Bayesian variable selection for identifying subgroups in cost-effectiveness analysis

Elías Moreno\textsuperscript{1}  Francisco–Javier Girón\textsuperscript{2}  Francisco–José Vázquez–Polo\textsuperscript{3}  Miguel Negrín\textsuperscript{3}

\textsuperscript{1}University of Granada, Spain
\textsuperscript{2}University of Málaga, Spain
\textsuperscript{3}University of Las Palmas de Gran Canaria, Spain
Outline

1. **Context**
   - Nixon and Thompson (2005) model

2. **Model**

3. **Simulation exercise**

4. **Results**

5. **Application with real data**

6. **Conclusions**
Analysis of subgroups

- Policy–makers interest cost–effectiveness for patient subgroups (NICE Decision Support Unit, 2007)
Analysis of subgroups

- Policy–makers interest cost–effectiveness for patient subgroups (NICE Decision Support Unit, 2007)
- Heterogeneity in incemental cost–effectiveness analysis (Sculpher, 2010)

References: Willan et al. (2004), Nixon and Thompson (2005), Vázquez–Polo et al. (2005), Hoch et al. (2006), Manca et al. (2007), Willan and Kowgier (2008), Moreno et al. (2012)
Analysis of subgroups

- Policy–makers interest cost–effectiveness for patient subgroups (NICE Decision Support Unit, 2007)
- Heterogeneity in incremental cost–effectiveness analysis (Sculpher, 2010)
- Regression methods have been proposed as an appropriate method in cost–effectiveness analysis where the subgroups analysis can be carried out with the inclusion of interactions between treatment and subgroup.
Analysis of subgroups

- Policy–makers interest cost–effectiveness for patient subgroups (NICE Decision Support Unit, 2007)
- Heterogeneity in incremental cost–effectiveness analysis (Sculpher, 2010)
- Regression methods have been proposed as an appropriate method in cost–effectiveness analysis where the subgroups analysis can be carried out with the inclusion of interactions between treatment and subgroup.

Analysis of subgroups

- Policy–makers interest cost–effectiveness for patient subgroups (NICE Decision Support Unit, 2007)
- Heterogeneity in incemental cost–effectiveness analysis (Sculpher, 2010)
- Regression methods have been proposed as an appropriate method in cost–effectiveness analysis where the subgroups analysis can be carried out with the inclusion of interactions between treatment and subgroup.
- Moreno et al. (2012) proposed an analysis of subgroups based on an optimal Bayesian variable selector.
Analysis of subgroups

- Policy–makers interest cost–effectiveness for patient subgroups (NICE Decision Support Unit, 2007)
- Heterogeneity in incremental cost–effectiveness analysis (Sculpher, 2010)
- Regression methods have been proposed as an appropriate method in cost–effectiveness analysis where the subgroups analysis can be carried out with the inclusion of interactions between treatment and subgroup.

- Moreno et al. (2012) proposed an analysis of subgroups based on an optimal Bayesian variable selector.
- In this work we show a simulation study to compare both methods.
Nixon and Thompson (2005) model

Differences between subgroups
Modelization for a patient $j$ in arm $i$.

\[
E_{ij} \sim \text{Dist}(\phi_{Eij}, \sigma_{Ei}) \\
C_{ij} \sim \text{Dist}(\phi_{Cij}, \sigma_{Ci})
\]

\[
\phi_{Eij} = \mu_{Ei} + \beta_i(C_{ij} - \phi_{Cij}) + \sum \gamma_E x_{ij} + \sum \delta_E l_i x_{ij} \\
\phi_{Cij} = \mu_{Ci} + \sum \gamma_C x_{ij} + \sum \delta_C l_i x_{ij}
\]

Comments
- Covariates have the same influence for both treatments, except subgroups.
- Detecting subgroups is reduced to an hypothesis test about the statistical relevance of parameters $\delta$.
- Its modelization is appropriate for Normal and Gamma models.
Model proposed by Moreno *et al.* (2012)

### Differences between subgroups

Modelization for a patient $j$ in arm $i$.

\[(E_{ij}, C_{ij}) \sim MVN((\phi_{Eij}, \phi_{Cij}), \Sigma_i)\]

\[
\phi_{Eij} = \beta_0i + \sum \beta_i x_{ij}
\]

\[
\phi_{Cij} = \gamma_0i + \sum \gamma_i x_{ij}
\]

### Comments

- Objective Bayesian variable selection is carried out to detect the covariates with influence. Selecting covariates define a subgroup over the effectiveness and (or) cost.
- Normal and Log–normal distributions can be considered.
Bivariate Objective Bayesian Variable Selection

Posterior probability for each model

\[
P(M_j|Y, X_j) = \frac{B_{j1}(Y, X_j)}{1 + \sum_{k=2}^{2^p-1} B_{k1}(Y, X_k)}
\]

Intrinsic prior (Torres et al., 2011)

\[
\pi_1^j(B_1, \sigma_1) = c \frac{1}{\sigma_1}, \quad \pi_j^j(B_j, \sigma_j|B_1, \sigma_1) =
\]

\[
N_{j \times 2} \left[ B_j|\Delta_j, \frac{n}{j+1}(\sigma_j^2 + \sigma_1^2) \left( (X_j^t X_j)^{-1} \otimes V \right) \right] \times \frac{2\sigma_j}{\sigma_1^2 \left( 1 + \sigma_j^2 / \sigma_1^2 \right)},
\]

where \( \Delta = (0_{(j-1)\times 2} B_1) \).
Bayes factor for intrinsic priors

\[ B_{k1}(Y, X_k) = 2(k + 1)^{(k-1)} \int_{0}^{\pi/2} \frac{\sin(\varphi)^{2(k-1)+1} (n + (k + 1) \sin^2 \varphi)^{(n-k)}}{\cos(\varphi)^{-1} [(k + 1) \sin^2 \varphi + n\mathcal{B}_{k1}]^{(n-1)}} d\varphi. \]

where

\[ \mathcal{B}_{k1} = \frac{\text{tr}[H_{X_k} Y V^{-1} Y^t]}{\text{tr}[H_{X_1} Y V^{-1} Y^t]}, \]

and \( H_X = I_n - X(X^t X)^{-1} X^t. \)
Simulation

$X_1$, $X_2$ and $X_3$ covariates were simulated from a Uniform(0,10) distribution.

\[ E_{ij} \sim N(\phi_{Eij}, 1) \]
\[ C_{ij} \sim N \text{ or Gamma}(\phi_{Cij}, 1) \]

Bivariate normal distribution with $\rho = 0.5$ or FGM copula for Normal-Gamma simulation.

Treatment 1:
\[ \phi_{E_{i1}} = 1 + 0.7X_{1i} + 0.2X_{2i} \]
\[ \phi_{C_{i1}} = 5 + 1X_{1i} + 0.3X_{2i} \]

Treatment 2:
\[ \phi_{E_{i2}} = 2 + 0.7X_{1i} + 0.1X_{2i} \]
\[ \phi_{C_{i2}} = 8 + 2X_{1i} + 0.2X_{2i} \]
Simulation

\[ E_{ij} \sim N(\phi_{Eij}, 1) \]
\[ \log - C_{ij} \sim N(\phi_{Cij}, 0.1) \]

Bivariate normal distribution with \( \rho = 0.5 \)

Treatment 1:

\[ \phi_{C1i} = 1.74235 + 0.1X_{1i} + 0.03X_{2i} \]

Treatment 2:

\[ \phi_{C2i} = 1.79444 + 0.2X_{1i} + 0.02X_{2i} \]
Different frameworks for different sample–sizes were considered. We carry out 1,000 simulations and we define as an optimal selection when:

- **Objective variable selection:** The model with the highest posterior probability is intercept, X1 and X2. The selection is carried out for the Treatment 1 and 2.
- **Nixon and Thompson model:** Only the variable X2 is detected as a subgroup for effectiveness and X1 and X2 are detected as subgroups for the cost model.

Simulations were carried out with Mathematika and WinBUGS using the R2WinBUGS package.
Results: Normal data

[Graph showing normal data with different scenarios and outcomes labeled by Elias et al. and Nixon and Thompson.]

- Normal costs
- Asymmetric costs

N1=N2=50, N1=70, N2=30, N1=N2=200, N1=300, N2=100
Results: Gamma data

Graph showing results for Normal and Asymmetric costs with different values of N1 and N2.
Results: Log–normal data

Graph showing normal costs and asymmetric costs with different scenarios.
Example with real data

- Data from a randomized clinical trial (Hernández et al., 2003) that compares two alternative treatments for exacerbated chronic obstructive pulmonary disease (COPD): home hospitalization or conventional.
- Effectiveness: Difference between the score at the beginning and at the end of the study of the St. George’s Respiratory Questionnaire (SGRQ).
- Potential covariates: Age, sex, smoking habit, forced expiratory volume in one second (FEV), exacerbations requiring in–hospital admission (HOSV) and the score at the beginning of the study (SGRQ1).
Example with real data: Variable Selection

**Treatment 1**
SGRQ1, Age, FEV

**Treatment 2**
SGRQ1, FEV
Example with real data: Posterior analysis
Conclusions

- Cost–effectiveness analysis based on regression methods facilitates the analysis of subgroups with the inclusion of interactions terms in the model.
- The identification of subgroups is reduced to an hypothesis test about the relevance of these parameters.
- Bayesian Variable Selection is proposed as a natural way for the identification of subgroups.
- Simulation study shows the preference for the Bayesian Variable Selection.
- Bayesian Variable Selection obtains good results even with small sample sizes.
- Bayesian Variable Selection is less sensitive to the distribution assumption.