



**Comparison of Geographic Distribution and Correlates of Adult Mortality in South Africa using Civil Registration System and Census Datasets, 2011**

**Samuel Manda <sup>1,2</sup>, Thandi Kapwata<sup>1</sup>, Henry Mwambi<sup>2</sup>, Nada Abdelatif<sup>1</sup>**

**<sup>1</sup>South African Medical Research Council, South Africa**

**<sup>2</sup>University of KwaZulu-Natal, School of Mathematics, Statistics and Computer Science, PMB Campus, Pietermaritzburg, South Africa.**

**Presentation at International Association of Official Statistics (IAOS) Conference**

**Abu Dhabi, United Arab Emirates**

**6-8 December, 2016**

# Outline

- Data Sources Mortality in South Africa
- Census and Vital Registration
- Multivariate Mapping of the Mortality Data in South Africa
- Results
- Discussions and conclusions

## MORTALITY DATA IN SOUTH AFRICA

- South Africa has an enormous and wide-ranging mortality data sources that provide a huge potential for comparing results from different sources
- An in-depth of mortality data sources in the post-democracy South Africa is presented in Joubert, et al. (2012).
- Estimation method (direct or indirect) of mortality depends largely on data source and availability
- Uses and application of mapping modelling to mortality data have appeared (see, a review by Manda, Ngandu & Abdelatif, 2016)

# Principle Mortality Data Sources: Joubert, et al. (2012).

- **Large-scale nationally representative surveys may collect mortality data**
  - **South African Demographic and Health Surveys (1998, 2003, 2016), South Africa Community Survey (2006, 2016)**
- **Vital registration system-based mortality and causes of death.**
  - **Collected using death notifications by the Department of Home Affairs (DHA)**
  - **Statistics South Africa (StatsSA) compiles yearly reports of cause of death statistics**
  - **The best and most-reliable source for mortality data in the country, with very high level of completeness (Groenewald et al, 2012; Bradshaw et al, 2012).**
- **Census mortality data (1996, 2001, 2011)**
  - **Census mortality can be computed at lower geographical areas in the absence of complete vital registration system.**

## Death Totals for Adult (age 15-64 years), 2011

- A total of just over 280545 deaths were reported using Census 2011; 151970 male deaths; 128575 female deaths
- The number of notified and registered deaths for 2011 were about 297,039, of which 165,490 and 131,549 were male and female deaths, respectively
- This age group was chosen mostly affected by the HIV epidemic and emerging NCD morbidity and mortality in South Africa.

# ADULT MORTALITY STATISTICS AND CONTEXTUAL DATA, SOUTH AFRICA 2011

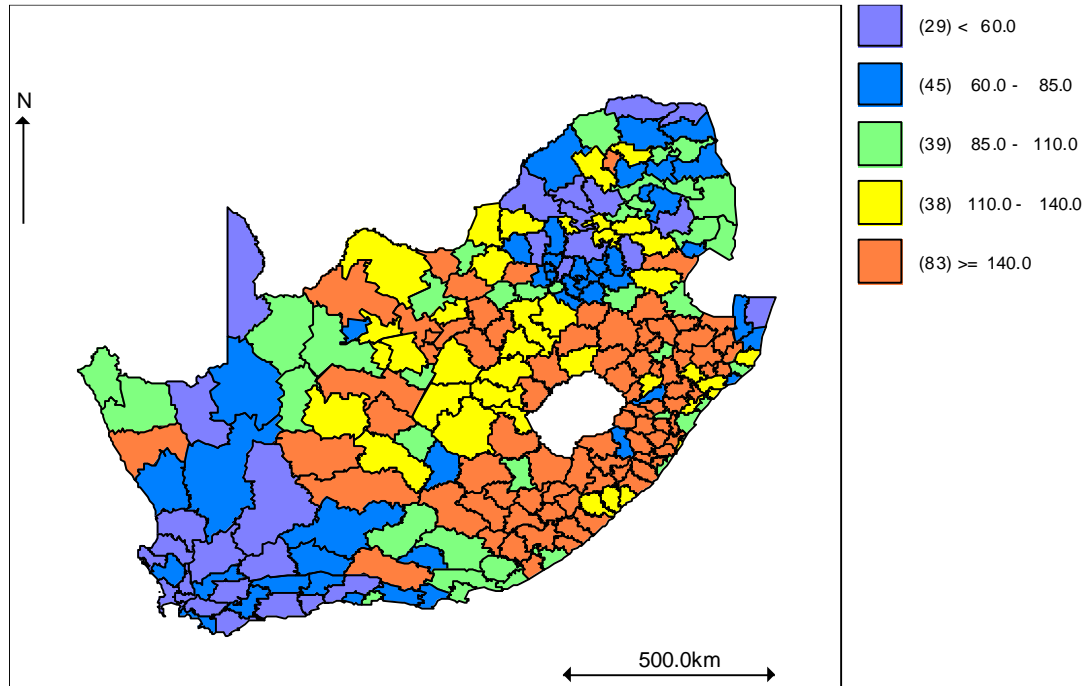
	Min	Max	Median	Mean	STD
No of deaths, Males, Census	8	10154	342	649	1184
No of deaths, Females, Census	10	8164	294	554	967
Mortality rate, Males, Census	15.6	243.3	113	118	50
Mortality rate, Females, Census	0	176.6	92	92	38
No of deaths, Males, Civil	7	10129	337	707	1380
No of deaths, Females, Civil	5	7698	275	562	1041
Mortality rate, Males, Civil	12.2	585.3	110	124	76
Mortality rate, Females, Civil	11.5	420.5	84	93	57
PHCR, Census	21.0	72.1	43.8	45.4	12.9
GI, Census	0.58	0.78	0.64	0,68	0.04
LS PHRC, Census	13.3	96.5	41.2	50.4	26.2

## POSSIBLE CONTEXTUAL FACTORS BASED ON CENSUS 2011

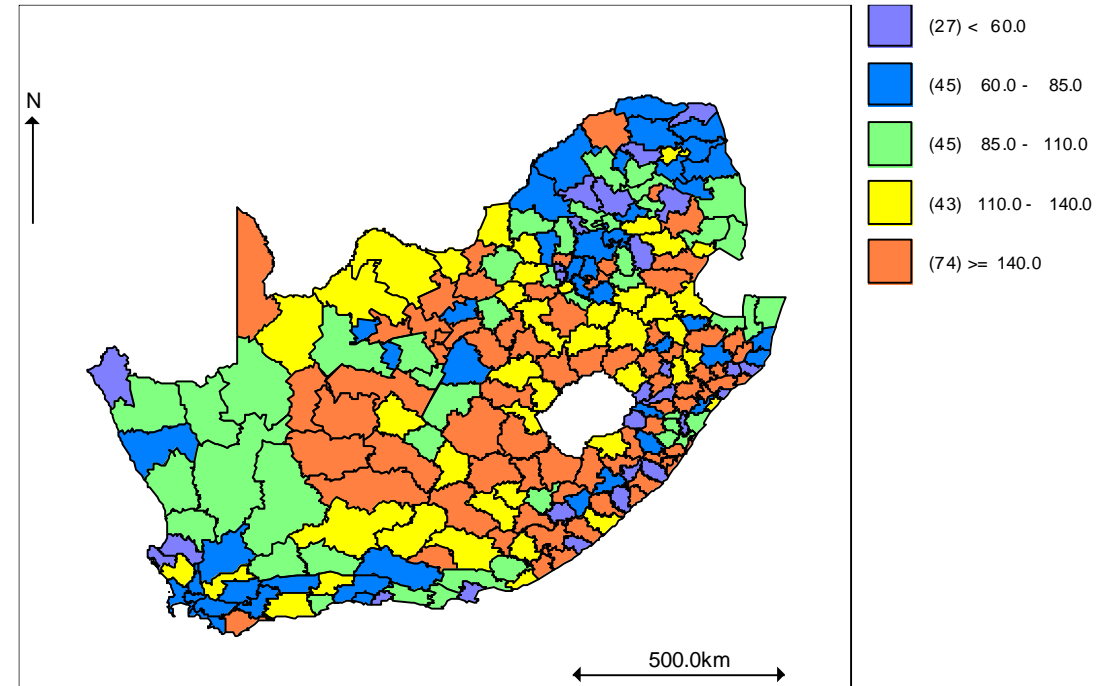
- South African Index of Multiple Deprivation (SAIMD)
  - A weighted score of four domains: income and material, employment, education and living environment deprivations measured at either the individual or household level (Noble, et al, 2013)
- Poverty headcount ratio (PHCR) is the proportion of population that live below the poverty line
- A living standard index (LSI) constructed from household data measures of wellbeing of people by accounting different factors beyond income. Similar to household wealth index
  - Households in the first two quintiles are categorised as poor (Living standard poverty headcount ratio (LS PHCR)).
- The Gini coefficient or index (GI) measures the level of income inequality in a society. The GI varies between 0 and 1, with 0 corresponding to perfect equality and 1 corresponds with perfect inequality
  - Values near 1 reflects unequal access to healthcare, nutrition and other services which linked to mortality
- Zewdie (2014) for PHCR, GI data

# Male Mortality rates

## Mortality rates, Males Census, 2011



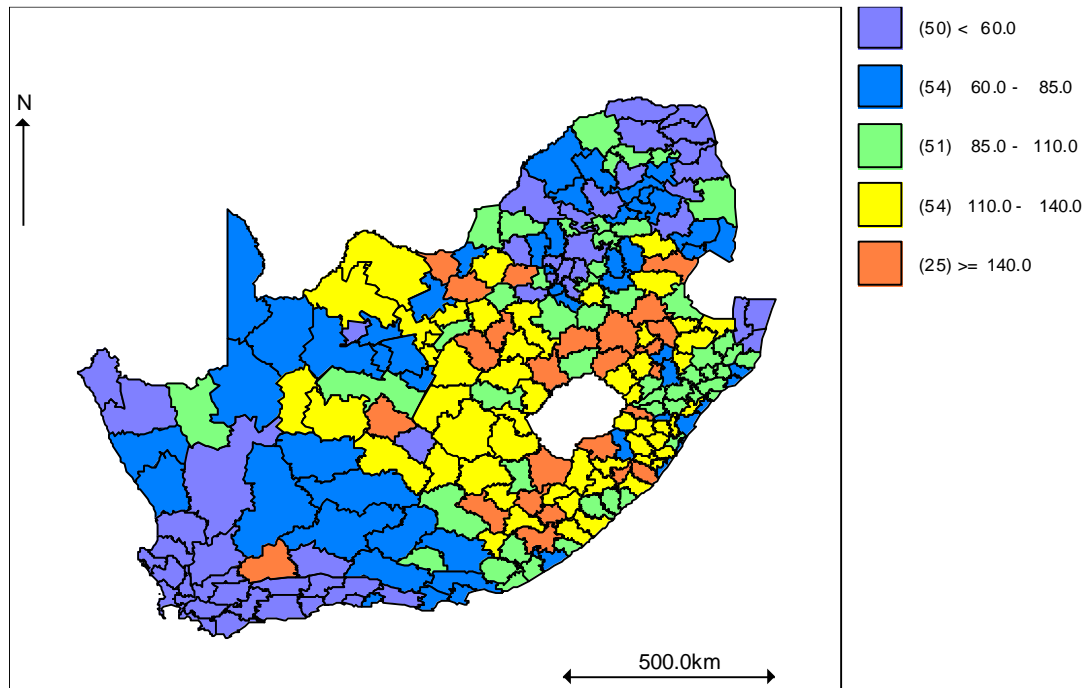
## Mortality rates, Males Civil Data



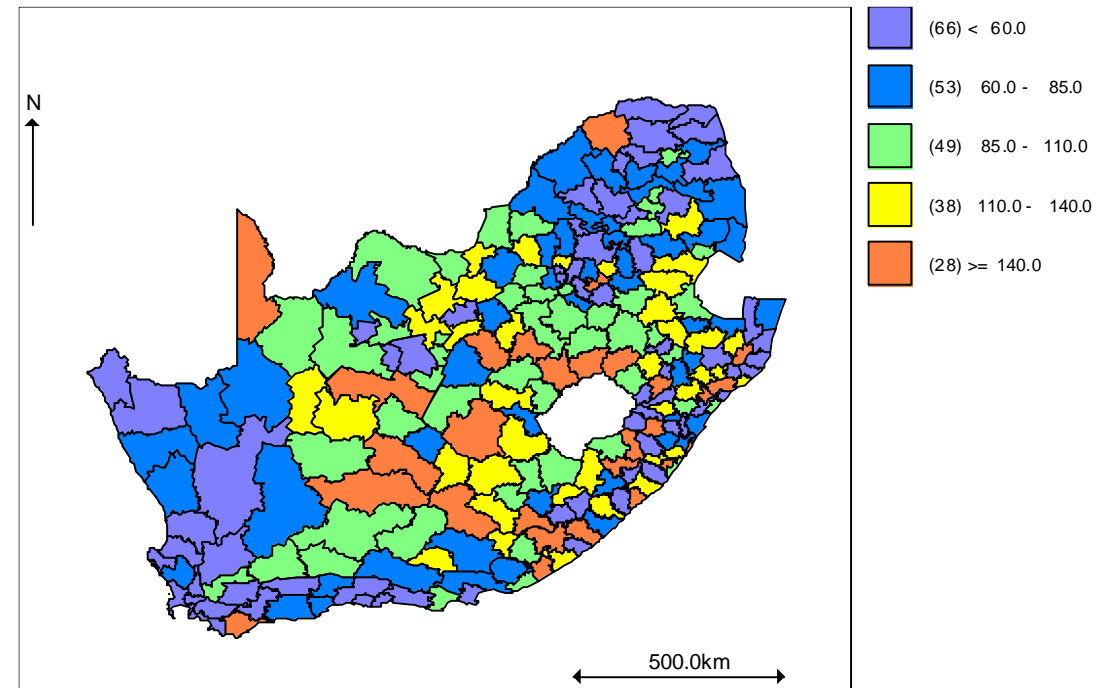


# Female Mortality rates

## Mortality rates, Female Census, 2011

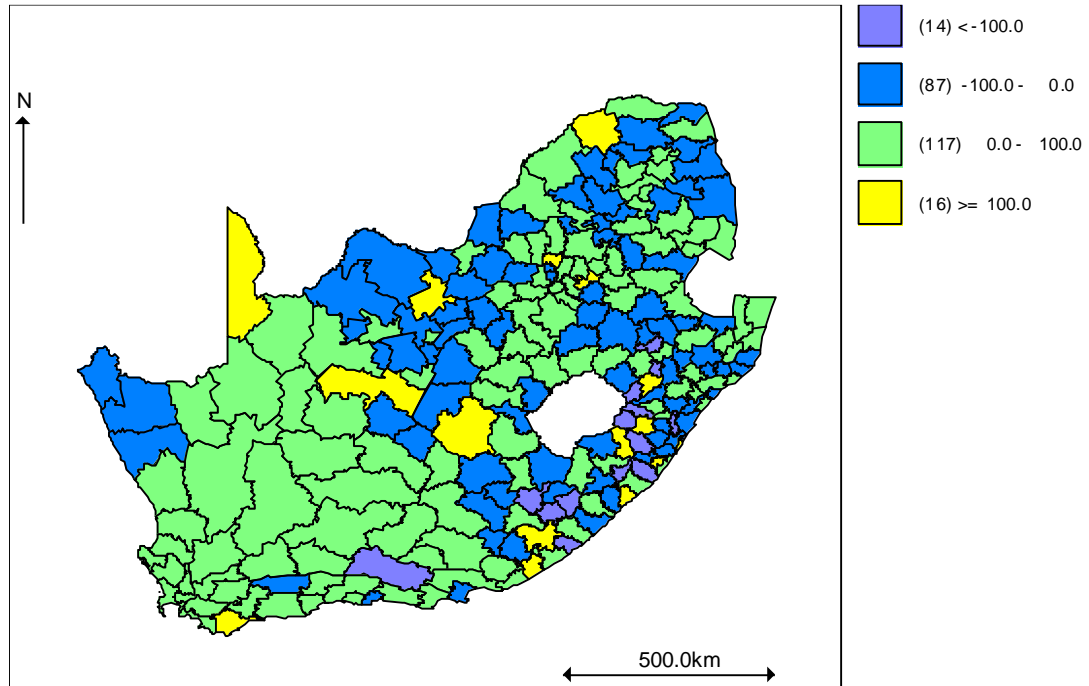


## Mortality rates, Female Civil Data

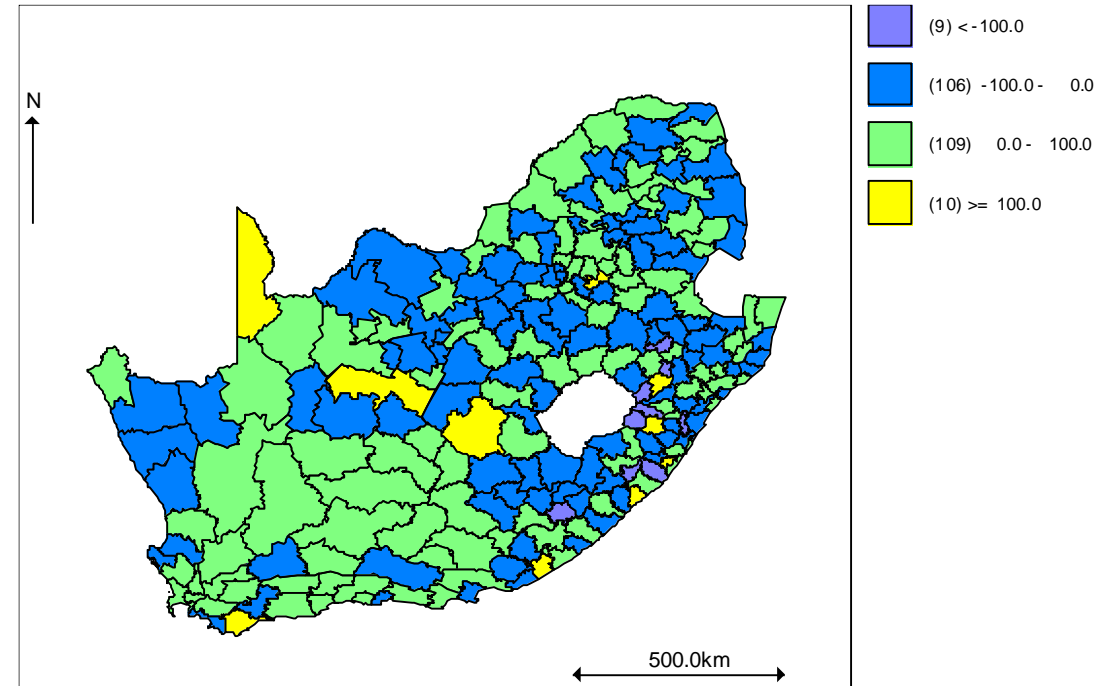


# Differences in rates, Civil-Census, 2011

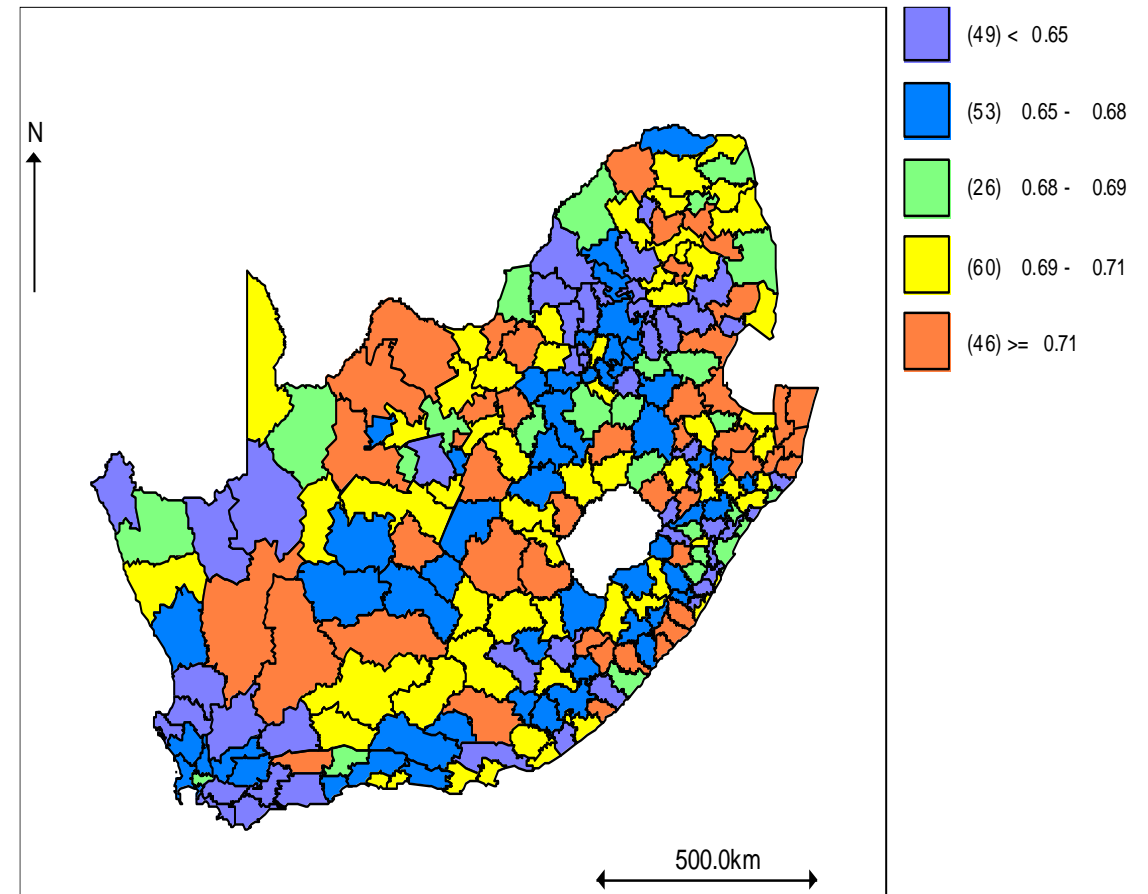
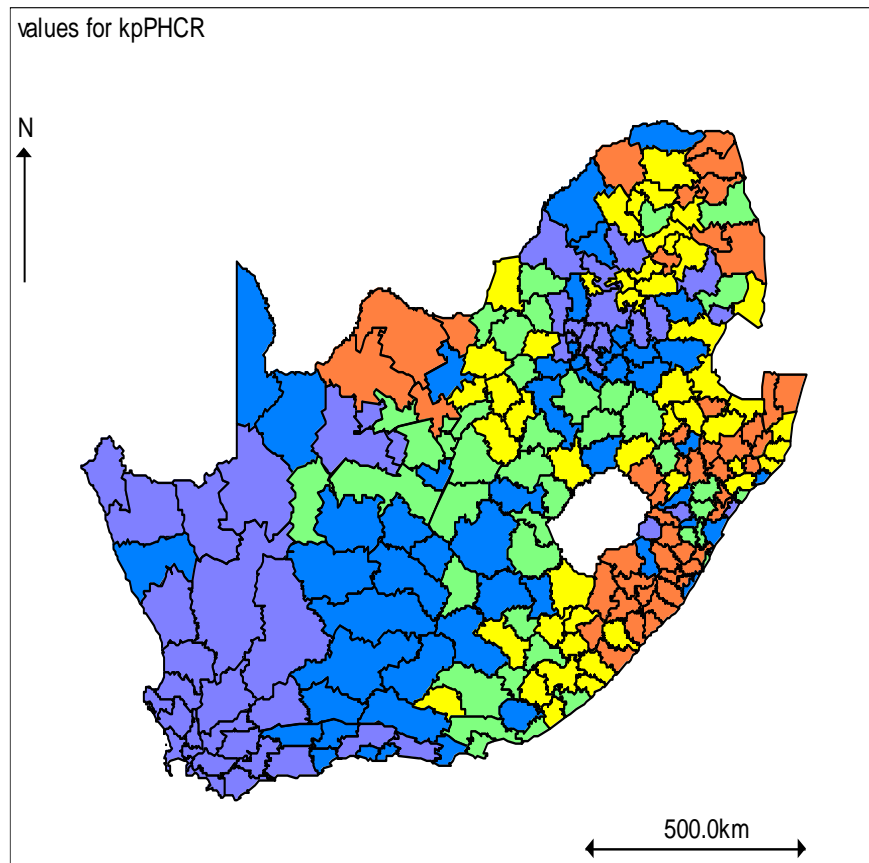
## Males



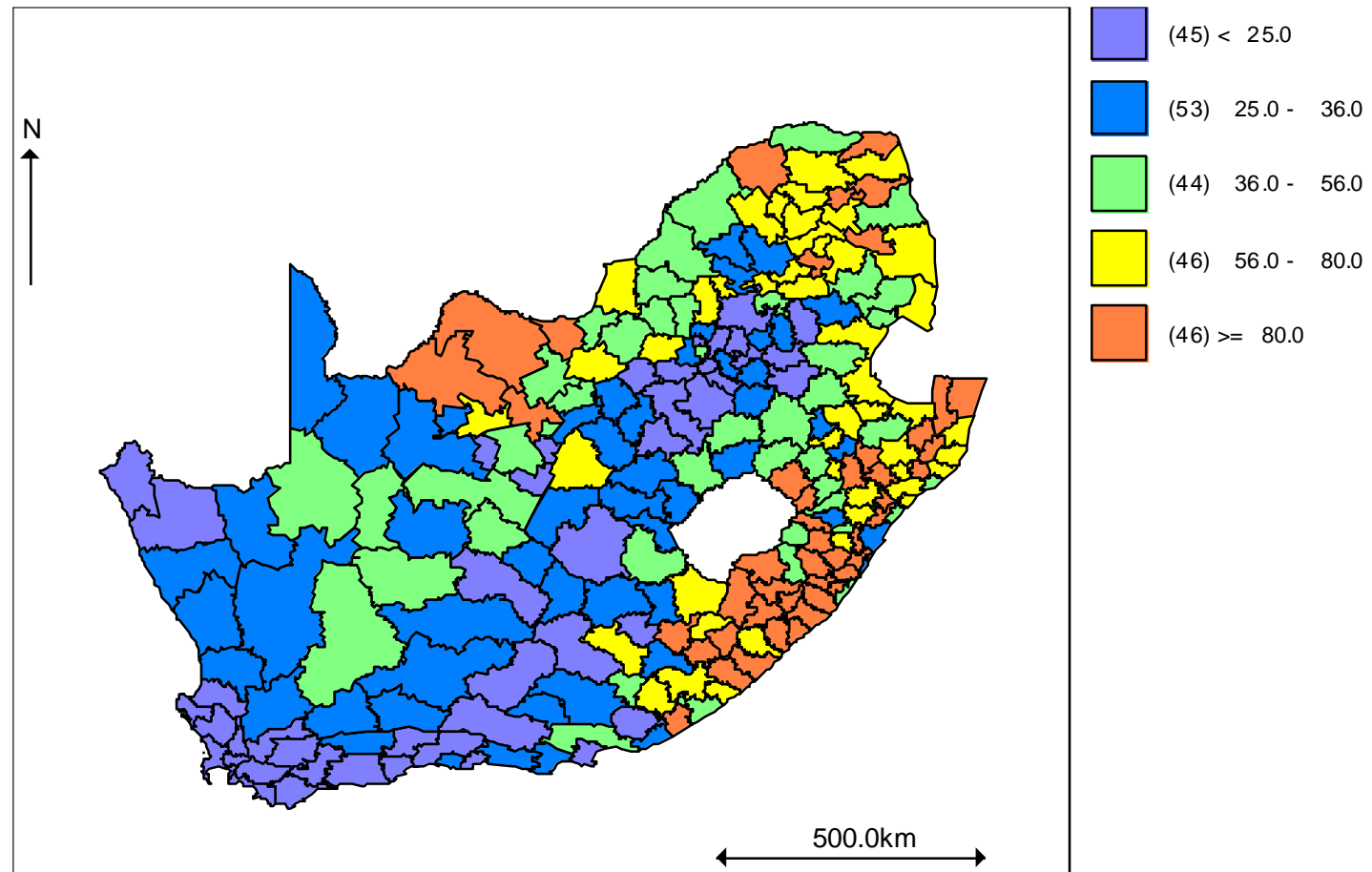
## Female



# Contextual Factors: left PHCR, right GI



# Contextual factor: LSPHCR



## Some considerations for mapping mortality data

- Most analyses model age and gender groups or cause-specific mortality using univariate spatial models (one at a time)
- BUT mortality risks from certain diseases are age and gender-dependent (Bradshaw et al, 2010)
  - ✓ Infectious aetiology (e.g. TB, pneumonia, meningitis, diarrhoea and HIV/AIDS) among young adults and children
  - ✓ Non-communicable disease death rates higher among the elderly

# BAYESIAN RELATIVE RISK MODELS FOR AREA SPATIAL DATA

- Bayesian models for area disease risks now widely applied (to detect smooth underlying risk surface over region).
- Assume observed disease counts  $O_i$  Poisson distributed,  
$$Y_i \sim \text{Poisson}(e_i r_i), (e_i = \text{expected counts})$$
- Relative risks  $r_i$  have average 1 when  $\text{sum}(\text{expected}) = \text{sum}(\text{observed})$ .
  - Expected counts based on applying region-wide disease rates to each small area population
- Modelling area relative risks
  - $\log(r_i) = \alpha + s_i + u_i,$
  - Spatial error:  $s_i \sim$  Conditional Autoregressive (CAR) prior
  - Heterogeneity/overdispersion error:  $u_i \sim$  Unstructured normal White Noise

## MULTIVARIATE SPATIAL MODELLING OF THE FOUR SPECIFIC MORTALITY RATES

- We could model the spatial effects without borrowing strength between the four mortality outcomes by using unrelated CAR for structured spatial effects and independent and identically distributed unstructured effect for each outcome.
- However, in order to model spatial co-mortality among the four mortality rates:
  - the four spatially structured effects are allowed to follow a four-way CAR model with the precision matrix  $\Sigma_s^{-1}$  taken to be Wishart distributed with 4 degrees of freedom and an appropriate scale matrix.
  - The four unstructured spatial random effects were modelled as multivariate normals within area covariance matrix  $\Sigma_v$  with  $\Sigma_v^{-1}$  taken to be Wishart distributed with 4 degrees of freedom and appropriate scale matrix.
- Joint mapping models Knorr-Held and Best, 2001; Richardson et al, 2006; Manda et al, 2009; 2012; Kandala et al, 2013)

# Neighbourhood Clustering in Elevated Risk

- Consider binary risk measures:

$b_i = 1$  if relative risk  $r_i > 1$ ,  $b_i = 0$  otherwise.

- These binary indicators are latent (unknown) as  $r_i$  are latent.
- Other thresholds (*e.g.*  $r_i > 1.5$ ) can be used
- Posterior exceedance probabilities of elevated disease risk

$$PE_i = Pr(r_i > 1 | data) = Pr(b_i = 1 | data)$$

in each area separately.

- Threshold of  $PE$ : Area  $i$  is hotspot if  $PE_i > 0.9$ , or  $PE_i > 0.8$ . Suitable thresholds may depend on data frequency (higher thresholds can be set for more frequent data)

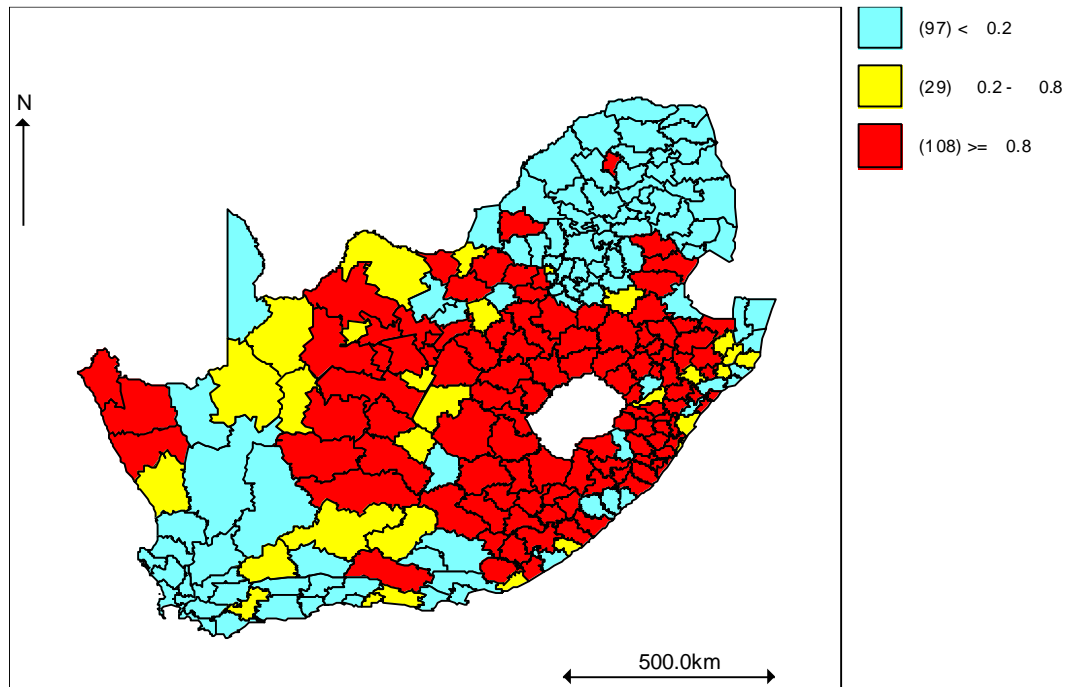


## MODEL EVALUATION AND VALIDATION

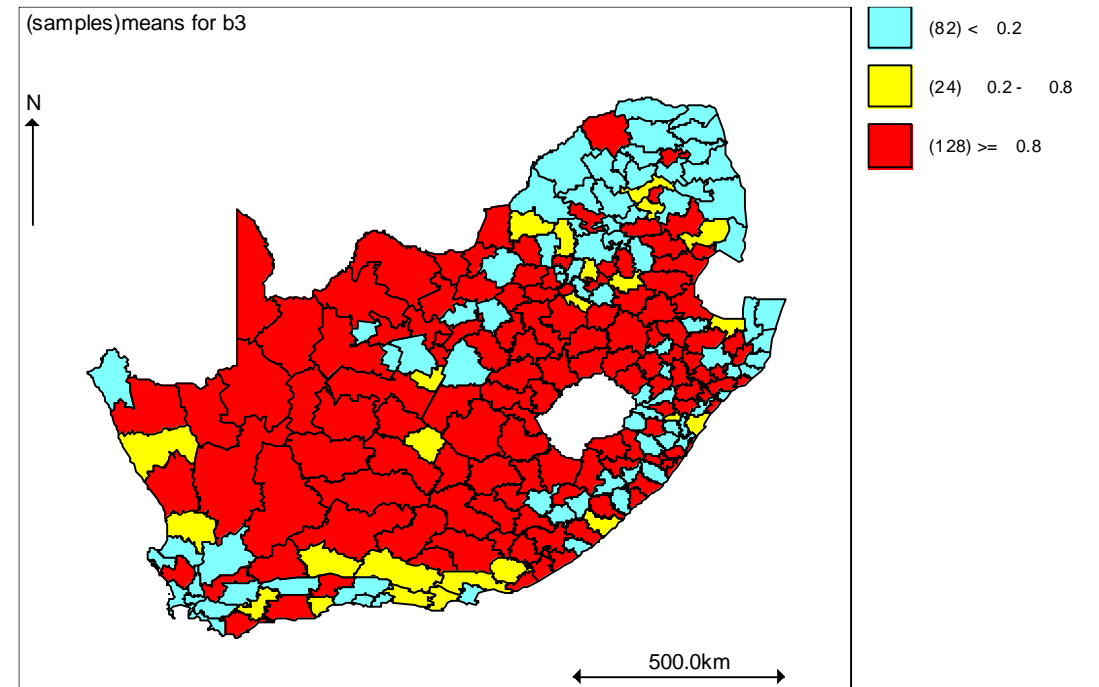
- **Conditional Predictive Ordinate (CPO), which is the marginal posterior predictive density. Also known as leave-one-validation.**
- **A large CPO indicates agreement between observations and the model.**
- **Calculated the product of individual area CPOs:  $CPO[1] * \dots * CPO[26]$  and took -log of this product (negative cross-validators log likelihood (NLLK), or negative log pseudo -marginal likelihood).**
- **Results: The multivariate model was more adequate than separate spatial specific-mortality models.**

# Posterior Probability of Smoothed Male SMR Exceeding 1

## Census, 2011

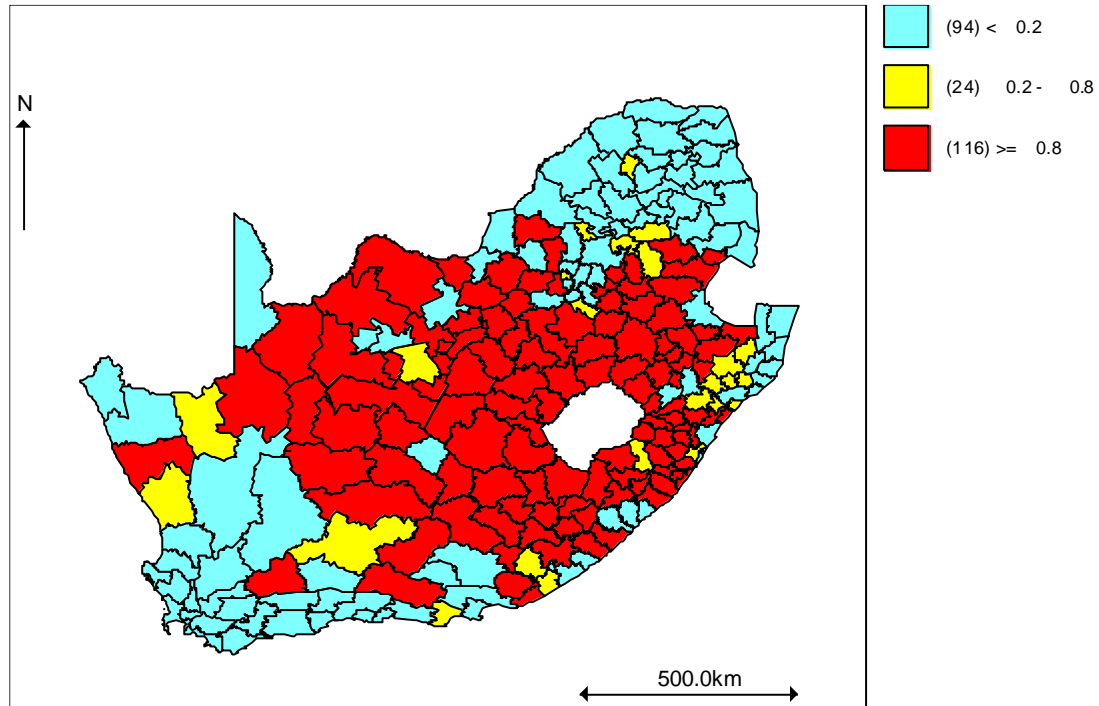


## Civil Data, 2011

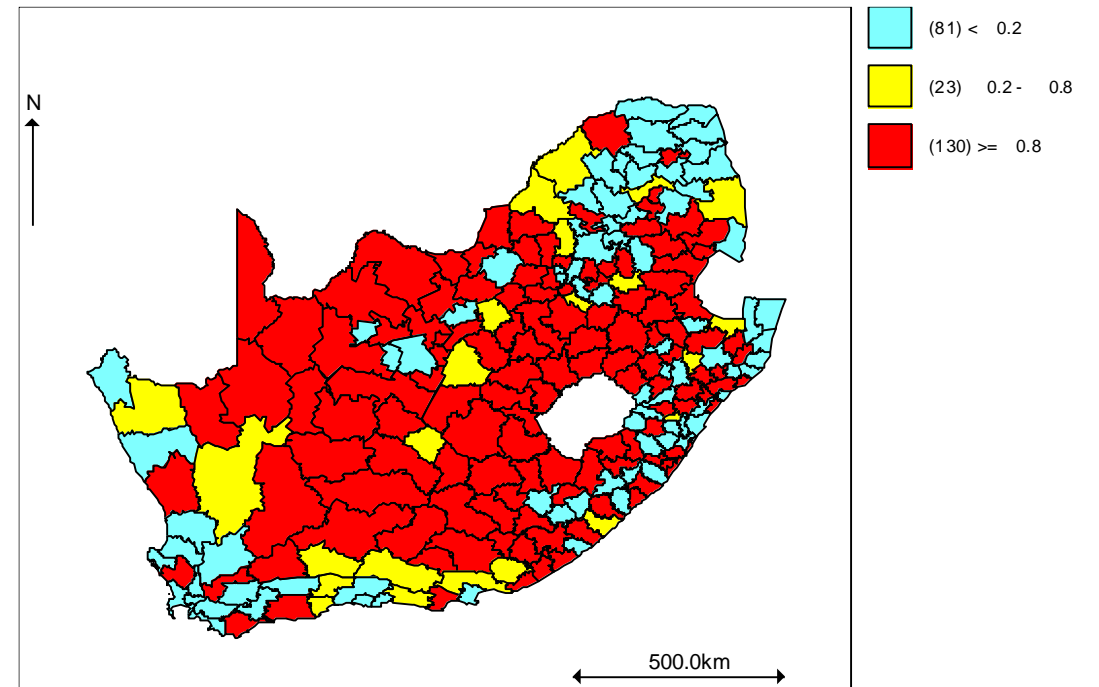


# Posterior Probability of Smoothed, Female SMR Exceeding 1

## Census, 2011



## Civil Data, 2011



# POSTERIOR ESTIMATES OF EFFECT OF PHRC ON MORTALITY RISK, SOUTH AFRICA 2011

PHCR		95% Lower Limit	RR	95% Upper Limit
2nd Fifth	<b>Male Mortality, Census</b>	<b>1.1</b>	<b>1.2</b>	<b>1.3</b>
	<b>Female Mortality, Census</b>	<b>1.2</b>	<b>1.2</b>	<b>1.4</b>
	<b>Male Mortality, Civil</b>	<b>1.3</b>	<b>1.4</b>	<b>1.6</b>
	<b>Female Mortality, Civil</b>	<b>1.3</b>	<b>1.4</b>	<b>1.6</b>
3rd Fifth	<b>Male Mortality, Census</b>	<b>1.4</b>	<b>1.5</b>	<b>1.6</b>
	<b>Female Mortality, Census</b>	<b>1.4</b>	<b>1.5</b>	<b>1.6</b>
	<b>Male Mortality, Civil</b>	<b>1.2</b>	<b>1.4</b>	<b>1.7</b>
	<b>Female Mortality, Civil</b>	<b>1.2</b>	<b>1.4</b>	<b>1.6</b>
4th Fifth	<b>Male Mortality, Census</b>	<b>1.8</b>	<b>1.9</b>	<b>2.1</b>
	<b>Female Mortality, Census</b>	<b>1.7</b>	<b>1.8</b>	<b>2.0</b>
	<b>Male Mortality, Civil</b>	<b>1.1</b>	<b>1.3</b>	<b>1.6</b>
	<b>Female Mortality, Civil</b>	<b>1.1</b>	<b>1.3</b>	<b>1.6</b>
5th Fifth	<b>Male Mortality, Census</b>	<b>1.9</b>	<b>2.1</b>	<b>2.3</b>
	<b>Female Mortality, Census</b>	<b>1.7</b>	<b>1.8</b>	<b>2.0</b>
	<b>Male Mortality, Civil</b>	<b>1.2</b>	<b>1.3</b>	<b>1.6</b>
	<b>Female Mortality, Civil</b>	<b>1.1</b>	<b>1.3</b>	<b>1.6</b>

# Summary

- Both show evidence of municipality variations in adult mortality, with south-eastern parts with increases risk
- Both highly and positively associated with levels of poverty

# Summary

- Presented a model that combines different data sources for mortality to estimate sub-national risk profiles in mortality
- Alternative methods exist, for example Dorrington and Timæus (year) proposed a method of combining the mortality sources of to produce reliable estimates of sub-national mortality
- Contextual health indicators e.g. HIV at lower levels, data source maybe ANC HIV