Epidemiology of Foreign Body Injuries in Children using a Data-Driven Bayesian Network

Dario Gregori, Silvia Snidero, Cecilia Scarinzi, Paola Berchialla

Department of Public Health, University of Torino
Injuries to Children Research Group - www.I2Crg.org

The problem: severity of FB jnuries and product characteristics

- Food items and unsafe consumer products are involved in thousand of injury in children caused by inhalation, ingestion, aspiration and insertion of foreign bodies.
- In general, probabilistic methods enable the characterization of the dimensions and the shape of the objects involved in injuries (Rimell, Thome Jr et al. 1996; Stool, Rider et al. 1998; Rider, Milkovich et al. 2000)
- The task is challenging because typically it is not known a priori
  - which aspects of product characteristics are relevant to the response;
  - the mathematical form of the relation among variables and response probabilities;
  - how unobserved variables can affect the observed relation (Cox 2002).

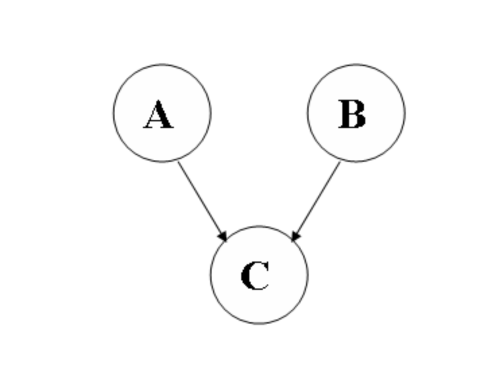
Our work: move toward more complex models

- A successful model useful for Quantitative Risk Assessment should be able to:
  - combine in the probabilistic framework different types of evidence including both subjective beliefs and objective data,
  - overturn previous belief in the light of the new information being received, and
  - make predictions also in case of incomplete data.
- The proposal is to use a Bayesian Network Model as a candidate to satisfy all such targets
- The BN model is thus compared with Artificial Neural Network, Regression Trees, Logistic regression to assess its performance

What is a Bayesian Network?

- A BN is a graphical representation of the joint probability distributions over a set of random variables

P (Severe Injury, All characteristics)



- It consists of a series of nodes representing variables connected by arrows forming a graph that has no cycles.
- The arcs specify the independence assumptions that must hold between the random variables: if no arc, then conditional independence occur
- Learning a BN from data involves the tasks of:
  - structure learning, that is, identifying the graphical structure of the network, and
  - parameter learning, that is, estimating the conditional probability distributions to be associated with the network's graph (Pearl 2000).

Estimation of a BN

- Among the algorithms created for performing structural learning, the K2 algorithm (Glymour and Cooper 1999) was chosen.
  The K2 algorithm is a widely used procedure which is based on a scoring function and a search procedure:
  - The scoring function tries to find a graph that best represents the data according to a specific criterion, in combination with a search method, which measure the goodness of each explored structure from the space of feasible solutions.
  - During the exploring process, the scoring function is applied to evaluate the fitness of each candidate structure to the data.
- Since the BN was set up to constantly receive new cases, we have chosen to treat it as an adaptive net giving to more recent cases a higher weight than to the older ones.
  As a result we built a BN that while receiving cases and updating information on foreign body features (size and shape) was able to quickly respond to changing product safety issues.

The Susy Safe Registry

- The Susy Safe Registry (Gregori 2006) collected data on FB injuries in children aged 0-14 according to the International Classification Disease ICD9-CM 931-935.
- A total of 7296 cases were registered in one Pakistani and 28 European hospitals at the end of March 2007.
- Data encompassed four main aspects of the FB injuries:
  - the characteristics of the children (age, gender);
  - the characteristics of the object (shape, consistency and dimensions);
  - circumstances of injury (presence of parents, activity);
  - hospitalization’s details.

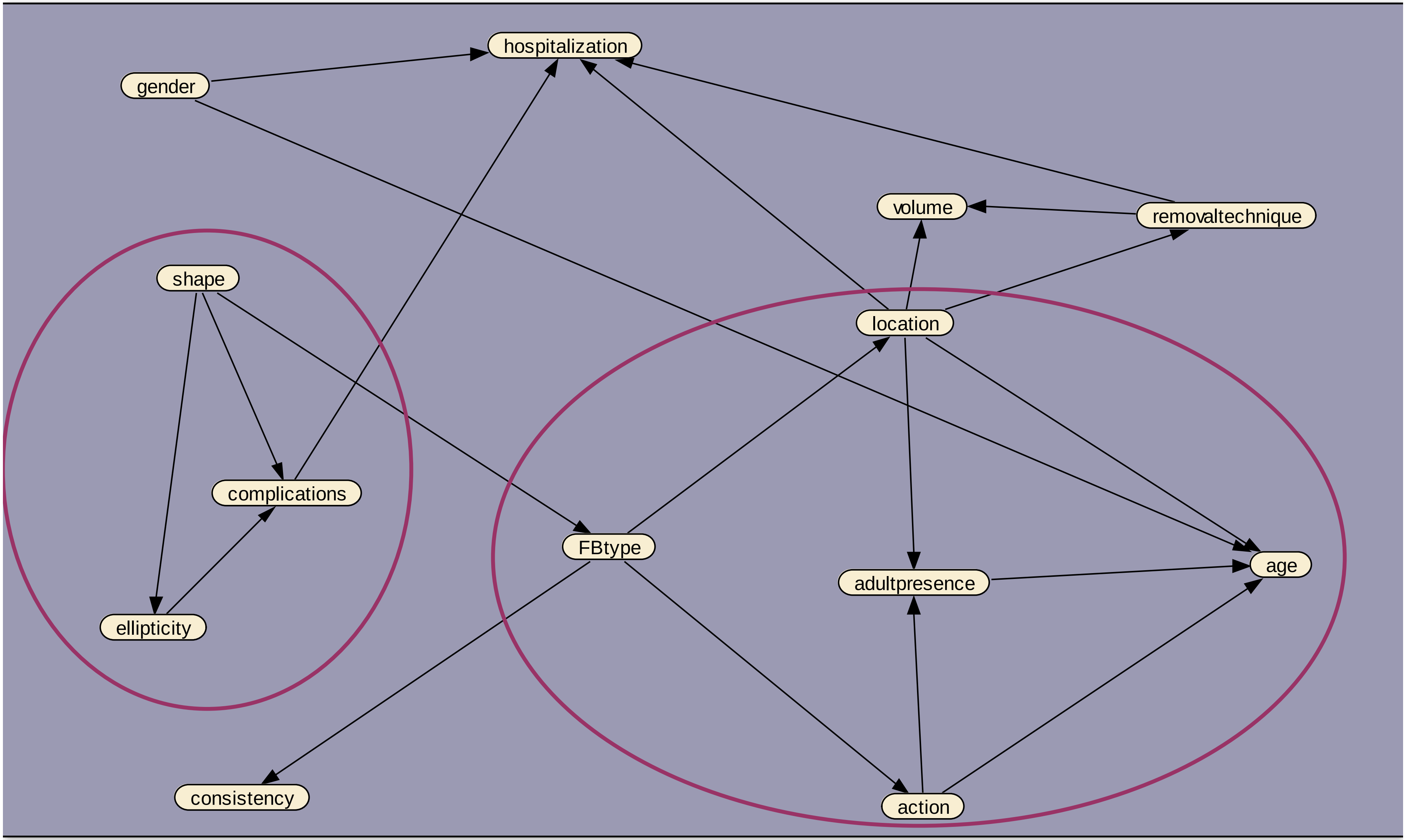
Department of Public Health and Microbiology
University of Torino

Variables used in the models

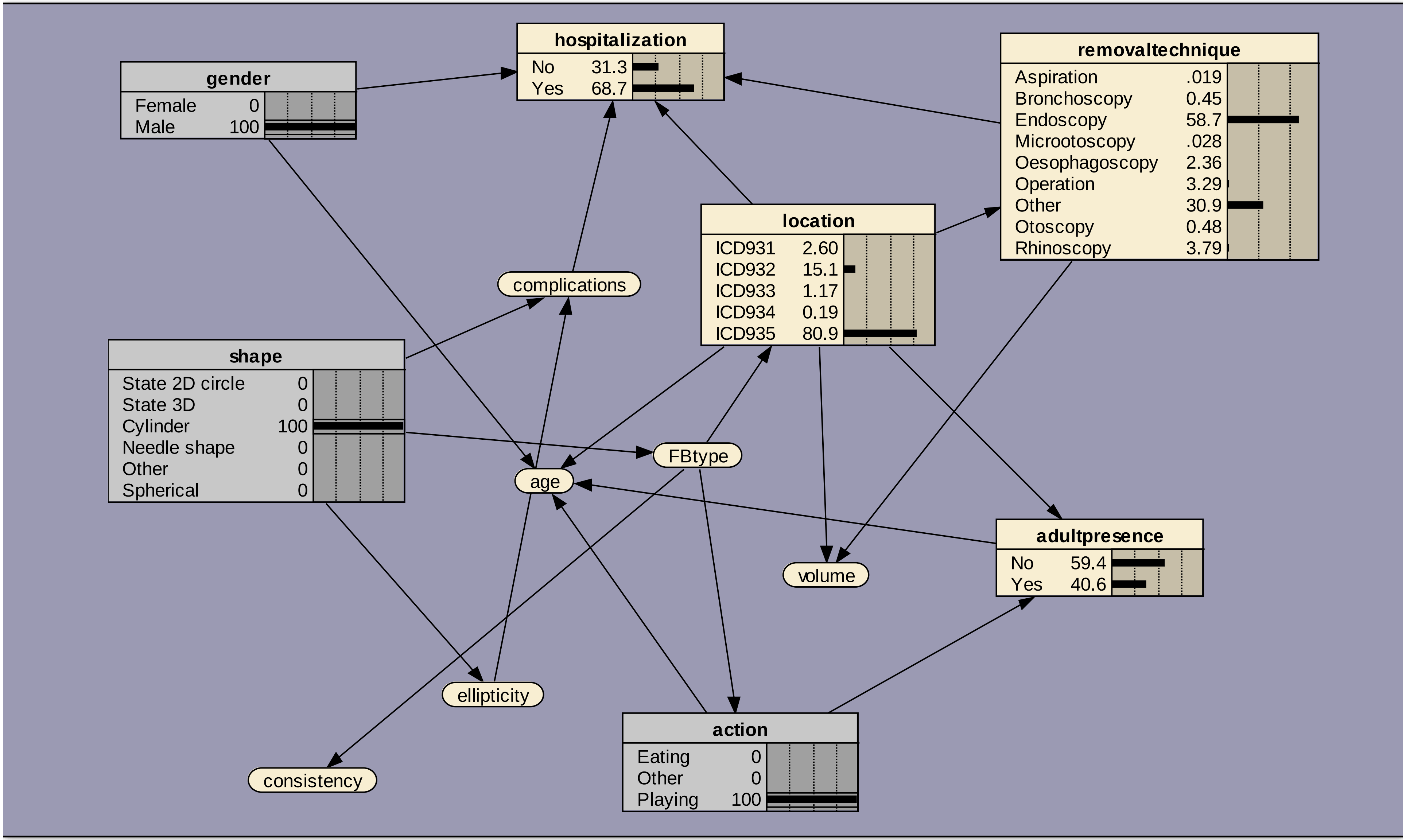
Node description	Variable description	State description
Age	Continuous	Age class: 0-1, 2, 3-4, 5-6, 6-14
Gender	Discrete	Female, Male
Location	Discrete	ICD931-935
Hospitalization	Discrete	No, Yes
Complications	Discrete	No, Yes
Removal tecnique	Discrete	Aspiration, Bronchoscopy, Endoscopy, Operation, Microotoscopy, Otoscopy, Other
Foreign body type	Discrete	Accessorize, arthropod, battery, bean and pea, bone, button, capo, coins; cotton, earplug, fruit and stone; jewellery; metal; nut; other inorganics; other organics; papers; pearl, ball and marble; pebble; pins and needle; plastic; stationery; stick, sweet, toys
Shape	Discrete	2D circle; 3D; Cylinder; Needle shape; Other; Spherical
Consistency	Discrete	Conforming, Rigid, Semi-rigid
Ellipticity	Continuous	1, score from 1 to 2, greater than 2
Volume	Continuous	Score up to 33.5, score from 33.5 to 65.4, score from 65.4 to 140, score from 140 to 400, up to 4710
Adult Presence	Discrete	Adult absent, Adult present
Activity before accident	Discrete	Eating, Playing, Other

Severe Injury (DTI, 1999) is that requiring at least one day of hospitalization

The structure of the Bayesian Network



Probability Estimates from a BN: example



Department of Public Health and Microbiology
University of Torino

Usage of BN in risk assessment

		Observation pattern					The most probable ICD	The most probable removal technique	Probability of hospitalization
foreign body type	age	gender	shape	consistency	volume	ellipticity			
batteries	2	m	cylinder	rigid	70	1	935 (90%)	endoscopy (81%)	91%
pebble	1	m	3D	rigid	140	1.4	932 (71%)	endoscopy (43%)	91%
plastic	3	f	3D	conforming	95	>2	932 (61%)	endoscopy (39%)	34%
fish bone	5	m	needle shape	conforming	140	>2	934 (52%)	endoscopy (41%)	65%
pearls	6	f	spherical	rigid	>400	1	934 (53%)	endoscopy (38%)	62%
stationery	4	m	cylinder	conforming	33.5	>2	934 (57%)	endoscopy (43%)	67%
toy	2	f	spherical	rigid	102	1.3	933 (57%)	other (39%)	48%
nut	6	m	3D	rigid	200	1	934 (95%)	bronchoscopy (62%)	93%
button	3	m	2D circle	rigid	>40	2	935 (64%)	other (65%)	41%
stick	4	m	needle shape	rigid	NA	NA	933 (92%)	other(62%)	21%

Sensitivity analysis

Node	Entropy reduction	%	Variance of Beliefs	%
location	0.48836	50.3	0.1440832	60.1
removaltechnique	0.2899	29.9	0.0855542	35.7
fbmacro	0.2333	24	0.0726672	30
shape	0.03803	3.92	0.0125463	5.23
consistency	0.02601	2.68	0.0080859	3.37
volume	0.01357	1.4	0.0044153	1.84
action	0.01298	1.43	0.0042976	1.79
age	0.00945	0.98	0.0030165	1.26
ellipticity	0.00592	0.61	0.0019679	0.82
adultpresence	0.00151	0.16	0.000502	0.21
gender	0.00015	0.02	0.0000504	0.02
complications	0.00005	0.005	0.0000169	0.007

Entropy is the capability to change the posterior probability of a targeted node they have when new evidence is entered.
The target node chosen in the severity of the injury

Bayesian Network Performance

	AUC	SENS	SPEC
BN	92.31% (89.94-94.68)	95.19%(93.1-97.28)	90.06%(88.08-92.04)
LR	87.03%(84.99-89.05)	89.2%(86.5-91.9)	83.1%(80.75-85.45)
MLP (2 layers)	78.67%(76.64-80.7)	85.33%(81.73-88.93)	54.47%(50.76-58.18)
MLP (1 layer)	74.45%(71.25-77-65)	81.42%(77.52-85.32)	55.16%(51.51-58.8)
CT	72.29%(70.27-74.31)	96.14%(94.06-98.22)	41.74%(39.58-43.9)

Area under the ROC curve was used to assess the performance of the models:

Bayesian Network (BN);
Logistic Regression (LR),
Multilayer Perceptron (MLP) with 2 and 1 hidden layers;
Classification tree (CT)

Final remarks

- Our analysis showed that BN outperformed all methods in terms of accuracy.
- Moreover, advantage of BNs relies on the fact the complex causal relationships among factors are explicated in a graphical model which incorporates uncertainty via the conditional probability associated to each node.
- In addition, relationships among variables are discovered by learning algorithms but at the same time they are allowed to be modified to embody prior knowledge.
- BN model gave a picture of the influence of critical factors on the injury severity. Also it allows for the comparison of the effect of different foreign body characteristics (volume, foreign body type, shape, consistency) and children features (age, gender) on the risk to experience a hospitalization.

Department of Public Health and Microbiology
University of Torino