Bayesian approach in Estimating Risk Determinants of Infectious diseases

Thembile Mzolo

University of KwaZulu Natal & Human Sciences Research Council

SUSAN Conference - Kabarak University, August 2009

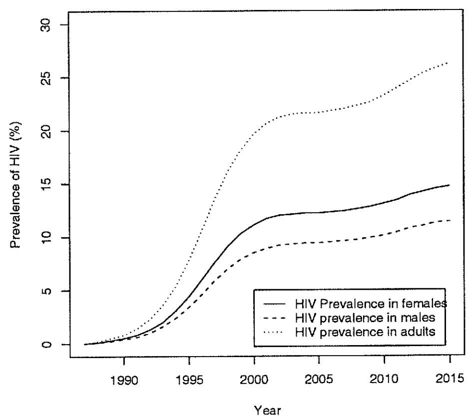
Outline

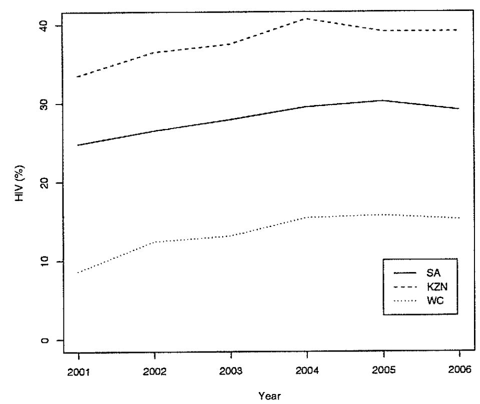
1. Introduction
   - Human Immune Virus and Tuberculosis
   - Objectives of the study
2. Methodology
   - Data description
   - Bayesian Hierarchical models
3. Analysis of the data
   - Results
4. Discussion
5. Conclusion

HIV and TB

- People who are co-infected with both HIV and TB are at increased risk of dying from TB (Bucher 1999; Corbett 2003) than HIV/AIDS.
- The burden of TB in countries with high rates of HIV has increased rapidly over the past decade.
- Life expectancy has dropped to 49 years for males and 53 years for females (Dorrington 2006).
- SA has the highest number of people living with HIV/AIDS in the world (UNAIDS 2008).
- Infection levels vary among provinces and by gender.

HIV status by gender and province





Incidence estimates of TB in South Africa (Weyer 2004)

Province	Incidence	Total cases	% of TB cases	% of HIV cases
Western Cape	1333	58577	2.28	50.4
Northern Cape	822	8033	10.4	52.0
Eastern Cape	1307	102152	1.3	58.8
KwaZulu Natal	1696	173944	0.9	83.4
Limpopo	647	41108	1.6	55.1
Mpumalanga	1052	35977	2.9	77.9
Free State	871	29790	2.9	70.5
Gauteng	1034	85855	1.2	63.6
North West	754	29472	2.6	64.3
SOUTH AFRICA*	1084	529320	2.9	66.4

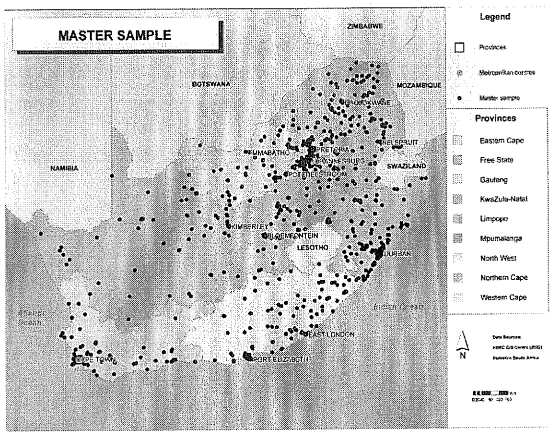
* *The totals of South Africa are calculated from incidence rounded to nearest full digit, therefore total differs from sum of provincial totals.*

Objectives

In South Africa, no coherent analysis of the determinants of HIV and TB has been conducted at a national level and thus this study seeks to mend this gab.

- Determine the risk determinants of HIV and those of TB.
- Incorporate heterogeneity between geographic areas in modelling by means of random effects models and Bayesian hierarchical modelling approaches.

Household based surveillance survey of HIV, HSRC in 2005



- Multi-stage disproportionate, stratified random sampling approach.
- Sample weights were introduced.
- 10 584 households participated in the study.
- 15 581 people aged 2^+ gave DBS for HIV testing.
- 16398 adults (15^+) participated in the study.

Bayesian Hierarchical models

- Our dataset is clustered at an EA level.
- Controlling for both fixed and random risk factors will help in quantifying any excess association between HIV & TB within an EA.
- Bayesian methods require prior information to estimate the posterior distribution.
- These methods involve integrating high-dimensional functions.
- Our focus is on the MCMC methods of simulating data.
- The roots of the MCMC methods come from the Metropolis Algorithm (Metropolis & Ulam 1949; Metroplis 1953).

Bayesian *cont.*

Gibbs sampling

- The Gibbs sampler (Geman & Geman 1984) is a MCMC method that is widely applicable which avoids the need of evaluating high dimensional integrals.
- In the Gibbs sampler one needs only to consider the univariate conditional distributions.
- These conditional distributions have simple forms and are easier to simulate than complex joint distributions.
- Consider a bivariate random variable (x, y) & suppose we wish to compute one or both marginals, $p(x)$ & $p(y)$.

Bayesian *cont.*
Gibbs sampling

- The sampler starts with some initial value y_0 for y & x_0 for x by generating random variable from the conditional distribution $p(x|y = y_0)$.
- The sampler uses x_0 to generate a new value of y_1. The sampler proceeds as follows:

$$x_i \sim p(x|y = y_{i-1})$$
$$y_i \sim p(y|x = x_i)$$

- This process is repeated k times, generating a Gibbs sampler of length k.

Bayesian *cont.*

Gibbs sampling

- In order to get the desired total m sample points, one samples the chain:
  - after a sufficient burn-in to remove the effects of the initial sampling values.
  - at a set time points (say, every n samples) following the burn-in.
- The Gibbs sampler converges to a stationary distribution which is a target distribution we are trying to simulate from (Tierney 1994).
- There are stationarity formal tests that can be used, e.g Geyer (1992), Geweke (1992), Raftery and Lewis (1992b), etc.

HIV Results

- For HIV model the ff. variables were used: sex, age, race, education, health & condom use at sexual debut.
- For TB model the ff. variables were used: HIV status, sex, education, income and health status.
- Priors for fixed effects were assumed multivariate normal centered at zero.
- Priors for random effects (EA) were assumed to follow a normal distribution.
- An estimated intraclass correlations for HIV and TB are $\rho_{HIV} = 0.169$ and $\rho_{TB} = 0.249$, respectively.

HIV model: Odds Ratios and Confidence intervals

Parameter	Bayesian results OR (95% CI)	GLMMs results OR (95% CI)
Sex of the respondent		
Male	0.686 (0.596 - 0.791)	0.708 (0.612 - 0.818)
Female	1	1
Age of respondent		
15 to 24	6.43 (4.697 - 8.802)	6.203 (4.393 - 8.758)
25 to 34	11.45 (8.415 - 15.580)	10.79 (7.768 - 14.999)
35 to 44	7.29 (5.425 - 9.806)	6.979 (5.053 - 9.641)
45 to 54	2.846 (2.0647 - 3.924)	2.801 (1.972 - 3.979)
55+	1	1
Race		
African	17.409 (8.594 - 35.269)	7.869 (4.023 - 15.394)
White	1.802* (0.736 - 4.415)	1.366* (0.546 - 3.418)
Coloured	2.945 (1.415 - 6.129)	2.445 (1.199 - 5.452)
Indian	1	1
Education		
None/primary	2.232 (1.619 - 3.077)	2.115 (1.501 - 2.980)
Secondary/matric	2.057 (1.514 - 2.793)	2.0442 (1.471 - 2.841)
Tertiary	1	1
Health status		
Good	0.536 (0.456 - 0.631)	0.525 (0.443 - 0.621)
Poor	1	1
Condom use at sexual debut		
Yes	0.69 (0.576 - 0.849)	0.698 (0.569 - 0.856)
No	1	1

* p-value>0.05

TB model: Odds Ratios and Confidence intervals

Parameter	Bayesian results OR (95% CI)	GLMMs results OR (95% CI)
HIV status		
Negative	0.239 (0.173 - 0.332)	0.236 (0.168 - 0.330)
Positive	1	1
Sex of the respondent		
Male	2.309 (1.684 - 3.168)	2.370 (1.725 - 3.258)
Female	1	1
Education		
None/primary	9.0702 (1.902 - 43.250)	7.841 (1.670 - 36.782)
Secondary/matric	4.637* (0.974 - 22.0652)	4.007* (0.854 - 18.803)
Tertiary	1	1
Income		
Unemployed	1.008* (0.747 - 1.584)	1.207* (0.846 - 1.721)
Employed	0.454 (0.278 - 0.743)	0.565 (0.351 - 0.909)
other	1	1
Health status		
Good	0.202 (0.147 - 0.278)	0.187 (0.135 - 0.259)
Poor	1	1

* p-value>0.05

Discussion

- In SA women are less empowered to negotiate safe sex with their partners (Worth *et al*. 1990; Stein *et al*. 1990).
- So far it has been reported that HIV is high among women than men based on antenatal clinic attendee studies and other studies (DOH 2007; Shisana *et al*. 2003) & confirmed in these results.
- Individuals in age group 25 to 34 engage in highly connected sexual networks.
- Africans are mostly found in informal areas where the socio-economic impact plays a huge role in the spread of HIV.
- Use of condom at sexual debut is an indication of being well informed about risk of HIV.

Discussion *cont*

- TB is one of the leading opportunistic infections in HIV infected individuals.
- Individuals infected with TB are likely to be infected with HIV due to the compromised immune system.
- Male individuals are likely to be exposed to poor working environments than females.
- While those who are educated have enough income to take good care of their health and less likely to contract TB.

Conclusion

- Results from the Bayesian and the GLMMs are quite comparable & this comparability is re-assuring.
- Inference drawn from the two modelling approaches provides some degree of confidence in the results.
- The positive correlation observed at an EA level for both HIV (ρ_{HIV}) and TB (ρ_{TB}) indicates that interventions should be aimed at an area level rather than only the individuals.
- Studies that intervene at community level. e.g Mwanza Trial (Grosskurth, 1995) should be encouraged to fight epidemics of diseases such as HIV and TB.
- Spatial modeling of these data can help map areas that are prone to HIV & TB.

Acknowledgements

- SUSAN for making sure that I attend this conference.
- SACEMA, for funding my MSc studies.
- HSRC, for allowing me to use their rich dataset.
- Dr Henry Mwambi, UKZN, supervisor.
- Dr Khangelani Zuma, HSRC, co-supervisor.
- Dr Samuel Manda, MRC, for helping with WinBugs software.

Thank You