorbidity; joint modelling

1. Introduction

Morbidity in children living in low-income countries is commonly characterized by more than one health condition [11], and is a challenge in many settings that may lead to death. On the other hand, the problem of coinfection due to overlapping risk factors such as nutrition, sanitation, and overcrowding [9] is highly prevalent among children under five years seeking care [10]. Specifically, diarrhea and ARI (Acute Respiratory Infection) are some of the diseases that share common risk factors in tropical developing regions [22].

Diarrhea is one of the key contributors to deaths for under-five children in Ethiopia. According to, the World Health Organization (WHO) estimates, diarrhea contributes to more

than 13% (one in every ten) child deaths in Ethiopia [5]. The trend from previous studies depicts, percentage of children under age 5 who had diarrhea decreased from 24% in 2000 to 12% in 2016 [3]. Furthermore, acute respiratory infection (ARI), and especially pneumonia, is one of the leading causes of morbidity and mortality that accounts for 18% of deaths in Ethiopia [5]. The country has made remarkable commitment to reduce the childhood morbidity and mortality of ARI [4]. Deaths can be reduced if health interventions are targeted at either one or more simultaneous diseases [11]. Usually programs for child care addressed single diseases such as diarrhea, acute respiratory, fever, and malaria infections [10]. Diarrhea and acute respiratory infections remain the leading cause of death and often co-occur in children under the age of 5 years old [24]. The illness of the two health conditions may be as a result of shared risk factors at childhood, or as a result of shared external factors. For instance, diarrhea and acute respiratory diseases may both share risk factors such as age, as a child-dependent risk factor, or poor sanitation and crowding as the environmental risk factors [25].

For child care programs under health institutions to be successful, statistical tools are essential to determine the characteristics of diseases that coexist. In recent time, a joint analysis that examines risk factors of such diseases has become more popular in identifying similar patterns of variation [26]. The joint modeling addresses the correlation of the two responses variables with the same set of covariates. The two outcomes provide better control over the type I error rates in multiple tests and increases efficiency in the parameter estimates.

Most of previously conducted health researches in Ethiopia have emphasized on studying a single disease at a time and joint response modeling of correlated outcomes focused on continuous, discrete or a mixture of discrete and continuous outcomes [1].

However, due to common and overlapping risk factors, separate analyses may fail to provide a comprehensive picture of the epidemiology of the diseases and the combined effects of childhood diseases on the population under consideration [15].

Epidemiological methodology and research have increased over the years, and have extended from studying a single disease to several diseases jointly at a time. Thus, in this paper, we specifically aimed at comorbidity among under-five children using joint response model that accommodates the interdependence between the two infections in assessing their risk factors, so that the government of Ethiopia and interested institutions engaged in the health sectors are able to use the findings, for better design and implementation of intervention strategies that address exposure to multiple illnesses in the country.

2. Methodology

2.1 Data source

The empirical analysis in this paper is based on data available from the 2016 Ethiopian Demographic and Health Survey (EDHS). The Demographic and Health Surveys (DHS) are a well-established source of reliable population data with a substantial focus on childhood diseases [3]. There were about 41,392 children under the age of five years records in the 2016 survey of Ethiopia. Each record consists of information on childhood diseases and the list of covariates that could affect the health status of children. However, to reduce possible bias in the conclusion, totally 9917 children were considered by omitting those cases with the missing data in this study.

2.2 Description of Study variables

Let \mathbf{Y} be a vector of two dichotomous dependent variables, that is Y_1 (Diarrhea Status) and Y_2 (Acute Respiratory Infection Status). If the child had diarrhea and acute respiratory infection, the two binary response variables would assume each a value '1' or '0' otherwise. The possible joint outcomes of Y_1 and Y_2 for a set of m -paired observations with their probabilities of occurrence can be illustrated as in Table1 below.

Table1: Possible combination of outcomes with corresponding probabilities of occurrence

		Acute Respiratory Infection (Y_2)		Total
Diarrhea		$y_2=0$	$y_2=1$	
(Y_1)	$y_1=0$	p_{00}	p_{01}	$1-p_1$
	$y_1=1$	p_{10}	p_{11}	p_1
Total		$1-p_2$	p_2	1

Similarly, we can represent the joint and separate probabilities of Y_1 and Y_2 indicated in the table1 as, $p_{ij}=P(y_1=i, y_2=j)$; $i,j = 0,1$ and $p_k = P(y_k=1)$, $k = 1,2$ respectively.

Where:

P_{11} = the child has both diarrhea and ARI

P_{01} = the child has no diarrhea but has ARI

P_{10} = the child has diarrhea but not ARI

P_{00} = the child has neither diarrhea nor ARI

In the final models, the following predictors were included: child's age in months, child's sex, residence, antenatal care visit, vaccination, mother's education, number of children, and wealth index.

2.3. Statistical models

2.3.1. Separate and joint models

Bivariate logistic regression model is not identical application wise to the ordinary binary logistic regression model, since the primary objective in binary regression model is to predict the categories of the response variable as function of covariates involved in the model. However, in bivariate logistic regression model we try to examine the relationship between two different dichotomous response variables (Y_1 and Y_2), by modelling jointly as a function of explanatory variables under consideration. As depicted in table1, each pair of dependent variables (Y_{i1}, Y_{i2}) has four possible outcomes, ($Y_{i1}=1, Y_{i2}=1$), ($Y_{i1}=1, Y_{i2}=0$), ($Y_{i1}=0, Y_{i2}=1$), and ($Y_{i1}=0, Y_{i2}=0$).

The joint probability for each of these four outcomes is modeled with three systematic components, these are:

- The marginal (separate) distributions, that is $P(Y_{i1}=1)$ and $P(Y_{i2}=1)$
- The odds ratio, which is denoted by ψ describes the dependence of one marginal distribution on the other.

Bivariate logistic regression model or bivariate logistic odds-ratio model proposed by [23, 27, 28, 32] is specified by modeling the marginal distribution of each of Y_j , and the common odds ratio. The odds ratio or ψ is defined as the ratio of the odds of $Y_1=1$ given that $Y_2=1$ and the odds of $Y_1=1$ given that $Y_2=0$, that is $\psi = \frac{p_{11}p_{00}}{p_{10}p_{01}}$. The odds ratio is used to measure the association between the two dichotomous response variables (Y_1 and Y_2); the value of ψ equal to one indicates lack of association between Y_1 and Y_2 .

Application of bivariate logistic regression provides adjusted estimates of concordance through simultaneous estimation of covariate effects on the odds ratio that describes the pair wise association structure. Furthermore, adopting bivariate logistic regression analysis may offer greater precision than unadjusted estimates obtain from other possible categorical distributions. The influence of the independent variables on the marginal probabilities that is, marginal probability of diarrhea (Y_1) and ARI (Y_2), and odds ratio can be expressed using regression models. Likewise, the association between the covariates and each disease can be examined using separate logistic regression models for each outcome:

$$\text{logit}[E(Y_{1j})] = \text{logit}[\text{Pr}(Y_{1j} = 1 | X_{1j}, \beta_1)] = \beta_1^T X_{1j} \quad (1)$$

$$\text{logit}[E(Y_{2j})] = \text{logit}[\text{Pr}(Y_{2j} = 1 | X_{2j}, \beta_2)] = \beta_2^T X_{2j} \quad (2)$$

These standard logistic regression models, indicated in equation (1) and 2, do not consider the correlation between the two childhood diseases. However, we assume a set of unobserved

random effects are shared by the two diseases of the same individual, that is, the two diseases share common unobservable features. Let u_j denote the random intercept shared by the two diseases of the j^{th} ($j=1,2,\dots, m$) individual. Let ϕ_{1i} and ϕ_{2i} define dummy variables with $\phi_{1i} = 1$ for $i=1$ and $\phi_{2i} = 1$ for $i=2$. Then, the joint response model for the bivariate logit is given by

$$\begin{aligned} \text{logit}[E(Y_{ij}|u_j)] &= \text{logit}[\Pr(Y_{ij} = 1|X_{ij}, \tilde{\beta}_i, u_j)] \\ &= \phi_{1j}(\tilde{\beta}_1^T X_{1j} + u_j) + \phi_{2j}(\tilde{\beta}_2^T X_{2j} + u_j) \end{aligned} \quad (3)$$

Hence, in this equation, the bivariate responses (Y_{1j}, Y_{2j}) of all individuals are stacked into a single response vector. The random intercepts are assumed to vary independently from one individual to another. In addition, the random intercepts are assumed to be normally distributed with zero mean and variance σ_u^2 . Let G denote the distribution of u_j . The joint response of an individual is assumed to be independent given the shared random intercept. By using conditional independence from this assumption, we can write the likelihood function of the joint response model as follows:

$$\begin{aligned} L(.) &= \prod_{j=1}^m \Pr(Y_{ij} = 1|X_{ij}, \tilde{\beta}_i, u_j) \\ &= \prod_{j=1}^m \left\{ \int \Pr(Y_{ij} = 1|X_{ij}, \tilde{\beta}_i, u_j) \phi G(u_j) \right\} \\ &= \prod_{j=1}^m \left\{ \int \prod_{i=1}^2 \Pr(Y_{ij} = 1|X_{ij}, \tilde{\beta}_i, u_j) \phi G(u_j) \right\} \\ &= \prod_{j=1}^m \left\{ \int \prod_{i=1}^2 \frac{e^{u_j + \tilde{\beta}_1^T X_{1j}}}{1 + e^{u_j + \tilde{\beta}_1^T X_{1j}}} \phi G(u_j) \right\} \end{aligned} \quad (4)$$

Following Muse (2015), we used maximum likelihood estimation to obtain estimates of the model parameters $(\tilde{\beta}_1, \tilde{\beta}_2, \sigma_u^2)$. The estimation was based on the maximization of the log-likelihood function i.e. maximizing the logarithm of the likelihood function. However, since the integrals do not have closed-form solutions for equation (4), parameters of the joint model were estimated simultaneously via maximum likelihood by evaluating the integrals using Gaussian adaptive quadrature approximation [17].

3. Results and Discussion

Analyses of this study were based on total valid observations of 9917 children under the age of five years. As provided in Table 2, 5.9% and 8.2% of the male child and also 5.1% and 7.8% of the female child diarrhea and ARI cases were reported respectively in the past fifteen days. Geographically, 4.6% of the reported childhood diarrhea and ARI lived in urban areas, as compared to 22.1% who lived in rural areas. Regarding vaccination, 63.3% of the children in Ethiopia were reported as received various types of vaccination. The covariates such as

residence, vaccination, antenatal visit, wealth index, mother's education, birth order, drink water source and region with ($P < 0.05$) were found significantly associated with diarrhea and ARI. Regionally, a low prevalence of childhood diarrhea was observed in the nation's capital Addis Ababa that could be due to the relative socioeconomic development advantages in terms of access to health services, and a relatively higher prevalence of childhood diarrhea in Oromia and SNNP regions but each disease has its typical geographical pattern of variation. Table 3 also indicates that the prevalence of 4.3% childhood comorbidity of diarrhea and ARI noted in Ethiopia in 2016 DHS. On the other hand, 11.7% of the children had only ARI and 6.7% of the children had only diarrhea morbidity.

Table 2: Distribution of risk factors for childhood comorbidity in Ethiopia (DHS 2016)

	Diarrhea N (%)			ARI N (%)		
	Yes	No	Sig.	Yes	No	Sig.
Sex of child						
Male	587 (5.9)	4476 (45.1)	0.05	811(8.2)	4252 (42.9)	0.855
Female	503 (5.1)	4352 (43.9)		771 (7.8)	4083 (41.2)	
Place of residence						
Urban	186(1.9)	1689 (16.9)	0.001	275 (2.7)	1601 (16.0)	0.001
Rural	904(9.0)	7137 (71.3)		1311 (13.1)	6736 (67.3)	
Vaccination						
Yes	2472 (9.1%)	2118 (54.2%)	0.001	2472(12.3%)	1993(51.0%)	0.002
No	1407(3.7%)	1264 (32.3%)		1407(5.4%)	1194(30.5%)	
Antenatal visit						
No	261 (3.8)	2054 (29.7)	0.001	384(5.6)	1933 (28.0)	0.029
Less 5 visits	445 (6.4)	2668 (38.6)		584 (8.4)	2531(36.6)	
More 5	211(3.1)	1227(17.7)		282 (4.1)	1156 (16.7)	
Wealth index						
Poorest	365 (3.6)	3311 (33.1)	0.001	514(5.1)	3163(31.6)	0.001
Porrer	192 (1.9)	1469 (14.7)		305 (3.0)	1359 (13.6)	
Middle	172 (1.7)	1206 (12.1)		228 (2.3)	1152 (11.5)	
Richer	159 (1.6)	1055 (10.5)		225 (2.2)	990(9.9)	
Richest	202 (2.0)	1785 (17.8)		314(3.1)	1673(16.7)	
Mother's Education						
No	664 (6.6)	5679 (56.8)	0.001	982(9.8)	5366(53.6)	0.001
Primary	315 (3.1)	2188 (21.9)		456 (4.6)	2049 (20.5)	
Secondary	74(0.7)	613 (6.1)		101(1.0)	586(5.9)	
Higher	37(0.4)	346 (3.5)		47(0.5)	336(3.4)	

Table 3: Cross-classification of children by diarrhea and ARI diseases in Ethiopia m (%)

		Had ARI		Total
		No	Yes	
Had diarrhea	No	7671(77.4%)	1156(11.7%)	8827(89.0%)
	Yes	664(6.7%)	426(4.3%)	1090(11.0%)
Total		8335(84.0%)	1582(16.0%)	9917(100.0%)

$$\chi^2 = 488.65, df = 1, P < 0.001$$

As summarized in Table 4, the common odds ratio (4.30) which exceeds one describes a positive relationship between the two childhood diseases. Hence, bivariate logistic regression model was considered suitable to investigate the related risk factors associated with diarrhea and ARI diseases. This model provides estimates with vibrant properties of concurrence through simultaneous estimation of effects of risk factors on the odds ratio. Besides, the advantage of the model is that it may give better precision than unadjusted estimates obtained by considering a multinomial distribution. The effects of risk factors on the marginal probability of the two diseases and the odds ratio are each described with regression equations [31]. The bivariate model is parametrically dependent if Y_1 and Y_2 share some or all explanatory variables, and the effects of the shared explanatory variables are jointly estimated.

Before building the models, covariates were screened as they show some significant association with diarrhea and ARI in the chi-square and descriptive analysis.

Maximum likelihood estimator of the model parameters obtained in R using the VGAM by [29] and Zeligchoice by [19] packages. As reported in Table 3, in the separate model, we found the covariates such as vaccination, mother's education, and mother's antenatal care statistically significant in affecting child diarrhea status whereas residence, number of children, vaccination, mother's education, and wealth index were found significantly significant factors influencing childhood acute respiratory disease status. Usually, standard models that ignore interdependence between the two or more responses result in relatively biased estimates [15]. Thus, we employed a joint model that captures the dependence between childhood comorbidity in assessing the potential risk factors for diarrhea and acute respiratory infection (ARI).

Table 4: Results for risk factors of childhood diarrhea and ARI in separate and joint models

Variable	Marginal Model				Joint Model			
	Diarrhea		ARI		Diarrhea		ARI	
	OR	SE	OR	SE	OR	SE	OR	SE
Constants	0.0786**	0.3218	0.1182**	0.2808	0.0766**	0.3216	0.1190**	0.2806
Residence1								
Residence2	1.2871	0.2470	1.7303**	0.2181	1.2865	0.2466	1.7235**	0.2178
Childsize	0.9425	0.0665	0.8788**	0.0585	0.9456	0.0664	1.0084**	0.0585
Vaccin1	1.3905**	0.1171	1.3059**	0.1006	1.4097**	0.1173	1.3034**	0.1006
Vaccin2								
Educ0								
Educ1	1.1618	0.1197	1.0591	0.1068	1.1595	0.1196	1.0605	0.1068
Educ2	0.6111**	0.2848	0.6008**	0.2438	0.6043**	0.2852	0.6001**	0.2439
Educ3	0.3933*	0.5408	0.6082	0.3847	0.3918**	0.5403	0.6128	0.3839
Wealth1								
Wealth2	1.0660	0.1424	1.2079	0.1235	1.0844	0.1419	1.2069	0.1235
Wealth3	1.1634	0.1519	1.0823	0.1381	1.1633	0.1520	1.0817	0.1381
Wealth4	.2446	0.1654	1.4342**	0.1450	1.2472	0.1654	.4323**	0.1450
Wealth5	.0362	0.2460	1.6035**	0.2075	1.0386	0.2458	.5990**	0.2073
Antenatal1	1.1037	0.1161	1.1482	0.1006	1.1089	0.1161	1.1447	0.1005
Antenatal2	1.5853**	0.1564	1.2003	0.1446	1.6112**	0.1560	1.1938	0.1447
Antenatal3								
Odds ratio								4.30

* Significant at 10% level of significance, ** Significant at 5% level of significance

In the joint response model, covariates such as vaccination, Mother's education, antenatal care were found significant factors for childhood diarrhea. Similarly, residence, number of children, vaccination, mother's education and wealth index were significant factors for acute respiratory infection (ARI) among children age under five year. Specifically, we noticed that children who had not received vaccinations 1.4 times more likely to develop diarrhea and ARI. Regarding parent's education level, children from mothers with secondary and higher education were 0.60 and 0.39 times respectively less likely to be affected by diarrhea than mothers with no education. These findings are consistent with a study conducted in Egypt [21]. Furthermore, children from mothers who visited less antenatal care centers during pregnancy were 1.61 times more likely to encounter childhood diarrhea than mother's who had normal antenatal care visit.

On the other hand, children living in rural areas were 1.72 times more likely to be infected by ARI disease than their counterparts. This may be associated with inadequate health facilities and poor lifestyle in rural areas. Regarding the number of children ever born, we found that the likelihood of ARI infection increases, as the number of children increases i.e. a child with more children in the family are more exposed to acute respiratory disease.

Moreover, children from mothers with secondary school education were 0.60 less likely to be infected by ARI than mothers with no education. Regarding wealth quantile, we found that a child from a rich family was less likely to be infected by ARI i.e. children in wealthier families were less likely to be infected by acute respiratory disease than children in less wealthy families. Being in a higher wealth status, compared to the lowest (poorest), was revealed to reduce the probability of occurrence of diarrhea and ARI which is in line with that was found in Uganda [22].

More notably, we observed that standard errors of the parameter estimates in the joint response model were smaller compared to the corresponding standard errors from the marginal models. This indicates efficiency gains in the joint model as compared to the marginal models. As a result, these findings validated the need for jointly modeling of two correlated responses.

3.1. Simulation Result

Simulation can help researchers understand the entire statistical model, take full advantage of the parameter estimates, and convey findings in a reader-friendly manner [30].

As reported in Appendix Table 1-Table 5, for each of the two categories of the dependent variable, the means of the predicted probabilities of an event were calculated. Then, the absolute value of the difference between those two means was taken. If a model makes good predictions, the cases with events should have high predicted values and the cases without events should have low predicted values [20]. Along with the tables and graphs of the quantities of interest in Graph 1(Appendix) indicated that the built model can make good prediction. Each row of statistics for the expected and predicted values includes the mean estimate for Y , the standard deviation, the median, and the 95% confidence interval measures of uncertainty around the average.

4. Conclusion

In this paper, we specifically focused on childhood comorbidity using a joint response model that accommodates the interdependence between the two diseases, diarrhea and ARI. Different socio-demographic as well as other biological risk factors of diarrhea and ARI were considered. Among 9917 under-five children from DHS (2016) in Ethiopia, 4.3% childhood comorbidity of diarrhea and ARI was noted. A positive correlation between the two infections was observed, and was taken into consideration in this study thereby obtaining unbiased estimates. Preceding studies of diarrhea and ARI comorbidity modeled each disease separately ignoring the potential correlation between the two diseases. Findings from this

study contributed to appreciate joint models that deflate parameters otherwise may be overestimated in separate models. In the joint model, residence, vaccination, mother's education, and antenatal care visits during pregnancy were found significant risk factors for diarrhea whereas residence, number of children ever born, vaccination mother's education, wealth index were statistically significant risk factors childhood for ARI. The two correlated binary response variables, diarrhea, and ARI were together affected significantly by the risk factors such as residence, vaccination, and mother's education. As a limitation, data set from Ethiopian DHS 2016 included most of the important covariates but spatial information was absent, thereby we couldn't investigate the effect of spatial variations. Future research should look into spatial effects in order to explain the variation as a result of geographical difference for diarrhea and ARI childhood morbidities.

Acknowledgements

The authors would like to acknowledge the Demographic and Health Survey (DHS) database team for providing data for this study. The comments of the reviewers and the editor will be gratefully acknowledged.

Authors' contributions

TAB conceptualized and designed this analysis. TAB and HKG interpreted the results, drafted and revised the manuscript. Both authors read and approved the final manuscript.

Funding

Not applicable.

Disclosure statement

No potential conflict of interest was reported by the authors.

Ethics approval and consent to participate

Not applicable.

Author details

¹Department of Statistics, Haramaya University, Dire Dawa, Ethiopia

Reference

- [1] C.McCulloch, Joint modelling of mixed outcome types using latent variables. *Statistical Methods in Medical Research*. 2008; 17:53–73.
- [2] S.L.Zeger and K.Y. Liang, Feedback models for discrete and continuous time series. *Statistica Sinica*. 1991; 1:51–64.

- [3] Central Statistical Agency [Ethiopia] and ICF International. Ethiopia Demographic and Health Survey 2011; Addis Ababa, Ethiopia and Calverton, Maryland, USA: 2012; Central Statistical Agency and ICF International.
- [4] Miller, P.Nathan,A. Amouzou, M. Tafesse, E. Hazel, H. Legesse, T. Degefie, C.G. Victora, R. E. Black, and J. Bryce, Integrated Community Case Management of Childhood Illness in Ethiopia: Implementation Strength and Quality of Care. The American Journal of Tropical Medicine and Hygiene.2014; 13:751.
- [5] World Health Organization (WHO) and UNICEF. Ending Preventable Child Deaths from Pneumonia and Diarrhea by 2025: The Integrated Global Action Plan for Pneumonia and Diarrhea (GAPPD). Geneva, Switzerland:2013; WHO and UNICEF.
- [6] N.B.Kandala,G.Gebrenegus, Ageo-addictive bayesian discretetime survival model and its application to spatial analysis of childhood mortality in malawi. Quality and Quantity 2006; 40: 935-957.
- [7] N.B.Kandala, M.A.Magadi, andN.J.Madise, An investigation of district spatial variations of childhood diarrhoea and fever morbidity in Malawi. Soc Sci Med 2006; 62: 1138-1152.
- [8] K.Khatib and N.B. Kandala, Latent variable modelling of risk factors associated with childhood diseases: Case study for Nigeria', Asian Pacific Journal of Tropical Disease. Asian Pacific Journal of Tropical Disease.2011; 2222–1808:169–176.
- [9] L.N.Kazembe, A.S.Muula, C.C.Appleton, I. Kleinschmidt, Modelling the effect of malaria endemicity on spatial variations in childhood fever, diarrhea and pneumonia in Malawi. Int. J. Health Geogr. 2007, 6, 33–43.
- [10] C.G.Victora, T.Adam, J.Bryce, andD.B. Evans, Integrated Management of the Sick; Child Disease Control Priorities Project: Washington, DC, USA, 2006. Available Online: <http://www.dcp2.org/pubs/DCP/63/Section/9372> (accessed on 15 July 2019).
- [11] B.Fenn, S. Morris, andR.E. Black, Comorbidity in childhood in Ghana: Magnitude, associated factors and impact on mortality. Int. J. Epidemiol. 2005, 34, 368–375.
- [12] L.N.Kazembe and J.J. Namangale, A Bayesian multinomial model to analyse spatial patterns of childhood co-morbidity in Malawi. Eur. J. Epidemiol. 2007, 22, 545–556.
- [13] TW. Yee, *Vector Generalized Linear and Additive Models: With an Implementation in R*. 2015; Springer, New York, USA.
- [14] E. Bbaale, Determinants of diarrhea and acute respiratory infection among under-fives in Uganda. *Australas Med J*. 2011; 4(7): 400–409.
- [15] M. Ghebremikael, Joint modeling of correlated binary out comes: HIV-1and HSV-2 co-infection. Journal of Applied Statistics 2015, 42:2180 <https://doi.org/10.1080/02664763.2015.1022138>
- [16] J. R. Dale, Global cross-ratio models for bivariate, discrete, ordered responses. Biometrics1986;42: 909-917.
- [17] Skrondal and S. Rabe-Hesketh, *Generalized Latent Variable Modeling*, Chapman & Hall, New York, 2004.
- [18] T. W. Yee and T. J. Hastie, Reduced-rank vector generalized linear models. StatisticalModelling, 2003; 3:15-41.
- [19] Kosuke Imai, Gary King, and Olivia Lau.. *VZelig: Everyone's Statistical. Software*2007; <http://GKing.harvard.edu/zelig>.
- [20] T. Tjur, Coefficients of determination in logistic regression models—A new proposal: The coefficient of discrimination. The American Statistician63: 2009; 366-372.
- [21] AH. El-Gilany and S. Hammad, Epidemiology of diarrheal diseases among children under age 5 years in Dakahlia, Egypt. Eastern Mediterranean Health Journal. 2000;11(4): 762-775.

- [22] E. Bbaale, Determinants of diarrhea and acute respiratory infection among under-fives in Uganda. AMJ 2011, 4, 7, 400-409 <http://dx.doi.org/10.4066/AMJ.2011.723>
- [23] P.McCullagh and J A.Nelder, Generalized Linear Models (second edition). London, United Kingdom: Chapman & Hall; 1989.
- [24] R.E. Black, S.S. Morris, and J. Bryce, Where and why are 10 million children dying every year? Lancet, 2003; 361(9376), 2226–2234, DOI 10.1016/s0140-6736(03)13779-8.
- [25] M. Kosek, C. Bern, and R.L. Guerrant, The global burden of diarrhoeal disease, as estimated from studies published between 1992 and 2000. Bull. World Health Organ, 2003; 81(3), 197–204.
- [26] A. Downing, D. Forman, M.S. Gilthorpe, K.L. Edwards, and S.O. Manda, Joint disease mapping using six cancers in the Yorkshire region of England. Int. J. Health Geogr. 2008; 7:41–55. doi: 10.1186/1476-072X-7-41.
- [27] J. Palmgren, Regression Models for Bivariate Binary Responses. Technical Report no. 101, Department of Biostatistics, University of Washington, Seattle. 1989.
- [28] S. Cessie, and J. C. Houwelingen, Logistic regression for correlated binary data. Applied Statistics, 1994; 43, 95–108.
- [29] T. WYee, and T. Dirnbock, Models for analysing species' presence/absence data at two time points. Journal of Theoretical Biology 2009; 259, 684–694.
- [30] Fair, Ray C, Estimating the Expected Predictive Accuracy of Econometric Models. International Economic Review 1980; 21:355–378.
- [31] Abud Darda Ghazanfar Ali. Modelling of African Farm Dynamics Using Bivariate Binary Logistic Regression in WinBUGS. Mater Thesis. 2009; Lund Univesity.
- [32] McCullagh, P. and Nelder, J. A. Generalized Linear Models. 2nd ed. 1989; London: Chapman & Hall.

Appendices

1. Quantities of interest from bivariate logit model

Simulation for x (lowest value of explanatory variable)

Table 1: Expected values

	mean	sd	50%	2.5%	97.5%
Pr(Y1=0, Y2=0)	0.74854988	0.014409098	0.74913524	0.71984276	0.77557475
Pr(Y1=0, Y2=1)	0.12278475	0.010412208	0.12256883	0.10387570	0.14359428
Pr(Y1=1, Y2=0)	0.07540105	0.007805833	0.07522129	0.06004650	0.09160451
Pr(Y1=1, Y2=1)	0.05326432	0.006478022	0.05286303	0.04232368	0.06682350

Table 2: Predicted values

	0	1
(Y1=0, Y2=0)	0.289	0.711
(Y1=0, Y2=1)	0.843	0.157
(Y1=1, Y2=0)	0.888	0.112
(Y1=1, Y2=1)	0.980	0.020

Simulation for x1 (highest value of the explanatory variable)

Table 3: Expected values

	mean	sd	50%	2.5%	97.5%
Pr(Y1=0, Y2=0)	0.83784969	0.04727187	0.84443911	0.729413394	0.9124655
Pr(Y1=0, Y2=1)	0.10031222	0.03759618	0.09337025	0.046248931	0.1913946
Pr(Y1=1, Y2=0)	0.04132368	0.02283666	0.03550673	0.011918510	0.1019103
Pr(Y1=1, Y2=1)	0.02051441	0.01307836	0.01775972	0.005519475	0.0505012

Table 4: Predicted values

	0	1
(Y1=0, Y2=0)	0.160	0.840
(Y1=0, Y2=1)	0.897	0.103
(Y1=1, Y2=0)	0.951	0.049
(Y1=1, Y2=1)	0.992	0.008

Table 5: Fist difference

	mean	sd	50%	2.5%	97.5%
$\Pr(Y_1=0, Y_2=0)$	0.08929981	0.04562680	0.09516639	-0.01514334	0.16144689
$\Pr(Y_1=0, Y_2=1)$	-0.02247253	0.03606915	-0.02729431	-0.07446998	0.06537456
$\Pr(Y_1=1, Y_2=0)$	-0.03407737	0.02257006	0.03796583	-0.06516763	0.02111796
$\Pr(Y_1=1, Y_2=1)$	-0.03274992	0.01293345	-0.03455233	-0.05159941	-0.00274422

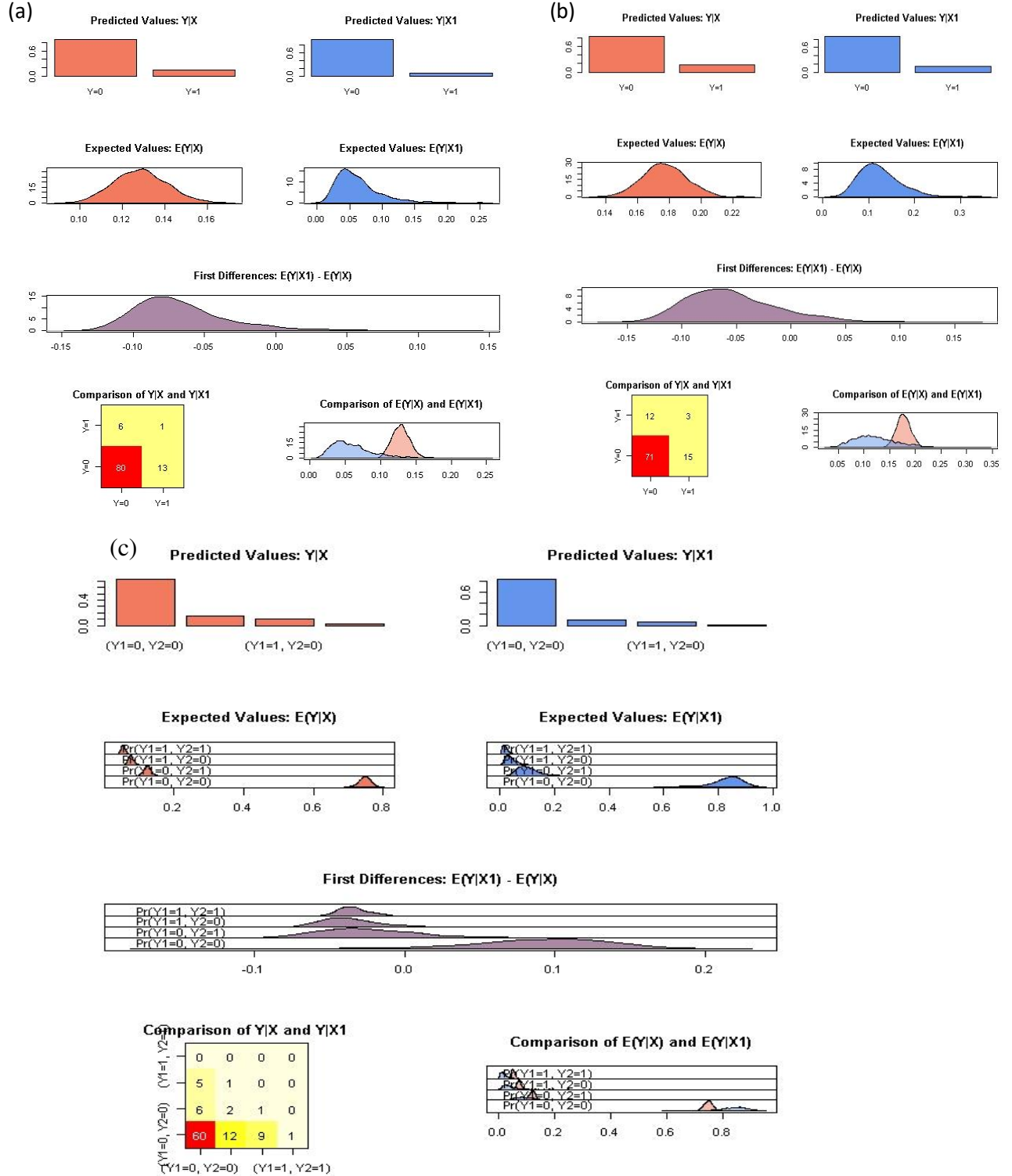


Fig 1: Graphs of quantities of interest for Diarrhea model (a), ARI model (b) and Joint model(c)

Figures

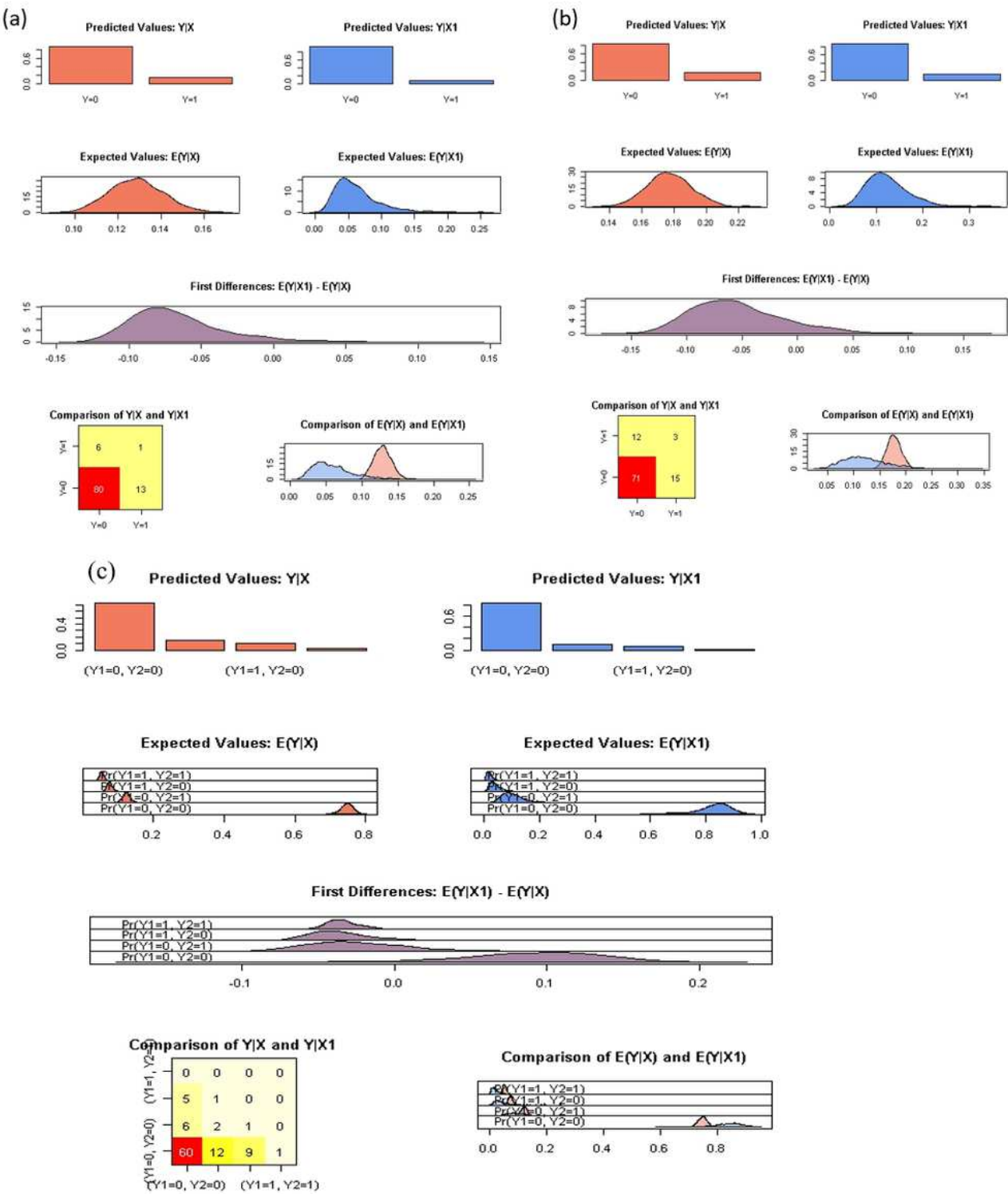


Figure 1

Graphs of quantities of interest for Diarrhea model (a), ARI model (b) and Joint model(c)

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Supplementarydata.sav](#)