Contents lists available at ScienceDirect

Asian Pacific Journal of Tropical Disease

journal homepage:www.elsevier.com/locate/apjtd

Document heading

Latent variable modelling of risk factors associated with childhood diseases: Case study for Nigeria

Khaled Khatab[1*], Ngianga-Bakwin Kandala[2]

[1]Institute of Occupational and Social Medicine, Medical Faculty, RWTH Aachen University, Pawelsstr.30 D-52074 Aachen, Germany
[2]Health Sciences Research Institute, University of Warwick, Warwick Medical School, Warwick Evidence, Gibbet Hill, Coventry CV4 7AL, UK

ARTICLE INFO

Article history:
Received 20 May 2011
Received in revised form 27 June 2011
Accepted 11 July 2011
Available online 28 September 2011

Keywords:
Associations between childhood diseases
Spatial analysis
Latent variable models

ABSTRACT

Objective: To investigate the impact of various bio-demographic and socio-economic variables on joint childhood diseases in Nigeria with flexible geoadditive probit models. **Methods:** Geoadditive latent variable model (LVM) was applied where the three observable disease (diarrhea, cough, fever) variables were modelled as indicators for the latent individual variable "health status" or "frailty" of a child. This modelling approach allowed us to investigate the common influence of risk factors on individual frailties of children, thereby automatically accounting for association between diseases as indicators for health status. The LVM extended to analyze the impact of risk factors and the spatial effects on the unobservable variable "health status" of a child less than 5 years of age using the 2003 Demographic and Health Surveys (DHS) data for Nigeria. **Results:** The results suggest some strong underlying spatial patterns of the three ailments with a clear southeastern divide of childhood morbidities and this might be the results in the overlapping of the various risk factors. **Conclusions:** Comorbidity with conditions such as cough, diarrhoea and fever is common in Nigeria. However, little is known about common risk factors and geographical overlaps in these illnesses. The search for overlapping common risk factors and their spatial effects may improve our understanding of the etiology of diseases for efficient and cost-effective control and planning of the three ailments.

1. Introduction

The prevalence of childhood diseases in Nigeria remain high during the last decade due to the burden associated with highly prevalent diseases such as diarrhea, malaria and acute respiratory infections (ARI). These three ailments are the leading causes of death globally among children under five years of age. There is much uncertainty surrounding the ascertainment of pneumonia as the underlying cause of death, principally from verbal autopsy-based studies in developing countries.

United Nations International Children's Emergency Fund (UNICEF) statistics in 2008 for Nigeria show that 186 out of 1000 children under the age of 5 years died, whilst 197 out of 1 000 children under the age of 5 years died in 2004[1]. Policy-makers and governments in Sub-Saharan African (SSA) countries rely on household surveys to inform public health policies.

*Corresponding author: Dr. Khaled Khatab, Institute of Occupational and Social Medicine-RWTH Aachen University, Pauwelsstraße 30- 52074 Aachen, Germany.
Tel: +49-2418036733
Fax: +49-241 8082587
E-mail: N-B.Kandala@warwick.ac.uk

The Demographic and Health Surveys (DHS) are a well-established source of reliable population data with a substantial focus on childhood diseases and health-seeking behaviour[2].

However, linking individual survey records with disease prevalence at the disaggregated small-scale, community level has not been possible because of the methodological restraints of traditional regression modelling.

There are few studies that have investigated the association between diseases and socioeconomic, environmental and individual risk factors in Nigeria and many other SSA countries[3-7]. Furthermore, recent studies have applied geoadditive regression models[7,8] to investigate the association between diseases and geographic areas[5,6,9].

These models can account for nonlinear covariate effects and geographical variation while simultaneously controlling for other important risk factors. They have been used in regression studies of risk factors, for morbidity[5,6,9] and for mortality in SA[10,11].

However, in these studies, regression analyses were carried out separately for each type of disease, such as cough, fever and diarrhoea, neglecting possible joint association among these response variables and without aiming at the detection of common latent risk factors. In this study, we take a different approach. We apply geoadditive latent variable models (LVMs) for mixed discrete response, considering binary indicators for cough, fever and diarrhea as observable outcomes of latent health status[12–14].

Health status is also known to have various risk factors including geographical location as a proxy for socioeconomic and environmental factors that affect disease prevalence and incidence.

Spatial heterogeneity in these factors influences the disease transmission pattern. Consequently, efforts to reduce the burden of childhood diseases should investigate the influence of the associations between the different type of diseases and their distribution among the locations on child health.

Indeed, the impact of geographic location on child health in developing countries is well documented in many studies[11].

Most of these studies have emphasized on studying a single disease at a time. However, because of common and overlapping risk factors, separate analyses may fail to give a comprehensive picture of the epidemiology of the diseases and the joint effects of childhood diseases at the population level.

Epidemiological methodology and research have increased over the years, and have extended from studying a single disease at a time[9,11] to joint analysis of several diseases using the LVMs or to a simple joint model for only two types of diseases[15–17].

Two approaches of LVMs (joint model) analysis of diseases have emerged: the measurement model, which accommodates and describes the effect of the latent variables and a set of observed covariates (*e.g.* child's sex, mother's educational attainment, working status, *etc.*) on the health indicators such as diarrhoea, fever and cough.

The structural model is linking a set of observed covariates, which have indirect effects (such as child and mother's age) on the latent variables.

In the latent variables, all risks are estimated using covariates. In addition, the correlation of risk between diseases can be quantified.

In this study, we considered the latent variables model to jointly analyse childhood fever, cough and diarrhoea in Nigeria, with the objective of highlighting spatial patterns of these illnesses and we focus on the first approach.

In geoadditive probit LVMs, the three observable binary disease variables are taken as indicators for the latent individual variable "health status" or "frailty" of a child. This approach is used in order to study the influence of risk factors on individual frailties of children, thereby automatically accounting for association between diseases as indicators for health status. Compared to previous results, our approach can provide new insights into childhood morbidity and mortality in developing countries in general and, more specifically, in Nigeria.

All computations have been carried out with a user written code in R programs using the MCMC package[12,13,15]. The most important covariate included in this analysis is the geographic location, where the child lives, that includes features of the wider socio–cultural and political context affecting both the child and his/her care givers. Other selected socio–demographic variables available in the data are grouped as individual child characteristics, mother characteristics, household economical level and community characteristics. Regarding the covariates, we were guided by the previous literature on the subject and the conceptual framework outlined in UNICEF (1998). Among the underlying determinants of disease, we considered a proxy measure of current or recent socioeconomic status (SES), having radio, television *etc*., the nutritional status of the mother measured by Body Mass Index (BMI), health knowledge and care practices measured by mother's education, mother's marital status, birth interval and place of delivery of children (which we included in the early stages of this study and it had no significant effect).

We also control for the sex of the child, urban rural location, and the age of child. Based on our previous own work as well as other literature, we investigated a potentially non–linear pattern of effects of the mother's BMI and age as well as the age pattern of the child on diseases. Empirical distributions of all factors used in the analysis are given in Table 2[18].

2. Materials and methods

The analysis in this work is based on data available from the 2003 Nigerian Demographic and Health Survey (NDHS)[2]. The NDHS uses standard survey instruments to collect data on household members such as sex of child, age of child, mother's age, current employment status of mother, mother's educational attainment, exposure to mass media, the type of toilet facility *etc*. It collects information on household living conditions such as housing characteristics, on childhood morbidity, malnutrition and child health from mothers in reproductive ages (15–49). The data is based on national samples that have been collected using questionnaires and allows for breakdowns by urban–rural and major regions and states.

The diseases of children included in this work were diarrhea, respiratory diseases, cough and fever. These diseases were still major causes of mortality among children in many developing countries, particularly in Sub–Saharan Africa.

Yet, except for some descriptive reports by National Statistics Offices of these countries, few systematic studies of factors that influence the prevalence of diarrhea, cough and fever among young children were carried out in these countries. The success of healthcare intervention depends on a correct understanding of the socioeconomic, environmental and cultural factors that determine the occurrence of diseases, undernutrition and deaths. The mapping of variations in risk of child morbidity and child malnutrition can help to improve the targeting of scarce resources for public health interventions. Bearing in mind that direct mapping of relevant environmental risk factors (which may vary considerably in both space and time) is difficult and this has led to investigations of environmental proxies[9].

Table 1 shows the prevalence of three ailments, fever, cough and diarrhea, which are used as response variables in this work.

Table 1
Overview of diseases in Nigeria [*n* (%)].

Variable	0: had no diseases	1: had diseases	Total
Fever	3 583 (69.09%)	1 603 (30.91%)	5 186 (100%)
Cough	3 967 (76.49%)	1 219 (23.51%)	5 186 (100%)
Diarrhea	4 257 (82.09%)	929 (17.91%)	5 186 (100%)

It also indicates the percentages of children under five years of age, who had diseases during two weeks before the survey.

2.1. Covariates

Based on preliminary analyses with separate geoadditive regression models[6], the following covariates were considered to study childhood morbidity in Nigeria.

2.1.1. Metrical covariates

Chage: Child's age in months.
BMI: Mother's body mass index.
Mageb: Mother's age at birth.

2.1.2. Categorical covariates (in effect coding)

Male: Child's sex: male or female (reference category).

Educ: Mother's educational attainment: incomplete primary, complete primary, and incomplete secondary school; or complete secondary school and higher eduction (reference category).

Trepr: Whether mother had treatment during pregnancy: yes or no (reference category).

Anvis: Whether mother had antenatal care: yes or no (reference category).

Water: Source of drinking water: controlled water or no (reference category).

Toilet: Flush toilet at household, or not (reference category).

Urban: Locality where respondent lives: urban or rural (reference category).

Radio: Radio at household: yes or no (reference category).

Elect: Electricity : yes or no (reference category).

Work: Mother's current working status: working or not (reference category).

2.1.3. Spatial covariate

Reg: state in Nigeria where the respondent resides

2.2. Statistical analyses

2.2.1. Latent variable models for binary responses

The basic idea of factor analysis and LVMs is that the vector of the p observable variables can be represented, at least partly, by one or more latent factors or variables

Table 2
Distribution of factors analyzed in childhood morbidity in Nigeria.

Factors		Diarrhea			Fever			Cough		
		Yes	No	*P*–value	Yes	No	*P*–value	Yes	No	*P*–value
Sex of child	Male	14.56 (432)	83.77 (2 565)	0.072	26.71 (818)	73.29 (2 244)	0.820	20.05 (614)	79.95 (2 448)	0.740
	Female	6.23 (497)	85.44 (2 535)		26.46 (785)	73.54 (2 182)		20.39 (605)	79.61 (2 362)	
Place of residence	Urban	13.27 (281)	86.73 (1 837)	0.001	24.08 (510)	75.92 (1 608)	0.000	19.74 (418)	80.26 (1 700)	0.490
	Rural	16.57 (648)	83.43 (3 263)		27.95 (1 093)	72.05 (2 818)		20.48 (801)	79.52 (3 110)	
Working	Yes	14.86 (570)	85.14 (3 265)	0.120	26.41 (1 013)	73.59 (2 822)	0.680	21.38 (820)	78.61 (3 015)	0.003
	No	16.36 (359)	83.64 (1 835)		26.89 (590)	73.11 (1 604)		18.19 (399)	81.81 (1 795)	
Mother's education	Yes	10.20 (75)	89.80 (660)	0.001	23.27 (171)	76.73 (564)	0.030	19.73 (145)	80.27 (590)	0.720
	No	16.13 (854)	83.87 (4 440)		27.05 (1 432)	72.95 (3 862)		20.29 (1 074)	79.71 (4 220)	
Pregnancy's treatment	Yes	17.65 (516)	82.35 (2 407)	0.000	31.24 (913)	68.76 (2 010)	0.001	23.47 (686)	76.53 (2 237)	0.000
	No	13.30 (413)	86.70 (2 693)		22.22 (690)	77.78 (2 416)		17.16 (533)	82.84 (2 573)	
Drinking water	Controlled	13.57 (125)	86.43 (796)	0.090	25.84 (238)	74.16 (683)	0.570	19.11 (176)	80.89 (745)	0.360
	Not controlled	15.74 (804)	84.26 (4 304)		26.72 (1 365)	73.28 (3 743)		20.42 (1 043)	79.58 (4 065)	
Radio	Yes	14.42 (644)	85.58 (3 822)	0.000	26.11 (1 166)	73.89 (3 300)	0.150	20.26 (905)	79.74 (3 561)	0.880
	No	18.23 (285)	81.77 (1 278)		27.96 (437)	72.04 (1 126)		20.09 (314)	79.91 (1 249)	
Electricity	Yes	13.44 (365)	84.59 (2 350)	0.000	25.01 (679)	74.99 (2 036)	0.012	20.04 (544)	79.96 (2 171)	0.750
	No	17.02 (564)	82.98 (2 750)		27.88 (924)	72.12 (2 390)		20.37 (675)	79.63 (2 639)	
Toilet facility	Own flush toile facility	9.15 (54)	90.85 (536)		18.81 (111)	81.19 (479)		20.85 (123)	79.15 (467)	
				0.000			0.000			0.680
	Other and no toilet facility	16.09 (875)	83.91 (4 564)		27.43 (1 492)	72.57 (3 947)		20.15 (1 096)	79.85 (4 343)	
Antenatal visit	Yes	17.25 (416)	82.75 (1 996)	0.000	31.51 (760)	68.49 (1 652)	0.000	25.46 (614)	74.44 (1 798)	0.000
	No	14.18 (513)	85.82 (3 104)		23.31 (843)	76.69 (2 774)		16.73 (605)	83.27 (3 012)	

υ with a lower dimension. As in our case study, where we introduce the latent variable υ "health status" we only consider a one-dimensional latent variable for simplicity. Extension to multi-dimensional latent variables and models with different types of observable responses are presented in Raach, 2005 and Khatab, 2007[12, 14].

LVMs in our study of a common latent variable υ to the linear predictor, is resulting in the observation models:

$$P(y_j=1 \mid \chi_j, \upsilon)=\Phi(\beta_{0j}+\chi_j'\beta_j+\lambda_j\upsilon) \quad (1)$$

Where $y_1,\ldots,y_p$ denote p observable binary responses, such as the three disease indicators in our study. Φ is the standard normal distribution function.

The covariates χ_j have direct effects β_j on the responses y_j, similar as in separate probit models. The common latent variable υ automatically induces correlation among the responses. The effects λ_j of υ are usually called factor loadings. The observation is supplement through a structural model for the latent variable. Geoadditive latent variable probit models assume a geoadditive structural model

$$\upsilon=u'\alpha+f_1(w_{i1})+\ldots+f(w_{iq})+f_{geo}(S_i)+\delta_i \quad (2)$$

with i.i.d. Gaussian errors $\delta_i \sim N(0,1)$. For identifiability reasons it is assumed that var $(\delta_i)=1$ and that the predictor for υ contains no intercept term. The additional covariates u, $w_1,\ldots,w_k$ and the location variable s act directly on the latent variable υ, but indirectly on the observable responses. The nonlinear functions and the spatial effect are modelled through similar priors as for separate geoadditive probit models, see aslo Khatab and Fahrmeir, 2009 and Khatab, 2010[15,16]. LVMs of the form (1) and (2), restricted to a linear structural model $\upsilon=u'\alpha$ have been suggested previously by Sammel *et al*[19].

We applied geoadditive probit LVMs to analyze the data. While the DIC is now commonly accepted as a standard tool for selecting probit or logit models, its performance for LVM model choice is not yet well understood. It was decided to choose the covariates used in equation (1) for the measurement model, which have direct effects on the disease indicators; or in the case of the structural equation (2), those have indirect effects via their common impact on the latent variable "health status", we therefore proceeded more informally: if the effects of covariates turned out to be significantly different (in terms of confidence intervals) for the three diseases, we decided to keep them in the measurement model, otherwise covariates were included in the geoadditive predictor of the structural equation for the latent variable. The results are presented in next section.

We present results for the following probit models, selected from a longer hierarchy of models. The responses y_j, j=1 (diarrhea), 2 (fever), 3 (cough) are coded as:

$$y_i=\begin{cases}1 & \textit{if child had disease 2 weeks prior to the survey}\\ 0 & \textit{if not}\end{cases}$$

In this section, we investigate how the three diseases can be interpreted as indicators of a latent variable υ ''health status'' of children, how much of the variation of υ can be explained through a geoadditive predictor, and which covariates have a direct effect on the disease indicators. This concept does automatically introduce a correlation among disease indicators. Further, it allows us to analyze the impact of covariates on health status. To demonstrate the latter property, we first consider a classic model without any covariates, *i.e.* in turns of auxiliary variables.

(LVM0):

$$P(y_{ij}=1 \mid \upsilon_i)=\Phi(\lambda_j\upsilon_i),\ \upsilon_i: N(0,1) \quad (3)$$

and $\eta=0$, so that υ_i: $N(0,1)$. Table 3 shows the estimates for the factor loadings λ_i, j=1,2,3 implying considerable (marginal) correlation.

Table 3
Results of model LVM0 with $\eta=0$ for Nigeria.

Parameter	Factor loadings	2.5%	97.5%
1. Fever λ_{1l}	1.48±0.120	1.30%	1.72%
2. Cough λ_{2l}	0.99±0.064	0.87%	1.12%
3. Diarrhea λ_{3l}	0.69±0.042	0.61%	0.77%

Factor loadings were expressed as Mean± SD.

Our next model was selected on the basis of the separate analyses as explained at Khatab, 2007[14]. In that work, we have used separate geoadditive probit models for the binary target variables for diarrhea, cough and fever using covariate information from the 2003 NDHS. The computations for the separate models were carried out using the Bayes X program[20]. We're showing here the results of Model 3 (Tables 1 - 3), which was selected from a long hierarchal analyses based on its DIC (the value of deviance information criterion).

This leads to the LVM

$$P(y_{ij} \mid x_{ij})=\Phi(\beta_{0j}+\alpha'_{ij}\beta_j+\lambda_j\upsilon_i), j=1,2,3 \quad (4)$$

with the structural model

$$\upsilon_i=u'\alpha+f_1(Chage_i)+f_2(BMI_i)+f_3(Mageb_i)+f_{geo}(reg_i)+\delta_i \quad (5)$$

for the latent variable. The vector α' (measurement model) comprises the covariates with direct effects (such as urban, availability of electricity, antenatal visits, working status and controlled water in LM1) on y_j, and u comprises the remaining categorical covariates (such as sex, mother's education, *etc.* in LM1) having common effects on the latent variable υ. Because the patterns for the nonparametric functions and the spatial effects were rather similar in the separate analyses, they were included in the geoadditive predictor for υ.

3. Results

The distribution of the categorical factors we used in this analysis and their association with the diseases were shown in Table 2. The following factors were significantly associated with diarrhoea: place of residence (P=0.001); mother's education (P<0.001); pregnancy treatment(P<0.001); availability of radio and electricity (P<0.001).

For fever, the significant factors were place of residence (P=0.001); pregnancy treatment (P<0.001); type of toilet (P< 0.001); mother's education (P=0.03); and availability of electricity (P=0.012). For childhood cough, significant factors were mother's working status (P=0.003), and pregnancy treatment (P<0.001).

Table 4

Results of LVM1 including direct and indirect effects for Nigeria.

Parameter		Mean±SD		2.5%	97.5%
Factor loadings	1. Fever λ_{11}	0.998*	0.084	0.845	1.162
	2. Cough λ_{21}	0.917*	0.072	0.770	1.063
	3. Diarrhea λ_{31}	0.753*	0.056	0.647	0.862
Parametric indirect effects	Male	0.049	0.068	−0.083	0.185
	Educ	0.0008	0.054	−0.104	0.110
	Toilet	−0.081	0.071	−0.219	0.058
	Radio	−0.028	0.040	−0.103	0.052
	Trepr	0.222*	0.076	0.070	0.370
Semi-parametric indirect effects	Chage	0.051*	0.050	0.0107	0.181
	BMI	0.005*	0.012	0.0003	0.032
	Mageb	0.003*	0.0048	0.0003	0.017
	reg	0.437*	0.156	0.211	0.812
Parametric direct effects	Urban (α_{11})	−0.224*	0.084	−0.387	−0.059
	Anvis (α_{12})	0.014	0.084	−0.150	0.180
	Elect (α_{13})	0.013	0.084	−0.150	0.177
	Work (α_{14})	0.030	0.075	−0.114	0.182
	Water (α_{15})	0.048	0.050	−0.052	0.146
	Urban (α_{21})	−0.114	0.083	−0.280	0.048
	Anvis (α_{22})	0.255*	0.083	0.092	0.416
	Elect (α_{23})	0.046	0.081	−0.115	0.209
	Work (α_{24})	0.161*	0.070	0.023	0.300
	Water (α_{25})	−0.032	0.049	−0.130	0.065
	Urban (α_{31})	−0.038	0.080	−0.199	0.119
	Anvis (α_{32})	−0.274*	0.076	−0.420	−0.122
	Elect (α_{33})	−0.032	0.079	−0.184	0.124
	Work (α_{34})	−0.033	0.067	−0.161	0.098
	Water (α_{35})	−0.037	0.049	−0.135	0.058

In Nigeria (Table 4), the parametric indirect covariate of *trepr* (whether the mother had treatment during pregnancy) had a significant positive effect on the latent variables. With regard to the parametric direct covariates, the results showed that urban location had a significant negative effect on an indicator of fever λ_{1l}, antenatal visits during pregnancy and current employment status of mother have significant effects on an indicator of cough (λ_{2l}). But, only the covariate of antenatal visits during pregnancy has a significant effect on the indicator of diarrhea (λ_{3l}) and these results are quite consistent with the previous separate analyses. As further analysis, we excluded the parametric indirect covariates, which were nonsignificant in the previous results and include covariates of urban, antenatal visits during pregnancy and current employment status of mother as indirect effects. The results of LVM2 (Table 5) show that all the parametric indirect covariates have a significant effect on the latent variable "health status" of children.

Table 5

Results of LVM2 including direct and indirect effects for Nigeria.

Parameter		Mean	SD	2.5%	97.5%
Factor loadings	1. Fever	1.014*	0.078	0.871	1.189
	2. Cough	0.945*	0.076	0.808	1.112
	3. Diarrhea	0.730*	0.056	0.621	0.842
Parametric indirect effects	Trepr	0.217*	0.071	0.010	0.1623
	Work	0.054*	0.077	0.0004	0.024
	Urban	−0.154*	0.072	0.0003	0.014
	Anvis	0.0026*	0.086	0.205	0.770
Semi-parametric indirect effects	Chage	0.048*	0.0438	0.010	0.162
	BMI	0.0049*	0.007	0.0004	0.0244
	Mageb	0.0028*	0.004	0.0003	0.0145
	reg	0.421*	0.144	0.205	0.770
Parametric sirect effects	Elect (α_{11})	0.007	0.087	−0.168	0.174
	Water (α_{12})	0.053	0.051	−0.045	0.155
	Male (α_{13})	0.035	0.067	−0.091	0.167
	Educ (α_{14})	0.022	0.058	−0.093	0.138
	Toilet (α_{15})	−0.194*	0.066	−0.327	−0.065
	Radio (α_{16})	−0.0140	0.041	−0.093	0.071
	Elect (α_{21})	0.0371	0.082	−0.120	0.201
	Water (α_{22})	−0.035	0.050	−0.131	0.0654
	Male (α_{23})	0.0013	0.066	−0.125	0.133
	Educ (α_{24})	0.045	0.055	−0.063	0.154
	Toilet (α_{25})	0.081	0.064	−0.040	0.210
	Radio (α_{26})	0.009	0.040	−0.068	0.0915
	Elect (α_{31})	−0.0124*	0.077	−0.164	−0.062
	Water (α_{32})	−0.024	0.047	−0.115	0.066
	Male (α_{33})	0.114	0.062	−0.004	0.235
	Educ (α_{34})	−0.098	0.055	−0.206	0.0095
	Toilet (α_{35})	−0.120	0.063	−0.248	0.0024
	Radio (α_{36})	−0.075	0.037	−0.148	0.0045

With regards to non-linear effects (Figure 1) on the latent variable health status, the pattern showed that the prevalence of disease was high among children 6-12 months of age. Mother's age seemed to have a slight effect on the child's health in Nigeria. Whilst there was a signifcant association between the child's health and a mother with an BMI less than 22.

Figure 2 displayed the results of spatial effects in model LVM1 and LVM2. This suggested that the high risk of all three health status rates was associated with the northeastern

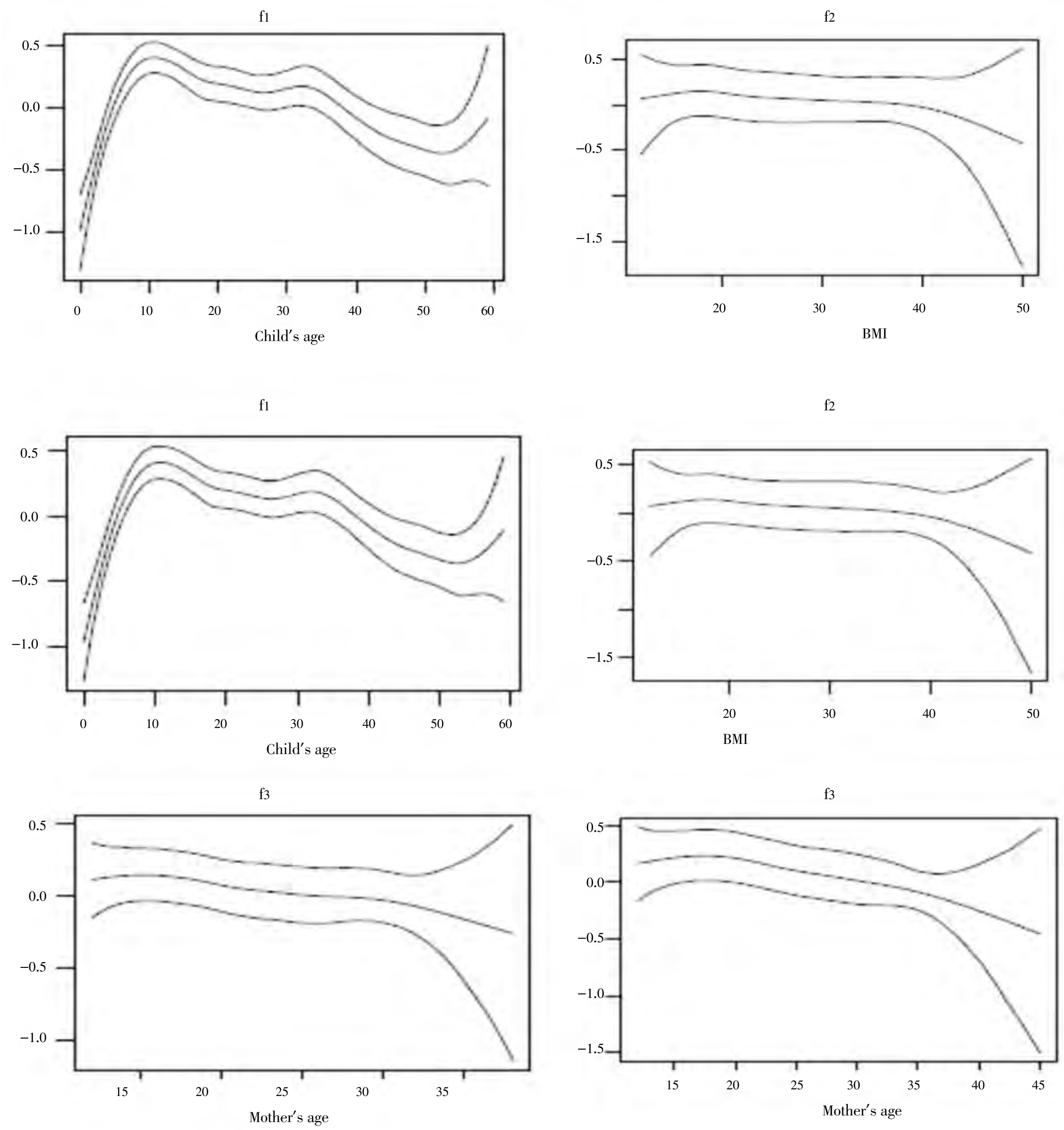

Figure 1. Non-linear effects from top to bottom: child's age, mother's BMI and mother's age at birth (for model LVM1), child's age, mother's BMI and mother's age at birth (for model LVM2) on the indicators of a latent variable ''health status'' of children disease for Nigeria using Bayesian LVM for binary responses.

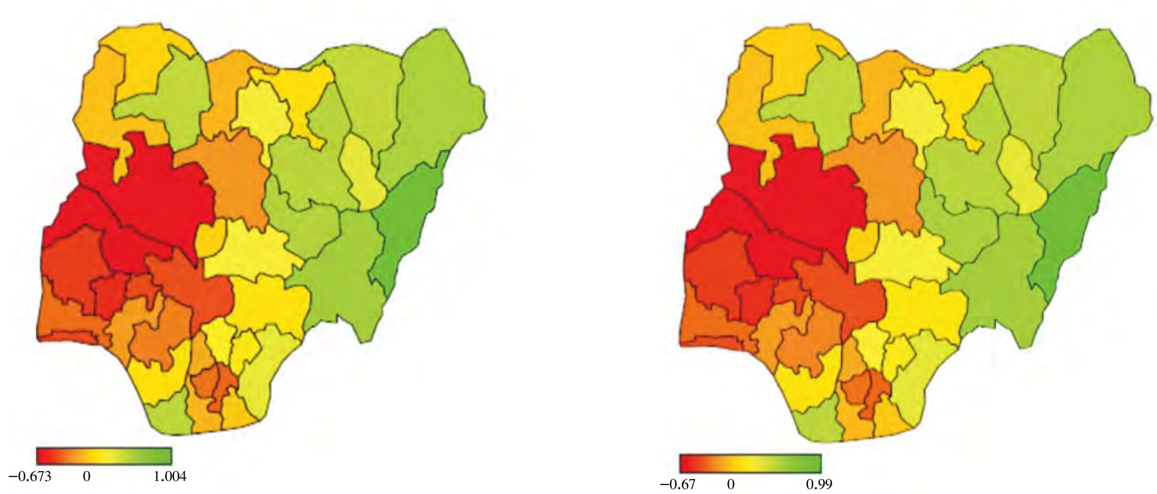

Figure 2. Posterior mean LVM for LVM1 (left panel) and LVM2 (right panel) on diseases in Nigeria.

Table 6
Fixed effects of separate model 3 using Bayesian geoadditive model on diarrhea, fever, cough–Nigeria.

Variable	Diarrhea					Fever					Cough				
	Mean	SD	2.5%	Median	97.5%	Mean	SD	2.5%	Median	97.5%	Mean	SD	2.5%	Median	97.5%
Const	−1.320*	0.165	−1.650	−1.320	−1.012	−0.600*	0.125	−0.844	−0.605	−0.363	−0.791*	0.143	−1.076	−0.793	−0.513
Male	0.047	0.026	−0.002	0.046	0.101	0.017	0.023	−0.028	0.017	0.064	0.0025	0.023	−0.048	0.002	0.049
Urban	−0.052	0.034	−0.119	−0.054	0.018	−0.048	0.030	−0.106	−0.047	0.010	−0.060	0.031	−0.119	−0.062	0.0031
Work	0.0145	0.028	−0.040	0.014	0.069	0.025	0.026	−0.025	0.025	0.079	0.015	0.027	−0.039	0.014	0.065
Trepr	0.035	0.034	−0.032	0.035	0.100	0.071*	0.030	0.010	0.072	0.130	0.105*	0.032	0.045	0.105	0.170
Anvis	−0.064*	0.033	−0.128	−0.065	−0.0009	0.018	0.030	−0.041	0.016	0.080	0.033	0.033	−0.030	0.034	0.102
Radio	−0.039	0.032	−0.103	−0.039	0.025	−0.029	0.031	−0.092	−0.029	0.026	0.003	0.030	−0.050	0.0012	0.067
Elect	0.0168	0.033	−0.046	0.0159	0.079	−0.019	0.031	−0.075	−0.020	0.0418	0.002	0.031	−0.060	0.0033	0.062

part of Nigeria as already indicated by the previous separate results with geoadditive probit models[14].

As presented in this paper, LVMs were based on the results of the separate models which presented in Khatab, 2007. In that stage, we have used separate geoadditive probit models the binary target variables for diarrhea, cough, and fever using covariate information from the 2003NDHS. The computations for the separate models were carried out using BayesX program (Brezger A, Kneib T, Lang S, 2005)[20]. Results of Model 3 was shown in Table 6, which was selected from a long hierarchal analyses based on its DIC (the value of deviance information criterion).

4. Discussion

4.1. Fixed effects

As for child's gender, it is widely believed that probability of disease is higher for males due to biological reasons. However, some studies show higher female mortality, indicating gender discrimination. The variable is insignificant for our three types of diseases.

The effects of urban versus rural place of residence has a significant effect on the health status in Nigeria. This refers to the fact that living in urban areas lowers the risk for children of having diseases *vs.* rural areas. These results support the important role of the public health policy in rural–urban disparities.

Mothers who attended a clinic to receive antenatal care during the period of pregnancy are expected to have lower problems in comparison to those who had not received any care.

The variable *trepr* has a positive significant effect on the three types of disease in Nigeria.

The ownership of radio facilitates the acquisition of disease and vaccination information, allowing a more effective allocation of resources to produce child health. However, in Nigeria it has non–significant effect.

Concerning current working status of mother, it has an association with child's health in Nigeria. The problem as mentioned in some studies is that when mothers engage in out–of–home employment, it curtails the duration of full breastfeeding and necessitates recently introduced supplementary feeding, often by the illiterate care–takers, and that could have an effect on the health of child in the early months.

Availability of the flush toilet in the household is associated only with a lower risk of fever. The same for availability of electricity with diarrhea.

4.2. Non–linear effects

In general, the results show that the risk of having diseases in the two–week reference period reaches its peak at 11 months and then begins to fall with increasing age of the child. This pattern resembles those found in many studies of Sub–Saharan Africa. The prevalence of disease was found to be the highest among children 6–12 months of age, the period when most children are weaned. In addition to breast milk, inborn immunity and less exposure to contaminated agents during the early period also contributes to the lower prevalence of diarrhea. On the other hand, prevalence is quite high when the child has lost it's inborn immunity and when it is exposed to different types of infections by eating food prepared with contaminated water and from an unhealthy environment.

The results reflect a slight effect of mother's age at birth on the morbidity of children. In Nigeria, the impact of mother's age is slight and almost linear.

In the literature, the influence of the Body Mass Index (BMI) of the mother is sometimes expected to be inversely U–shaped. Parents with low BMI values are malnourished and are therefore likely to have undernourished and weak children. At the same time, very high BMI values indicate poor quality of the food and hence, may also imply weakness of the children in our study. It has a slight effect on cough and fever.

4.2. Spatial effects

In Nigeria, there is a sizeable difference between disease in the eastern parts of the country and the significantly better health status in the northern, and central parts. We can see from the results that southeastern regions through some regions in the north part are associated with a high rate of childhood disease. That is because, as suggested by previous studies, in those regions a high level of pollution due to petroleum production is present. Perhaps, the pollution in this area affect the health of children through the water pollution that influences access to drinkable water sanitation.

Our results also have some policy relevant implications. It is recommended that mothers exclusively breastfeed their babies for the first 2 years of life. In developing countries, continued breastfeeding is recommended up to at least 2 years of age with the timely addition of appropriate complementary food at 6 months of age.

The spatial effect suggests the need to give more attention to some areas that have high rates of diseases, such as the southeastern regions as well as some regions in the north part which are associated with a high rate of diseases. These areas are more likely to have a higher proportion of morbidity compared to other areas, caused by poor health facilities and complications during childbirth or even careless and misdiagnosis during hospital care. Therefore, the most important issues to address in these areas are health care, proper food, and raising the educational level of parents.

The spatial effects also suggest that further work is needed in order to assess the association between the morbidity and the malnutrition among the Nigerian children.

Comorbidity with conditions such as cough, diarrhoea and fever is common in Nigeria. However, little is known about common risk factors and geographical overlaps in these illnesses. The search for overlapped common risk factors and the spatial effects may improve our understanding of the etiology of the diseases for efficient and cost-effective control and planning.

The analysis suggests some strong underlying spatial patterns of the three ailments with a clear southeastern divide of childhood morbidities and this might result in shared and overlapping risk factors, one of which is probably general environmental risk associated with pollution.

This study is able to jointly analyse childhood fever, cough and diarrhoea in Nigeria by highlighting a clear spatial patterns of these illnesses. We postulate that LVM offer a new methodology opportunity for modelling joint diseases risks factors, in addition to flexibly modelling covariates and the spatial effects.

Conflict of interest statement

We declare that we have no conflict of interest.

References

[1] UNICEF. *The state of the world children*. New York: Editorial, Design and Publications Section, UNICEF; 2008.

[2] Survey DaH. *Demographic and health survey for Nigeria*. Nigeria: National Population Commission. 2003. [Online] Available form: http://www.measuredhs.com/pubs/pdf/FR148/00FrontMatter.pdf.

[3] Oyedeji GA, Olamijulo SK, Osinaike AI, Esimai VC, Odunusi EO, Aladekomo TA. Anthropometric measurement in children aged 0-6 years in a Nigerian village. *East Afr Med J* 1995; **72**: 523-526.

[4] Vaahtera M, Kulmala T, Maleta K, Cullinan T, Salin ML, Ashorn P. Epidemiology and predictors of infant morbidity in rural Malawi. *Paediatr Perinat Epidemiol* 2000; **14**: 363-371.

[5] Yohannes AG, Streatfield K, Bost L. Child morbidity patterns in Ethiopia. *J Biosoc Sci* 1992; **24**: 143-155.

[6] Kandala NB, Ji C, Stallard N, Stranges S, Cappuccio FP. Spatial analysis of risk factors for childhood morbidity in Nigeria. *Am J Trop Med Hyg* 2007; **77**: 770-779.

[7] Fahrmeir L, Lang S. Bayesian inference for generalized additive mixed models based on Markov Random Field Priors. *Applied Statistics* 2001; **50**: 201-220.

[8] Fahrmeir L, Kneib T, Lang S. Penalized structured additive regression of space-time data: a Bayesian perspective. *Statistica Sinica* 2004; **14**: 731-761.

[9] Kandala NB, Magadi MA, Madise NJ. An investigation of district spatial variations of childhood diarrhoea and fever morbidity in Malawi. *Soc Sci Med* 2006; **62**: 1138-1152.

[10] Adebayo SB, Fahrmeir L. Analysing child mortality in Nigeria with geoadditive discrete-time survival models. *Stat Med* 2005; **24**: 709-728.

[11] Kandala NB, Gebrenegus Ghilagaber. A geo-addictive bayesian discretetime survival model and its application to spatial analysis of childhood mortality in malawi. *Quality And Quantity* 2006; **40**: 935-957.

[12] Raach A. *A Bayesian semiparametric latent variable model for binary, ordinal and continuous response*. Munchen: Dissertation. 2005.

[13] Fahrmeir L, Raach A. A Bayesian semiparametric latent variable model for mixed responses. *Psychometrika* 2007; **72**: 327-346.

[14] Khatab K. Analysis of childhood diseases and malnutrition in developing countries of Africa. Dr Hut Verlag-Munich, Germany 2007.

[15] Khatab K, Fahrmeir L. Analysis of childhood morbidity with geoadditive probit and latent variable model: a case study for Egypt. *Am J Trop Med Hyg* 2009; **81**: 116-128.

[16] Khatab K. Childhood malnutrition in Egypt using geoadditive Gaussian and latent variable models. *Am J Trop Med Hyg*; **82**: 653-663.

[17] Kazembe LN, Muula AS, Simoonga C. Joint spatial modelling of common morbidities of childhood fever and diarrhoea in Malawi. *Health Place* 2009; **15**: 165-172.

[18] Kandala NB, Emina JB, Nzita PD, Cappuccio FP. Diarrhoea, acute respiratory infection, and fever among children in the Democratic Republic of Congo. *Soc Sci Med* 2009; **68**: 1728-1736.

[19] Sammel MD, Ryan LM, Legler JM. Latent variable models for mixed discrete and continuous outcomes. *J Roy Statistical Society* 1997; **197**: 667-678.

[20] Brezger AKT, Lang S. BayesX-Software for bayesian inference based on Markov Chain Monte Carlo simulation techniques. [Online] Available from: http://www.statuni-uenchende/lang/BayesX 2005;1.4. [Accessed on 8th Aug, 2010]