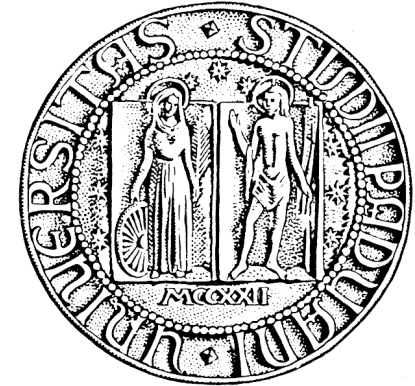


Investigating Complex Dependencies in Biochemical Data Using Bayesian Networks: a Reanalysis of the Fetal Growth Restriction Study



**Francesca Foltran¹, Paola Berchiolla², Sergio Bernasconi³,
Enzo Grossi⁴, Dario Gregori⁵, Maria Elisabeth Street⁶**

¹Department of Surgery, University of Pisa, Italy ²Department of Public Health and Microbiology, University of Torino, Italy ³Department of Pediatrics, University of Parma, Italy ⁴Centro Diagnostico Italiano, Milan, Italy ⁵Department of Environmental Medicine and Public Health, University of Padova, Italy ⁶Department of Pediatrics, University of Parma, Italy

Fetal

**complexities of
the
phenomenon**

Growth

**complexities
in statistical
modelling**

Restriction

Recent important papers (Street et al., 2008) showed that joining advanced statistical techniques with up-to-date databases including **biochemical (cytokines and IGF system) and clinical information, lead to relevant improvements in managing the complexity of such phenomenon**

In combination with non-parametric, distributional-free statistical approaches like **ANNs, it has been shown that **IL6, IGF2 and IGFBP2 concentrations in placental lysates** are all associated in complex, non-linear relationships with fetal growth.**

ANN

very useful for exploratory data analysis, as shown in some recent papers, their predictive capability is often limited by the small samples size and the absence of a predefined set of probability distributions to which refer to

BN

Bayesian approaches, which rely on knowledge of the conditional probability distributions governing the phenomenon, are known to have the best performances in predictive exercises providing that such knowledge is built on solid grounds.

A

to apply a BN model
to a known set of data on fetal growth restriction,
previously analyzed using Artificial Neural Networks (ANN)

by Street et al. in 2008*, in order to

(A) understand if BN were able to improve the previous
predictive performances

(A) verify if models implemented with various techniques
give similar

interpretation for the same covariates.

M

* M.E. Street, et al, Changes in interleukin-6 and IGF system and their relationships in placenta and cord blood in newborns with fetal growth restriction compared with controls, Eur. J. Endocrinol. 155 (2006) 567-574.

21 Intra-Uterine Growth Restricted and

25 Appropriate for Gestational Age

pregnancies

were followed.

M&M

Information was collected on: maternal age,

gestational age, neonatal gender, previous

gynaecological history, protein content, IGF2,

IGFBP2, IL6, and TNF α concentrations in

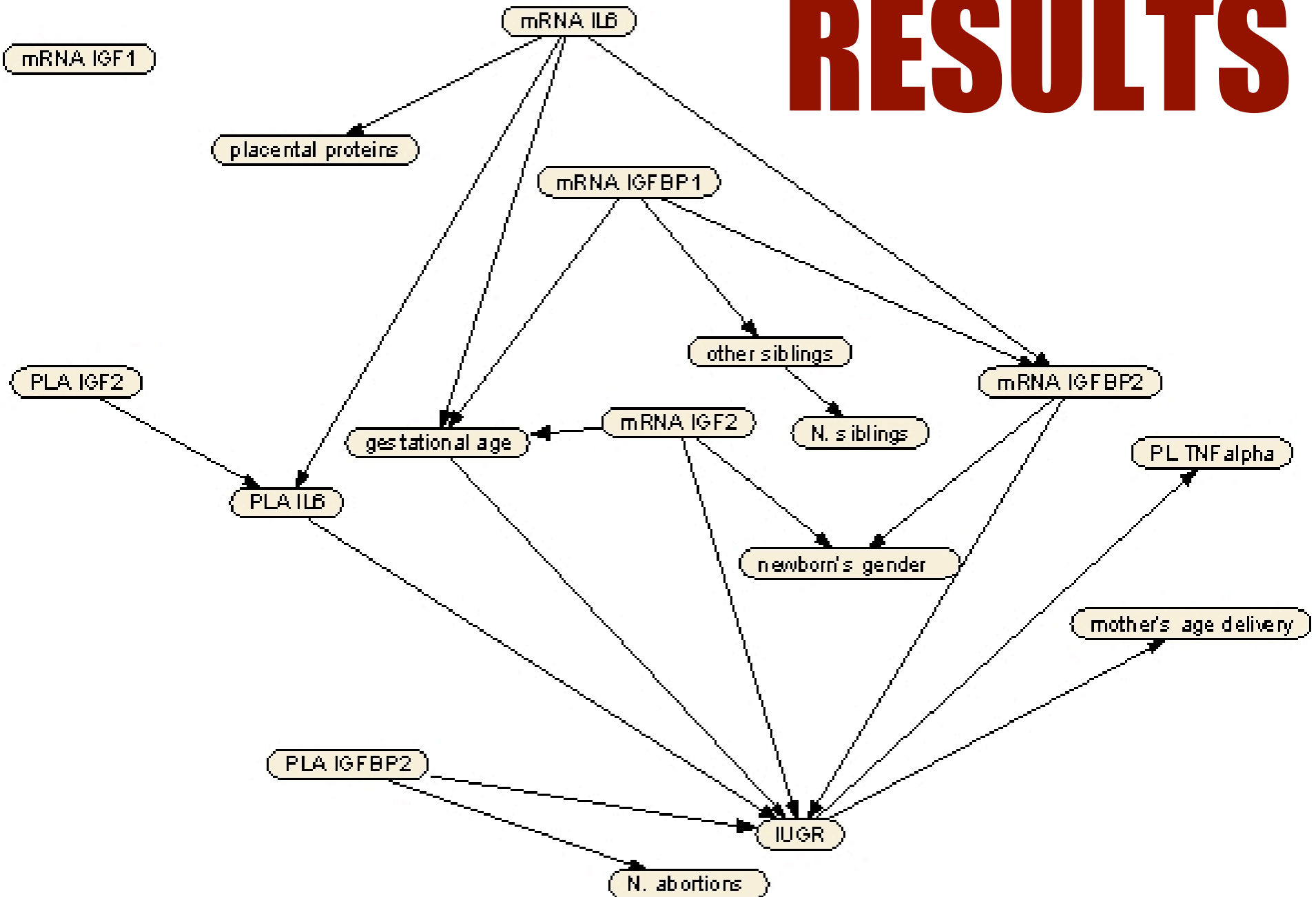
placental lysates were measured.

M&M

A Bayesian Network model and a hybrid forecasting system combining classification trees and naïve Bayes classifier were implemented.

Predictive ability of BN, hybrid system and ANN were compared.

RESULTS



Models	Mean Accuracy (%)
ANN Models	
Random Protocol	
BM	80
BP	80
SN	76
Optimized Protocol	
BM 7 VAR	88
BM 9 VAR	84
BP 7 VAR	88
BP 9 VAR	88
SN 7 VAR	84
SN 9 VAR	80
Bayesian Network	89.62
Hybrid forecasting system	73.91

RESULTS

Compared with other techniques, BN showed a better percentage mean accuracy. BN were able to provide an improvement with respect to ANN in terms of predictive accuracy.

BN confirmed their utility in discovering and graphically representing variable connections and hierarchy, particularly where, as in Fetal Growth Restriction, many interactions among predictors exist.

Conclusions