Bayesian Modelling of the Unobserved and Observed Factors affecting Under-five Mortality Rate in Zambia

38th Annual Conference of the International Society for Clinical Biostatistics, Students' Day, VIGO –SPAIN

Isaac Fwemba;
Department of Biostatistics-SPH/Univ. of Ghana, 2:I-BIOST, LEUVEN University
Department of Epidemiology & BIOST-SPH-UNZA
Department of Statistics-Univ. of Ghana

13 July, 2017

•

Outline

What is known in the SSA Region?

In the past 15 years, Governments have implemented various interventions in

Maternal Mortality

Child Mortality

In SSA, child mortality rate varies considerably

Ghana=90; Zambia=83; Rwanda=54 ,Chad=147.50 per 1000 live births

Challenges & Data structure

Existing Challenges/Problem

- Figures still high despite 15 years serious investment from
  - WHO, UN agencies
  - National Govts.
- Could there be a problem with implementation strategies?
  - If that is not known
    - what will happen to SDGs?
    - risk effects not well understood?

Data Structure

- Hierarchically arranged data

  - Regional factors (Kazembe,2014)
  - Clustering effect/Community
    - Environmental factors
    - Sanitation and water sources
  - Clustering effect/Family
    - Shared genetic factors

Standard Approaches & Challenges

Known Approaches

- Logistic Regression
  - Subject state overtime
- Basic Survival fun
  - Hazard of the child
  - independence across subject
  - Proportionality
- Life table Approach

Appropriate Approaches

- GLMM
  - Include Random effect
- GAM
  - Semi parametric
  - Nonparametric approach
- GAMM
  - Random effect
  - Semi parametric

Challenges with standard methods

- Correlational structure is neglected
  - S.E Underestimate Inflate of sig
- Fail to capture risk of dying

Fitting ZDHS Dataset

Model(1): PH Weibull

$h_i(t) \quad h_0(t)\exp(\ _i), t \quad 0$

$h_0(t) \sim \quad t \quad ^{1}, \quad 0$

$l_i \quad \ _i \log h(t_i) \quad \int_0^t h(u)du$

$\ _i (\log \quad (\quad 1) \log t_i \quad \ _i) \quad \exp(\ _i) t_i$

$\ _i \quad ^T z_i$

We assign Gaussian priors to all the elements

$b_i \quad \log(w_i) \square \ N(0, \quad ^{1})$

$\sim (0, 10^4 I)$

$\ _0 \sim N(0, 0.001 \ ^{1})$

$\sim \quad (10 \ ^3, 10 \ ^3)$

$\sim \quad (0, 10 \ ^3)$

Semi parametric baseline hazard

Model(2), PHM:Piecewise Cox Model

$h_0(t) \qquad {}_k$ $\qquad t \quad (s_{k\ 1}, s_k), k \quad 1,......k., t_i, \quad {}_i, i \quad 1,......., m$

$h_i(t) \quad h_0(t)\exp(\ {}^T z_i) \quad \exp(\ {}^T z_i \quad w_i \quad u_k) \quad \exp(\ {}_{ik})$ $\qquad t \qquad (s_{k\ \ 1,} s_k)$

$b_k \quad \log(\ {}_k)$ $\qquad l \quad \log h(t) \quad {}_0^t h(u)du \quad {}_k \ (t \quad s_k)e^{\ k} \quad {}_{j\ 1}^{k\ 1} (s_{j\ 1} \quad s_j)e^{\ j}$

Prior Specification

$b_{k\ \ 1,} \quad b \sim N(0, t^{\ 1}), k \quad 1..k \quad 1, j \quad 1,...j$

$\sim N_q(0, \ )$ $\qquad W_i \ \& \ {}_k \sim \log Normal(\ ,\ )$ $\qquad t \sim \quad (0.001, 0.001)$

Specification of Model

Model1 $: \;{}_{1} \qquad {}_{0} \qquad {}_{1}.Sex \qquad {}_{2}.Educ \qquad {}_{3}.BORD \qquad {}_{4}\Pr ec \qquad .Multiple$

Model2 $: \;{}_{2} \qquad {}_{0} \qquad {}_{1}.Sex \qquad {}_{2}.Educ \qquad {}_{3}.BORD \qquad {}_{4}\Pr ec \qquad .Multiple$

Model3 $: \;{}_{3} \qquad {}_{0} \qquad {}_{1}.Sex \qquad {}_{2}.Educ \qquad {}_{3}.BORD \qquad {}_{4}\Pr ec \qquad .Multiple \quad \log(w_i)$

Model4 $: \;{}_{4} \qquad {}_{0} \qquad {}_{1}.Sex \qquad {}_{2}.Educ \qquad {}_{3}.BORD \qquad {}_{4}\Pr ec \qquad .Multiple \quad \log(u_i)$

Fitting models Using Laplace Approximation(RUE, 2009)

Latent Gaussian Models $x \mid \ _i \quad N(0, Q^{\ 1}(\ _1)).$

Hyper parameter $\ _i \quad \{\ _1, \ _2\}.$

Likelihood for ith observation $(y_i \mid \ _i, \ _2)$

Structure predictor $\ _i$ accounts for various covariates in additive way $\ _i \quad \ _0 \quad \sum_{j\ 1}^{nf} w_{ij} f^{(j)}(\ _{ij}) \quad \sum_{k\ 1}^{n} \ _{ki} z_{ki} \quad \ _i$

Linear effect covariates, group specific frailty, weights for each observation $\ _k, \ f^{(j)}(.), \ w_{ij}$

Unstructured random effects $\ _i$

Posterior distribution reads as $(x, \ \mid y) \quad (\ ) \quad \prod_{i \ x} (y_i \mid x, \ )$

Model Assessment

No	Variance	DICs	pD	WAIC	pD
M1: PH Cox Piecewise	-	7385254	3678758	4.33	2.16E+19
M2: PH Weibull	-	6051808	3012049	2.29	1.15E+19
M3: PH Weibull +Family frailty	1.01	6902262	3435455	2.20	1.10E+19
M4: PH Weibull +Community frailty	7.40	8013655	3989473	1.17	5.83E+19
M5: PH Weibull+ community + family frailty	15.02	8995584	4481655	4.98	2.49E+19

$$DIC \qquad D \qquad pD$$

Better model with small DIC=Weibull & family frailty

Variation =1/standard deviation

Community level=1/7.40=13% **Family=1/1.013=98%**

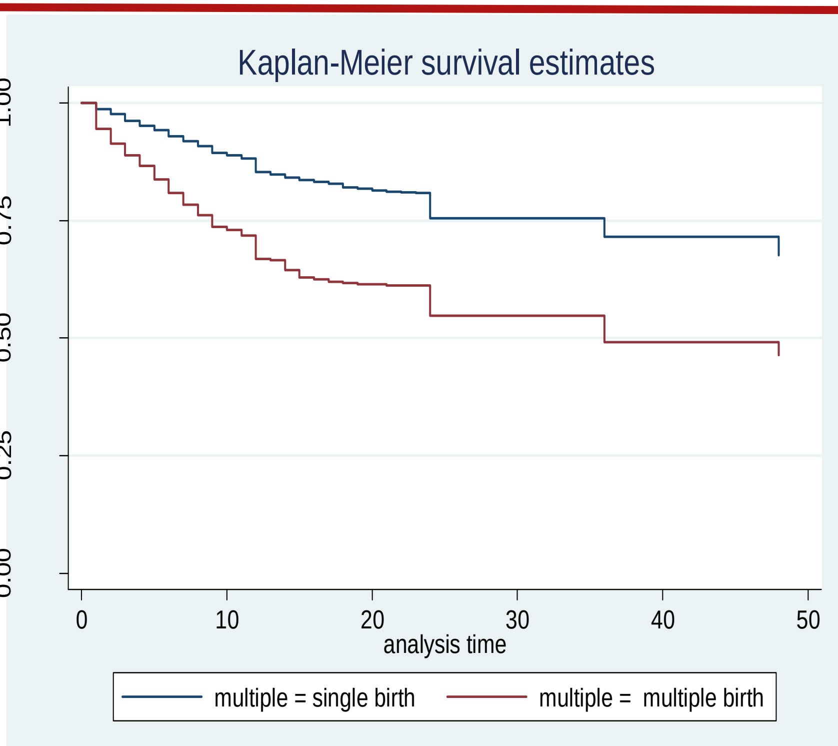
Selected Results

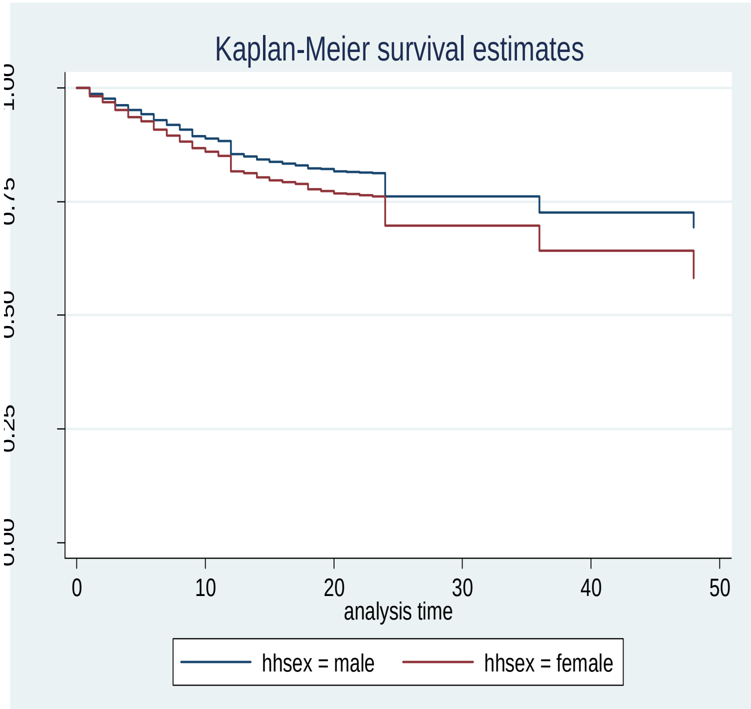
Provinces	Deaths
central	409(8.07)
Copper belt	407(8.03)
Eastern	875(17.36)
Luapula	630(12.43)
Lusaka	332(6.55)
Muchinga	517(10.20)
Northern	698(13.77)
North Western	406(8.01)
Southern	423(8.34)
Western	372(7.34)

Mother’s Education	Education
No Education	3243(64.01)
Primary	858(16.94)
Secondary	783(15.48)
Higher	182(3.59)

- Varying mortality rate across provinces Presence
  - Education level affected mortality rates too
- Standard errors were consistently small for random effect model compared to standard Weibull

Description of under-five child survival





- Singleton births survived more than multiple births of children
- Male children had higher survival probability than females

Implication of the findings

- Results consistent with previous studies in the region (Uganda(Naseji,2014);Malawi(Manda,2001;
- Ghana,(2014)

- Variation at family level suggest that there are other factors unexplained by the covariates in the model

Conclusions

- Clear evidence of heterogeneity
  - Family and Community Level
- Family variation showed more variation
  - Indicating that interventions should emphasize family structure
- It is clear that interventions should focus on
  - Households rather than communal
- Our findings also suggest that using robust advanced statistical analysis methods
  - May provide further insight into risk factors for mortality

Key messages

What are the disadvantages of using INLA compared to MCMC approach

What do you do when your DIC significantly differs from WAIC

How do you handle missing data in complex surveys such as DHS

What are the challenges encountered when analyzing complex survey studies

What do you do when the data does not fit the assumption

Thank You !!
GRACIAS

Reference continued

Martino, S., Akerkar, R., & Rue, H. (2011). Approximate Bayesian inference for survival models. *Scand J Stat*, *38*.

Nasejje, J. B., Mwambi, H. G., & Achia, T. N. O. (2015). Understanding the determinants of under-five child mortality in Uganda including the estimation of unobserved household and community effects using both frequentist and Bayesian survival analysis approaches. *BMC Public Health*, *15*(1), 1003. http://doi.org/10.1186/s12889-015-2332-y

Simpson D, Lindgren F, R. H. (2011). Simpson D, Lindgren F, Rue H: Fast approximate inference with INLA: the past, the present and the future. Technical Report at arxiv.org, Department of Mathematical Sciences, Norwegian University of Science and Technology (NTNU), Trondheim, Norway 2011.