A Bayesian change point analysis of prevalence rates of chronic diseases by gender and race

Trivellore Raghunathan
Chair and Professor of Biostatistics, School of Public Health
Research Professor, Institute for Social Research
University of Michigan

Presented at the National Conference on Health Statistics, August 16-18, 2010
Setup

• Need for assessing the health of the nation cannot be understated
• NCHS has been in the forefront of collecting data needed
• A look back on one such long series
  – National Health Interview Survey (as it is currently known)
    – Began in 1957
    – Limited Data available prior to 1962
• Relate the prevalence information to time line of our history
Data

• Diseases
  – Asthma and Hypertension (1962-2008)

• Populations
  – Overall
  – Male/Female
  – White/Black

• Time line
  – Important events (health and economic)

• The analysis is preliminary
Prevalence of Diabetes

- Civil Rights (64)
- Medicaid/AA (65)
- Medicare (66)
- Vietnam War
- Recession (74)
- Oil crisis (73-75)
- Recession (79)
- Gulf war Recession (90)
- Recession (97)
- 9/11
Model

\[ p_t = \text{Estimate for year } t \]

\[ P_t = \text{Population value} \]

\[ \nu_t = \text{sampling variance} \]

\[ n_e = p_t (1 - p_t) / \nu_t \]

\[ \pi(\alpha, \sigma_b) \propto 1 \]

\[ \sin^{-1} \sqrt{p_t} \sim N(\theta_t = \sin^{-1} \sqrt{P_t}, n_e^{-1} / 4) \]

\[ \theta_t = \alpha_0 + \alpha_1 t + \alpha_2 t^2 + \sum_{j=1}^{L} b_j \{ (t - k_j)_+ \}^2 \]

\[ b_j \sim iid \ N(0, \sigma_b^2) \]

\[ (t - k_j)_+ = \max(t - k_j, 0) \]
Analysis

• Fit the model using MCMC or MLE
• Calculate $E(P_t \mid Data)$
• Choose knots based on the time-line
• For this analysis choose 5 knots (1965, 1974, 1980, 1990, 2001)
• Modification (in progress)
  – Treat knots as unknown parameters and estimate them from the data
Diabetes by Gender
Diabetes by Race
Prevalence rates of Asthma
Prevalence rates of Hypertension
Diabetes by administration

Prevalence

Year


0.02 0.04 0.06 0.08
Issues and Concerns

• Major design changes, instruments, definitions etc
  – Reflecting such changes in the model development
• Any policy change takes time to have an impact. Incorporation of time lag in the model
• Outliers
• Regardless of these concerns:
  – Persistent differences between race
  – Lessening of differences between men and women
Discussion

• Relating the time series data to important epics/events can be useful to understand potential effects of policy/cultural changes

• Investigation of health disparities across various subpopulations

• Subgroups can be determined based on economic, social, cultural aspects of the society

• Useful exercise for any democratic society to look at the health of the nation

• Thank you NCHS, for this great data source!!