Usefulness of Bayesian modeling in risk analysis and prevention of Home Leisure and Sport Injuries (HLIs)

Madelyn Rojas Castro, Marina Travanca, Marta Avalos, David Valentin Conesa, Emmanuel Lagarde

To cite this version:

Madelyn Rojas Castro, Marina Travanca, Marta Avalos, David Valentin Conesa, Emmanuel Lagarde. Usefulness of Bayesian modeling in risk analysis and prevention of Home Leisure and Sport Injuries (HLIs). III Jornadas Científicas de Estudiantes de la Sociedad Española de Biometría, Jan 2018, Bilbao, Spain. hal-01964484

HAL Id: hal-01964484
https://hal.archives-ouvertes.fr/hal-01964484
Submitted on 22 Dec 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Usefulness of Bayesian modeling in risk analysis and prevention of Home Leisure and Sport Injuries (HLIs)

Madelyn Rojas Castro¹, Marina Travanca¹, Marta Avalos¹, David Conesa², Emmanuel Lagarde¹

¹Bordeaux Population Health Research Center INSERM U1219, Injury epidemiology, transport, occupation (IETO).
²Departamento de Investigación Operativa, Universidad de Valencia

18 January 2018
Home Leisure and Sport Injuries (HLIs)

![Diagram of Home Leisure and Sport Injuries (HLIs)](image)
Injuries Epidemiological Context

World Health Organization (WHO)

- **Injuries** are the 4th cause of mortality in the EU.
- **230 thousand** annual deaths (EU), 58 % are HLIs.

In France:
- **HLIs are the 3th cause of mortality.**
- **Leading cause of childhood mortality.**

Each year:
- **20 thousand** deaths, 5 times more than traffic accidents.
- **5 million** emergencies.
- **11 million** injuries
L’ Observatoire MAVIE

- Prospective online cohort study of HLIs
- Currently, MAVIE has more than 26 thousand volunteers in France during 3 years of recruitment (target sample 100 thousand volunteers).

Objectives

- Identify the risk factors associated with the HLIs occurrence and severity.
- Implement prevention measures to reduce the number of victims.
MAVIE-Lab (mHealth)

- **Mobile app** including a DSS (Decision Support System)
- To self-management of HLIs risks (Evaluation).
- To experience personalized **prevention solutions** to reduce the risk of injury (Mitigation).
MAVIE-Lab Development

Data Collection

System Evaluation

Implementation

Update App

Identify risk factors and protectors

Risk Model

Figure: Steps for MAVIE-Lab Development
Modeling Problems MAVIE data

1. Reduced number of injuries declared by each activity.
2. Missing values.
3. Under-representation between injuries reported and those occurred.
4. Complex relationships between risk factors.

Risk Model

\[ p(\theta \mid X, y) \propto p(\theta) \ l(\beta, \sigma \mid X, y) \]

*posterior \propto prior \times likelihood*
Bayesian Generalized Linear Models (Logistic Regression)

**Bayesian Approach**

- **Current study information**
  - **+Prior Information**
    - **+Elicitation**
      - **Results Previous Studies**
      - **Published literature**
      - **Experts Information**
  - **+Raw Data: Similar Studies**
    - **Hierarchical analysis**
    - **General Information Databases**
- **Prior no informative**
  - **Distr. Normal**
  - **Jeffrey’s Priors**
- **Weakly Informative**

Madelyn Rojas Castro, Marina Travanc, Marta Avalos, David Conesa, Emmanuel Lagarde

Bayesian Modeling of HLIs
Bayesian Generalized Linear Models (Logistic Regression)

Methodological Proposal

**Priors experts elicitation**

**Reason for using Bayesian experts elicitation**

1. The knowledge is limited or incomplete.
2. The evidence is inconsistent, missing or ambiguous.
3. The questions are complex and the relations between variables.
4. To deal with bias and uncertainties.
5. To integrate different sources of knowledge.

(A. B. Knol et al. 2010)
Bayesian experts elicitation (MAVIE-Lab)
Exploratory analysis objectives

Principal Objective
To explore **Bayesian modeling methodologies** for being used in MAVIE-Lab development

- To explore *experts elicitation* to improve the estimation of model parameters.
- To explore the use of automatic *selection models methods* since the Bayesian approach.
Model Selection  BMA  Bayesian Model Averaging

The BMA is a weighted average of the posterior distributions for each parameter for all possible models.

Figure taken from (FitzGerald et al. 2014).
Sample of MAVIE Cohort: N = 4,345 (March 2017). Volunteers over 18 years old, who had completed and validated the questionnaire.

Injuries: 603 Reported Injuries (13.87%).

Variables: Explanatory variables: 20 categorical or categorized variables (Factors associated with HLIs occurrence).

- Demographic: Age and sex.
- Previous Injuries.
- Physical and mental health: BMI, Health problems, depression, anxiety, hyperactivity, drowsiness and concentration.
- Consumption: Medicines, alcohol, tobacco and cannabis.
- Sport Practice: Sports, use of compression and maintenance accessories.
Methodology

MAVIE Data

Multiple Imputation

BY Logistic Model (Non informative priors)

Model Selection (BMA Bayesian Model Averaging)

Elicitation Prior Distributions

BY Logistic Model (Informative priors)
Results

BY Logistic Model (non informative priors)

(a) CI 95%
(b) ROC (AUC=0.58)
(c) Separation Graphic
Model Selection BMA: Better models according to BIC criteria

Selected Variables: Sex, Health Prob., Previous HLIs, Sports.
Model BMA Variables:
BY Logistic Model (Non informative priors)

(d) CI 95%  
(e) ROC (AUC=0.64)  
(f) Separation Graphic

<table>
<thead>
<tr>
<th>Variables</th>
<th>OR</th>
<th>Mean $\beta$</th>
<th>Variance $\beta$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex (F)</td>
<td>0.44</td>
<td>-0.36</td>
<td>0.04</td>
</tr>
<tr>
<td>Age 50-60</td>
<td>1.20</td>
<td>0.08</td>
<td>0.05</td>
</tr>
<tr>
<td>Age 40-50</td>
<td>1.40</td>
<td>0.15</td>
<td>0.04</td>
</tr>
<tr>
<td>Age 30-40</td>
<td>1.40</td>
<td>0.15</td>
<td>0.05</td>
</tr>
<tr>
<td>Age 18-30</td>
<td>2.10</td>
<td>0.32</td>
<td>0.04</td>
</tr>
<tr>
<td>Health Prob.</td>
<td>1.74</td>
<td>0.24</td>
<td>0.07</td>
</tr>
<tr>
<td>Smoker</td>
<td>1.64</td>
<td>0.21</td>
<td>0.03</td>
</tr>
<tr>
<td>Ex-smoker</td>
<td>1.34</td>
<td>0.13</td>
<td>0.03</td>
</tr>
<tr>
<td>Alcohol</td>
<td>5.72</td>
<td>0.76</td>
<td>0.28</td>
</tr>
</tbody>
</table>
Informative Prior Distributions

(g) CI 95%

(h) ROC (AUC=0.6)

(i) Separation Graphic
## Results BY model: (BMA, Prior Information)

<table>
<thead>
<tr>
<th></th>
<th>OR</th>
<th>Devest</th>
<th>2.5%</th>
<th>97.5%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Age (50-60)</td>
<td>0.95</td>
<td>0.035</td>
<td>0.883</td>
<td>1.020</td>
</tr>
<tr>
<td>Age (40-50)</td>
<td>0.93</td>
<td>0.037</td>
<td>0.855</td>
<td>1.002</td>
</tr>
<tr>
<td>Age (30-40)</td>
<td>0.89</td>
<td>0.043</td>
<td>0.812</td>
<td>1.973</td>
</tr>
<tr>
<td>Age (18-30)</td>
<td>1.29</td>
<td>0.076</td>
<td>1.154</td>
<td>1.443</td>
</tr>
<tr>
<td>Sex (F)</td>
<td>0.80</td>
<td>0.024</td>
<td>0.760</td>
<td>0.851</td>
</tr>
<tr>
<td>Sport</td>
<td>1.58</td>
<td>0.065</td>
<td>1.467</td>
<td>1.719</td>
</tr>
<tr>
<td>Health Prob.</td>
<td>1.57</td>
<td>0.050</td>
<td>1.476</td>
<td>1.668</td>
</tr>
<tr>
<td>Smokers</td>
<td>1.54</td>
<td>0.074</td>
<td>1.397</td>
<td>1.688</td>
</tr>
<tr>
<td>Ex-smokers</td>
<td>1.44</td>
<td>0.060</td>
<td>1.331</td>
<td>1.567</td>
</tr>
<tr>
<td>Physic Health</td>
<td>1.35</td>
<td>0.052</td>
<td>1.252</td>
<td>1.453</td>
</tr>
<tr>
<td>Psc.Medic.</td>
<td>1.22</td>
<td>0.038</td>
<td>1.146</td>
<td>1.294</td>
</tr>
<tr>
<td>Mantenim. Acc.</td>
<td>1.29</td>
<td>0.072</td>
<td>1.150</td>
<td>1.434</td>
</tr>
<tr>
<td>Comp. Acc.</td>
<td>1.31</td>
<td>0.065</td>
<td>1.188</td>
<td>1.440</td>
</tr>
<tr>
<td>Previous HLIs</td>
<td>2.41</td>
<td>0.095</td>
<td>2.236</td>
<td>2.605</td>
</tr>
<tr>
<td>Alcohol</td>
<td>1.18</td>
<td>0.037</td>
<td>1.112</td>
<td>1.254</td>
</tr>
</tbody>
</table>
• The model does not allow separating profiles of greater or lesser injury risk.

• The variables included do not explain the occurrence of injuries.

Bayesian approach remains appropriate for the development of MAVIE-Lab.
Solutions and ongoing work

1. To make **more specific models by type of injury**, including the most important variables in each case (Example: **Sport Injuries PhD Project**).

2. To perform a **formal elicitation** (Devilee & A. Knol 2011):
   - Experts selection
   - Uncertain evaluation
   - Elicitation protocol

3. To perform **Bayesian Network models** including besides the relationships between variables (**probabilistic and graphical modeling**) (Mujalli *et al.* 2016).
Variables relations in Running Injuries (DAG)
In conclusion, **Bayesian statistic** and the expert’s **elicitation** are powerful tools for the construction of **expert system** to be included in mHealth. This methodology makes possible to combine **statistical data** and experts information as for example **medical advice**.

**Figure: mHealth** (Figure taken from web-site UNC Gillinds School of Global Public Health)
Bibliography


Usefulness of Bayesian modeling in risk analysis and prevention of Home Leisure and Sport Injuries (HLIs)

Madelyn Rojas Castro, Marina Travanca, Marta Avalos, David Conesa, Emmanuel Lagarde

1 Bordeaux Population Health Research Center INSERM U1219, Injury epidemiology, transport, occupation (IETO).
2 Departamento de Investigación Operativa, Universidad de Valencia

18 January 2018