

# Using probabilistic graphical models to enhance the prognosis of health-related quality of life in adult survivors of critical illness

Cláudia Camila Dias, Cristina Granja,  
Altamiro Costa-Pereira  
CINTESIS  
Faculty of Medicine, University of Porto  
Porto, Portugal  
camila@med.up.pt

João Gama  
LIAAD - INESC TEC  
Faculty of Economics, University of Porto  
Porto, Portugal  
jgama@fep.up.pt

Pedro Pereira Rodrigues  
CINTESIS & LIAAD - INESC TEC  
Faculty of Medicine, University of Porto  
Porto, Portugal  
pprodriques@med.up.pt

**Abstract**—Health-related quality of life (HR-QoL) is a subjective concept, reflecting the overall mental and physical state of the patient, and their own sense of well-being. Estimating current and future QoL has become a major outcome in the evaluation of critically ill patients. The aim of this study is to enhance the inference process of 6 weeks and 6 months prognosis of QoL after intensive care unit (ICU) stay, using the EQ-5D questionnaire. The main outcomes of the study were the EQ-5D five main dimensions: mobility, self-care, usual activities, pain and anxiety/depression. For each outcome, three Bayesian classifiers were built and validated with 10-fold cross-validation. Sixty and 473 patients (6 weeks and 6 months, respectively) were included. Overall, 6 months QoL is higher than 6 weeks, with the probability of absence of problems ranging from 31% (6 weeks mobility) to 72% (6 months self-care). Bayesian models achieved prognosis accuracies of 56% (6 months, anxiety/depression) up to 80% (6 weeks, mobility). The prognosis inference process for an individual patient was enhanced with the visual analysis of the models, showing that women, elderly, or people with longer ICU stay have higher risk of QoL problems at 6 weeks. Likewise, for the 6 months prognosis, a higher APACHE II severity score also leads to a higher risk of problems, except for anxiety/depression where the youngest and active have increased risk. Bayesian networks are competitive with less descriptive strategies, improve the inference process by incorporating domain knowledge and present a more interpretable model. The relationships among different factors extracted by the Bayesian models are in accordance with those collected by previous state-of-the-art literature, hence showing their usability as inference model.

**Keywords**—quality of life; intensive care; Bayesian networks; critically ill patients; EQ-5D.

## I. INTRODUCTION

This study was built on two main concepts: health-related quality of life after critical illness, and Bayesian inference for outcome prognosis.

### A. Health-related quality of life after critical illness

During the past decade, health related quality of life (HR-QoL) has emerged as an important outcome in clinical research, and decision making in critical care patients has become highly influenced by the appreciation of patients' future HR-QoL. In critically ill patients, prolonging life could

result in a health outcome that may be considered worse than death [1]. In this setting, the ideal outcome for the patient is to return to their pre-existing state or that expected for a person of the same age and medical condition [2]. The interest in patients' perspectives in the evaluation of health care has led to the development of numerous measures. Measures of outcome have thus changed from objective measures toward subjective measures of functional status and HR-QoL, with data directly collected from patients [2].

HR-QoL can be conceived as a subjective concept, primarily reflecting individual patient attitude. Measuring HR-QoL is in essence evaluating the health state of individuals, both mental and physical, together with their own sense of well-being [3]. HR-QoL is more complex to record and analyze than survival or other objective measures due to a number of reasons: first, the timing of data collection from critically ill patients; then the choice of particular dimensions to be included; difficulties in consensus over the instrument to be used; and finally, the interpretation of the results [4,5]. EQ-5D has been recommended as one of the questionnaires suitable to apply to critical care survivors for measuring quality of life [6], and has been validated in critically ill patients [7]. It is a simple questionnaire, of easy comprehension and quick answers. The EQ-5D comprises two parts: a self-reported description of health problems according to five dimensions: mobility, self-care, usual activities, pain/discomfort and anxiety/depression; and the EQ-VAS, a self-rated health status using a visual analogue scale (VAS), to record perceptions of the current overall health state [8, 6].

### B. Bayesian networks for decision support

The definition of clinical decision support systems is now a major topic since it may help the diagnosis, the prognosis of mortality, the prognosis of quality of life, or even therapeutic options. However, the complicated nature of real-world biomedical data has made it necessary to look beyond traditional biostatistics [9] without losing the necessary formality. For example, naive Bayesian approaches are closely related to logistic regression [10]. Traditional statistical methods require that the model structure is given and only probabilistic information is learned from biomedical evidence, in the form of data, whereas machine-learning approaches

enable that both the structure of the models and the probabilistic information are evidence-based [11,12]. Bayesian statistical methods allow taking into account prior knowledge when analyzing data, turning the data analysis into a process of updating that prior knowledge with biomedical and health-care evidence [12], which offers a general and versatile approach to capturing and reasoning with uncertainty in medicine and health care [9]. Bayesian networks can be seen as an alternative to logistic regression where statistical dependence and independence are not hidden in approximating weights, rather explicitly represented by links in a network of variables [12]. On a general basis, a Bayesian network represents a joint distribution of one set of variables, specifying the assumption of independence between them, with the inter-dependence between variables being represented by a directed acyclic graph. The next step is how to calculate the new probabilities when new evidence is introduced in the network. This process is called inference and works as follows. Each variable has a finite number of categories greater than or equal to two. A node is observed when there is knowledge about the state of that variable. With the joint probabilities we can calculate the marginal probabilities of each unobserved variable [9].

### C. Bayesian networks in intensive care

Given their improved management of uncertainties, Bayesian networks have been successfully applied in healthcare domains [9], namely in different scenarios of intensive care [10]. Sakellaropoulos and Nikiforidis [13] developed in 2000 a system based on Bayesian belief networks for prognosis at 24 hours of head-injured patients. The performance of a Bayesian network was compared to other systems, including a simple belief network, and a human expert (neurosurgeon). The authors concluded that the performance of the developed expert system is not different from that of the neurosurgeon (expert) and was better than a simple belief network, suggesting that the system can be helpful for the prognostic performance of non-neurosurgeon intensive care unit (ICU) clinicians. Later on, Sierra et al [14] developed a bi-level multi-classifier using Bayesian networks for predicting the survival of ICU patients. The authors concluded that their approach outperformed traditional machine learning methods in the studied problem. To help in the diagnosis of ventilator-associated pneumonia (VAP), Schurink et al. [15] developed a decision support system based on Bayesian networks. This study compared the Bayesian network system's diagnosis of VAP with the reference diagnosis for each patient considered; the system's performance was good and the authors concluded that this system might become a useful tool to assist physicians in the ICU. Celi et al. [16] used a Bayesian network to predict fluid requirements in the ICU on day 2, given the physiological data from the previous 24 hours, developing a system that works as complementary knowledge to help the clinician. In 2009, Crump et al. [17] used vital sign data streams and known outcomes to train a Bayesian network and create the timeseries models, complemented by a rule-based trending, in order to predict patients status in the ICU. The main advantages of the system were the improvement in the monitoring of ICU patients, given the ability to provide high-sensitivity alerts,

fewer false alarms, and earlier interventions, thus enhancing the possibility to assist clinicians in their daily practice.

### D. Bayesian network inference of HR-QoL

Bayesian networks have become even more popular in health care due to their ability to deal with uncertain knowledge [11,10]. The measurement of quality of life is embedded in the same uncertainty as as other measurement in health, diagnosis, treatment or even mortality. That is why a Bayesian approach could be helpful since it controls the complexity of the model and gives us graphical representations that can be easily understood. This study aims to enhance the prognosis inference process of quality of life after ICU stay, by means of a probabilistic graphical model. For this purpose, we developed and tested several Bayesian network models which were then critically evaluated by an expert clinician.

## II. MATERIALS AND METHODS

All patients admitted in ICU of an urban hospital were considered for inclusion. After ICU stay, survivors went to a follow-up consultation where EQ-5D was applied. Given the change in the ICU follow-up protocol that occurred in that hospital, two independent samples were considered in different points of time: 503 (out of 1282 patients admitted between 1997 and 2001) had a 6 months follow-up and were retrospectively collected, and 60 (out of 560 admitted during 2005) had a 6 weeks follow-up and were prospectively collected. Since there was no 6 weeks follow-up prior to 2005, and most patients included during 2005 were lost to follow-up at 6 months, no paired comparison was possible. All patients participated in the follow-up study voluntarily. Hospital Pedro Hispano Ethics Committee approved this study. The methodology described in this section has been supervised by an intensivist clinical expert.

### A. Studied variables and outcomes

Data collection included patient-related background variables (age, gender, main activity and previous health state), ICU-related variables (admission diagnostic categories, APACHE II, Glasgow coma scale, and length of ICU stay - ICU LOS); and EQ-5D-related variables (mobility, self-care, usual activities, pain, anxiety/depression). The main outcomes of this study are the EQ-5D five main dimensions. Continuous variables, such as age, length of stay (LOS) and APACHE II, were discretized using clinical expert-supervised intervals. Outcome variables were also dichotomized on patients without and with problems, disregarding the severity of those problems. Clinically-judged highly correlated variables were also excluded from the study, such as the existence of sepsis on admission, the occurrence of sepsis after admission, and acute respiratory distress syndrome (ARDS) since this information was already considered in the diagnosis variable.

### B. Model building

Bayesian networks were learned using the ELVIRA software [18], for each of the five EQ-5D dimensions, for each

of the two periods in analysis. For each of the ten classification problems, three structures were learned corresponding to the application of Naive Bayes [19], TAN [20] and a 2-DBC [21], which differ on the number of conditional dependence allowed among variables (NB: zero dependences, the baseline; TAN: one dependence, baseline extension; and 2-DBC: two dependences, complex network), in order to enable the expert clinician with a wide range of different structures from which to choose. At this point, one of the most important tasks was to give a formal mean of assessing the representativeness of the model towards the real problem, so interaction with the clinical expert was initiated. The expert clinician had the possibility to analyze, evaluate and modify the learned structures in order to reflect the domain knowledge more accurately, according to his own opinion and experience, mostly by targeting artifact relationships which should not be considered. Finally, to assess the generalized accuracy of the learned models, 10-fold cross-validation was used.

### III. RESULTS

First, we present patient characteristics and the prevalence of problems regarding HR-QoL critical illness. After, we present accuracy comparison among different types of Bayesian networks for each prognosis horizon.

#### A. Patients characteristics

The main characteristics of patients for both groups (6 weeks and 6 months follow-up) are shown in Table I. Patients report more problems at 6 weeks than at 6 months, compatible with what was described by Dowdy et al. [22]. There is an improvement of the HR-QoL with time but this increase is not uniform across dimensions. In the mobility dimension, the percentage of patients with no problems at 6 weeks was 31%, going up to 67% at 6 months. In self-care at 6 weeks 52% (the highest value reported at 6 weeks over all dimensions) of patients had no problems, with this value going up to 72% at 6 months. In usual activities 23% of the patients reported no problems at 6 weeks, increasing to 50% at 6 months. As for pain/discomfort dimension, 39% of patients report no problems at 6 weeks, although this value grows up to 56% at 6 months. Finally, for anxiety/depression, at 6 weeks only 34% of the patients reported no problems, going up to 47% at 6 months. It can be observed that the dimension with higher improvement from 6 weeks to 6 months after ICU stay was mobility which rose 36%. The dimension which presented the lowest improvement was anxiety/depression, growing 13%. These differences between physical and mental dimensions had already been enunciated in previous works, using diverse methodologies, which reported concordant results [6,22].

#### B. Bayesian network model comparison

Results from the 10-fold cross validation for all the models are presented in Figure 1 (6 weeks) and Figure 2 (6 months).

TABLE I. MAIN CHARACTERISTICS OF PATIENTS FOR BOTH GROUPS (6 WEEKS AND 6 MONTHS FOLLOW-UP), N (%).

	6 weeks		6 months	
<b>Gender</b>				
Male	31	(52)	270	(57)
Female	29	(48)	203	(43)
<b>Age</b>				
[18-45]	16	(27)	139	(29)
[46-60]	16	(27)	125	(26)
[61-70]	10	(16)	112	(24)
>70	18	(30)	97	(21)
<b>Main Activity</b>				
Employed	23	(38)	112	(24)
Retired	30	(50)	249	(52)
Domestic/seeking work	5	(9)	47	(10)
Other	2	(3)	65	(14)
<b>Minimal education</b>				
No	12	(20)	270	(57)
Yes	48	(80)	203	(43)
<b>Smoking habits</b>				
Smoker	6	(10)	60	(13)
Ex-Smoker	27	(45)	184	(39)
Never smoke	27	(45)	229	(48)
<b>Previous health state</b>				
Healthy	33	(39)	132	(28)
Disease	35	(58)	241	(51)
Disabling disease	2	(3)	100	(21)
<b>Admission</b>				
Scheduled surgery	5	(10)	222	(47)
Non-scheduled surgery	7	(12)	63	(13)
Medical	44	(73)	169	(36)
Trauma	3	(5)	19	(4)
<b>Length of stay</b>				
<=1	2	(3)	187	(40)
2-7	31	(52)	189	(40)
>7	27	(45)	97	(20)
<b>Cancer</b>				
No	54	(90)	357	(75)
Yes	6	(10)	116	(25)

In some dimensions, TAN is significantly better than NB, while being never significantly worse than 2-DBC. Moreover, the networks created by TAN were found to better represent the domain knowledge as perceived by the expert clinician. Hence, in order to combine model accuracy and interpretability with the expert's opinion, the chosen model was the Tree Augmented Naive Bayes. As an example, Figures 3 and 4 present the learned network for the 6-weeks and 6-months mobility prognosis, respectively. Given this, from the chosen networks (TAN) for each prognosis problem, we can inspect the main characteristics of patients reporting less problems.

#### C. Patients with better HR-QoL 6 weeks after ICU stay

Concerning the mobility dimension, male patients, younger, or patients without cancer have higher probability of not having problems. In the same way, for the self-care dimension, male patients, younger, or with lower ICU LOS have higher probability of not having problems. Moreover, the probability of not having problems is also higher in patients with lower APACHE II. Relatively to usual activities, we observed the same characteristics encountered for the self-care dimension. Moreover, in this dimension, retired patients have also higher probability of not having problems. For the pain/discomfort dimension, also male patients, patients with less ICU LOS, or patients without cancer have higher probability of not having problems. Finally, for the anxiety/depression dimension a higher probability of not having problems is reported in male patients, with less ICU LOS or with lower APACHE II.

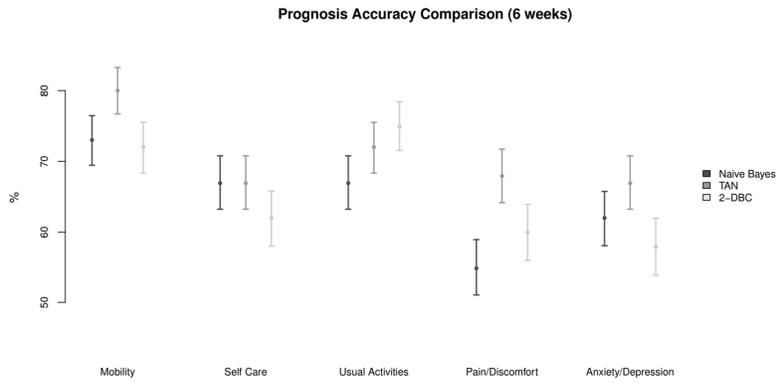


Fig. 1. Mean accuracy comparison among different types of Bayesian networks for the 6 weeks prognosis problem (error bars represent the 95% confidence interval).

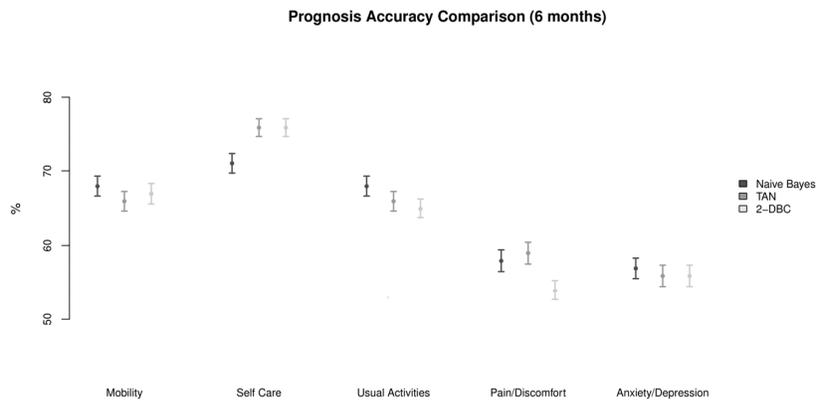


Fig. 2. Mean accuracy comparison among different types of Bayesian networks for the 6 months prognosis problem (error bars represent the 95% confidence interval).

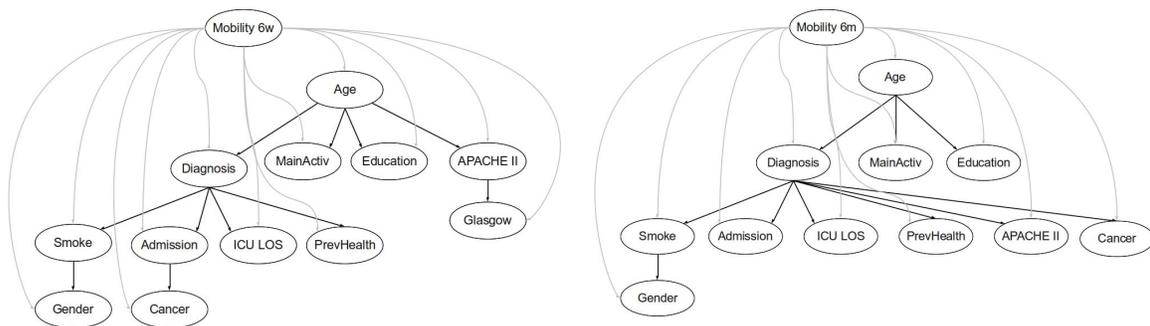


Fig. 3. TAN Bayesian networks for the 6 weeks (left) and 6 months (right) mobility prognosis (lighter arrows represent naïve Bayes connections).

#### D. Patients with better HR-QoL 6 months after ICU stay

For mobility, male patients, younger, with lower ICU LOS and those who were previously healthy have a higher probability of not having problems. Moreover, patients who were admitted for scheduled surgery or by medical reasons, or employed have also a higher probability of not having problems. Focusing on the self-care dimension, the patients who have a higher probability of not having problems are men, younger, with lower ICU LOS, scheduled surgery or the ones who had progressed studies longer than minimum education. Querying usual activities dimension, younger patients, or with lower ICU LOS have a higher probability of not having problems, as well as patients who were admitted by medical reasons or who had progressed studies longer than minimum education. For the pain/discomfort dimension, male patients, younger, with lower ICU LOS, lower APACHE II, admitted by non-scheduled surgery or medical reasons, or those who were previously healthy have a higher probability of not having problems. Concerning the anxiety/depression dimension, male patients, elder ones, patients without cancer, lower APACHE II, admitted by non-scheduled surgery or medical reasons, or retired patients have higher probability of not having problems.

#### E. Inference example

To better assess the enhancement achieved by our approach to the prognosis process, each structure is used to perform the 6 weeks or 6 months prognosis of the quality of life for a given patient. A new patient arrives at ICU and the physician, assuming that the patient will survive, wants to have an indication of the quality of life this patient will report 6 weeks and 6 months after is stay in the unit. Characteristics of the patient: sixty-one years old man, former smoker, with disabling disease, with pneumonia diagnosis, admitted by medical reasons and having an APACHE II value of 22 at the time of admission. One can inspect the Bayesian networks learned for the whole data, though separately for each dimension, introducing the evidence information in the network.

In Figure 5, structured inference for the prognosis at 6 weeks (left) and 6 months (right) is presented for the (exemplar) mobility dimension (other dimensions not shown for sake of readability): the admitted patient has a probability of 15% of not having problems at 6 weeks, going up to 79% at 6 months. Concerning self-care, the same patient has 25% probability of not having problems at 6 weeks, increasing to 82% for the 6 months prognosis. Concerning usual activities, the patient has 48% probability of not having problems at 6 weeks, increasing to 67% for the 6 months prognosis. Concerning pain/discomfort, the patient has 47% probability of not having problems at 6 weeks, increasing to 91% for the 6 months prognosis. Concerning anxiety/depression, the patient has 41% probability of not having problems at 6 weeks, decreasing to 31% for the 6 months prognosis.

## IV. DISCUSSION

This study presents four main findings. First, patients report more problems concerning their HR-QoL at 6 weeks than at 6 months. That is, HR-QoL improves with time after ICU stay, though not uniformly across different dimensions. This finding is consistent with several reports which have been analyzed in a systematic review by Dowdy et al. [22]. Second, factors that influence the HR-QoL of patients admitted in ICU are also identified, the most frequent being age, gender, length of ICU stay, previous health state, educations and APACHE II. Third, Bayesian networks present competitive performance, comparing with other less descriptive models like logistic regression proposed elsewhere in literature. Fourth, generated models expose relationships among variables which are concordant with previous works [6,22,23]: for the 6 weeks prognosis, the youngest and those who had a lower APACHE II have a lower risk of QoL problems; for the 6 months prognosis, with exception for anxiety/depression, the male patients, the younger, those with lower ICU LOS, or low APACHE II have a lower risk of QoL problems. For anxiety/depression, the oldest and retired patients are the ones with a lower risk of QoL problems. A more complex model, including all EQ-5D dimensions, could have been derived but the reduced number of patients included in the study would turn this an extremely hard task.

The main contribution of this work is the development of graphical representation models of uncertain knowledge existent in the characterization of QoL after ICU stay. This representational strategy allows an easy visualization of the model, reinforcing the experts with the possibility of a supported decision, not only by risk factors or discriminating cut points, but mostly supported in the interdependences of studied variables and their causality with evaluated dimensions. Models capturing these processes and methods for using these models are thus called for to support decision making in real-life practice [9]. This study has four main limitations. First, data on HR-QoL before intensive care for each patient is not available. This fact prevents the study of real impact of stay in the ICU. Second, the reduced amount of data at 6 weeks is too restrictive and, thus, decrease accuracy of change of quality of life from 6 weeks to 6 months after ICU stay. Third, no paired comparison was possible between 6 weeks and 6 months which may have biased the comparison due to possible confounding factors. Finally, only one expert clinician was consulted in the study, which might reduce the generalizability of the results.

## V. CONCLUSIONS

This study developed probabilistic graphical models to represent the uncertain knowledge regarding HR-QoL after critical illness. The networks enhance the prognosis of HR-QoL, 6 weeks and 6 months after ICU stay, supporting the clinicians' decision with the visualization of the interdependences among different factors. It adds to critical

care knowledge by contributing to the prognosis of the HR-QoL after critical illness, presenting an inference tool to support the clinicians in their daily practice.

REFERENCES

[1] DL Patrick, HE Starks, RF Uhlman, and RA Pearlman. Measuring preferences for health status worse than death. *Medical Decision Making*, 14:9:18, 1994.

[2] NA Black, C Jenkinson, JA Hayes, K Vella, KW Rowan, K Daly, and S Ridley. Review of outcome measures used in adult critical care. *Critical Care Medicine*, 19:2119–2124, 2001.

[3] RR Fernandez, JJ Cruz, and GV Mata. Validation of a quality of life questionnaire for critically ill patients. *Intensive Care Medicine*, 22:1034–1042, 1996.

[4] TM Gill and AR Feinstein. A critical appraisal of the quality of quality-of-life measurements. *JAMA*, 272:619–626, 1994.

[5] JA Hayes, NA Black, C Jenkinson, JDYKM Rowan, and KDS Ridley. Outcome measures for adult critical care: a systematic review. *Health Technology Assessment*, 4:19–81, 2000.

[6] Cristina Granja, Armando Teixeira-Pinto, and Altamiro Costa-Pereira. Quality of life after intensive care - evaluation with eq-5d questionnaire. *Intensive Care Medicine*, 28:898–907, 2002.

[7] LF Azevedo, C Dias, A Costa-Pereira, A Amaro, E Gomes, and C Granja. Validity and reliability of eq 5d questionnaire in critically ill patients. *Critical Care Medicine*, 35(12 suppl):A172, 2007.

[8] The EuroQol Group. Euroqol-a new facility for the measurement of health-related quality of life. *Healthy Policy*, 16:199–208, 1990.

[9] Peter Lucas, LC van der Gaag, and AA Hanna. Bayesian networks in biomedicine and health-care. *Artificial Intelligence In Medicine*, 30:201–214, 2004.

[10] CAM Schurink, PJF Lucas, IM Hoepelman, and MJM Bonten. Computer-assisted decision support for the diagnosis and treatment of infectious diseases in intensive care units. *Lancet Infectious Diseases*, 5:305–312, 2005.

[11] Tom Mitchell. *Machine Learning*. McGraw-Hill, 1997.

[12] Peter Lucas. Bayesian analysis, pattern analysis, and data mining in health care. *Current Opinion in Critical Care*, 10:399–403, 2004.

[13] GC Sakellaropoulos and GC Nikiforidis. Prognostic performance of two expert systems based on bayesian belief networks. *Decision Support Systems*, 27:431–442, 2000.

[14] B Sierra, N Serrano, P Larrañaga, EJ Plasencia, I InzaI, JJ JiménezJ, P Revuelta, and ML Mora. Using bayesian networks in the construction of a bi-level multi-classifier. a case study using intensive care unit patients data. *Artificial Intelligence in Medicine*, 22:233–248, 2001.

[15] CAM Schurink, S Visscher, PJF Lucas, HJ van Leeuwen, E Buskens, RG Hoff, AIM Hoepelman, and MJM Bonten. A bayesian decision-support system for diagnosing ventilator-associated pneumonia. *Intensive Care Medicine*, 33:1379–1386, 2007.

[16] LA Celi, LC Hinske, G Alterovitz, and P Szolovits. An artificial intelligence tool to predict fluid requirement in the intensive care unit: a proof-of-concept study. *Critical Care*, 12:151, 2008.

[17] C Crump, S Saxena, B Wilson, P Farrell, A Rafiq, and CT Silvers. Using bayesian networks and rule-based trending to predict patient status. In *Intensive Care Unit. AMIA Symposium Proceedings*, pages 124–128, 2009.

[18] TE Consortium. An environment for creating and using probabilistic graphical models. *Proceedings of First European Workshop on Probabilistic Graphical Models*, pages 1–11, 2002.

[19] P Domingos and M Pazzani. On the optimality of the simple bayesian classifier under zero-one loss. *Machine Learning*, 29:103–130, 1997.

[20] N Friedman, D Geiger, and M Goldszmidt. Bayesian network classifiers. *Machine Learning*, 29:131–163, 1997.

[21] M Sahami. Learning limited dependence bayesian classifiers. In *Proceedings of the Second International Conference on Knowledge Discovery and Data Mining*, 1996.

[22] DW Dowdy, MP Eid, A Sedrakyan, PAM Tello, PJ Pronovost, MS Herridge, and DM Needham. Quality of life in adult survivors of critical illness: a systematic review of the literature. *Intensive Care Medicine*, 31:611–620, 2005.

[23] FG Lizana, DP Bota, MD Cubber, and JL Vincent. Long term outcome in icu patients: What about quality of life. *Intensive Care Medicine*, 29:1286–1293, 2003.

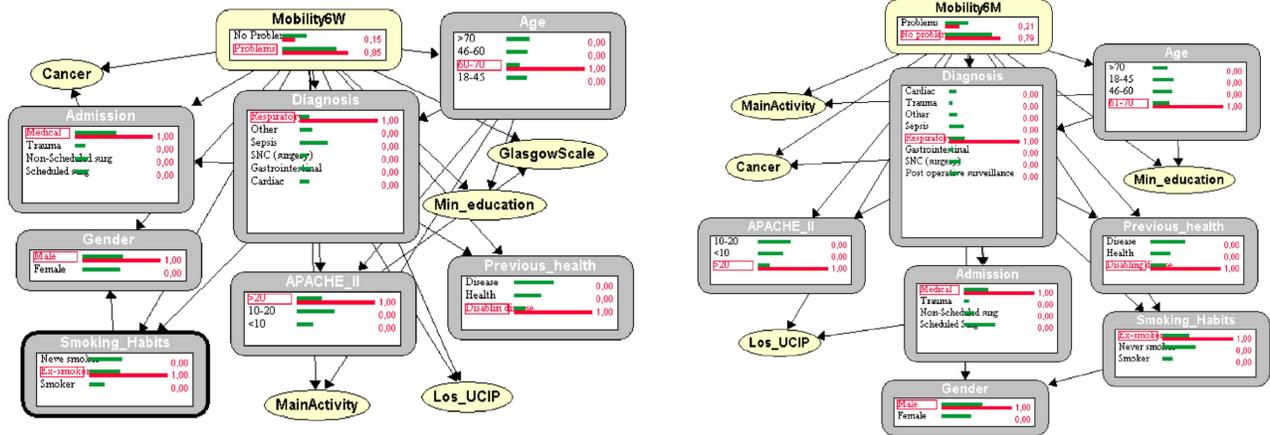


Fig. 4. Mobility inference example for a sixty-one years old man, former smoker, with disabling disease, with pneumonia diagnosis, admitted by medical reasons and having an APACHE II value of 22 at the time of admission. In all figures non-observed variables are represented by collapsed nodes. All other nodes are observed variables. Inside observed variables, and for each category, the top bar represents the prior probability, while the bottom bar represents the new evidence inserted in the network. If we expand the former category (on top) we can see on the bottom bar of each result what probability that this individual patient will have problems in his quality of life.