

# Automatic Early Risk Detection of Possible Medical Conditions for Usage Within an AMI-System

H. Joe Steinhauer and Jonas Mellin

**Abstract** Using hyperglycemia as an example, we present how Bayesian networks can be utilized for automatic early detection of a person's possible medical risks based on information provided by unobtrusive sensors in their living environments. The network's outcome can be used as a basis on which an automated AMI-system decides whether to interact with the person, their caregiver, or any other appropriate party. The networks' design is established through expert elicitation and validated using a half-automated validation process that allows the medical expert to specify validation rules. To interpret the networks' results we use an output dictionary which is automatically generated for each individual network and translates the output probability into the different risk classes (e.g., *no risk*, *risk*).

**Keywords** Ambient assisted living • Bayesian networks • Automated diagnosis

## 1 Introduction

A major part of the HELICOPTER (Healthy Life support through Comprehensive Tracking of individual and Environmental Behaviors, <http://www.helicopter-aal.eu>) is to develop information and communication technology (ICT) - based solutions that assist self-sufficient elderly people in early detection of the possible development of medical conditions, such as hyperglycemia or heart failure. The reason for this is to prevent complications arising from the medical conditions if they are not detected early enough. The main contribution of the HELICOPTER project is therefore the part of the system that can detect the risk of certain medical conditions

---

H. Joe Steinhauer (✉) · J. Mellin  
University of Skövde, Skövde, Sweden  
e-mail: [joe.steinhauer@his.se](mailto:joe.steinhauer@his.se)

J. Mellin  
e-mail: [jonas.mellin@his.se](mailto:jonas.mellin@his.se)

based on sensor readings and that we call the *automatic triage*. Its system architecture is closer described in [1]. The automatic triage should be as unobtrusive as possible and should not bother the patient with unnecessary interventions. Health surveillance for the automatic triage is achieved by deploying unobtrusive sensors (e.g., infrared sensors, pressure sensors, power meters, body weight scales, and food-inventory tools) and wearable sensors (e.g., fall detectors, individual identification tags). All data collected from these heterogeneous sensors are then interpreted within a data analysis engine in order to deduce the patient's current risk of developing an acute medical condition (e.g., hyperglycemia or hypotension).

In this project it is our objective to utilize well established existing methods, in this case Bayesian networks, deploy them within a case study in order to develop the specific network designs necessary for each medical condition, and validate the resulting networks. The remainder of this paper is organized as follows: In Sect. 2 we explain how a Bayesian network for the use in the automatic triage can be developed in cooperation with a medical expert. After that, in Sect. 3, we describe how the results of Bayesian networks are validated. Last, but not least, we discuss our work and give some suggestions for future work in Sect. 4.

## 2 Bayesian Networks for Automatic Triage Diagnosis

Generally, a diagnosis will be determined on available evidence  $E$  and is defined as in e.g. [2]:

$$d^* = \operatorname{argmax}_{d \in D} \Pr(d|E) \quad (1)$$

where  $D$  is the set of possible diagnoses, and  $d^*$  stands for the subset of diagnoses that have been chosen. Bayesian networks [3] have been used in the area of medical diagnostic reasoning, prognostic reasoning, treatment selection, and for the discovery of functional interactions, since the beginning of 1990 [2, 4, 5]. Some early examples can be found in [4, 6–8]. More recently, Bayesian networks are also applied in home care applications e.g. [9].

A Bayesian network [3] or causal probability network [6] is a graphical representation of a probability distribution over the set of random variables. Probabilistic inference can be done with Bayes rule (see e.g. [10]), which in our domain, where we want to infer the probability of a disease given that we observe one or several symptoms that are often caused by the disease, can be defined as:

$$P(\text{disease}|\text{symptom}) = \frac{P(\text{symptom}|\text{disease})P(\text{disease})}{P(\text{symptom})} \quad (2)$$

Due to their graphical representation, Bayesian networks are relatively easy to understand and to create and can therefore be used, developed, and interpreted by

domain experts [9]. They can often be seen as a model of cause-effect relationships [4] whereby their structure and the underlying probability distribution can be learnt from data or be created by hand. Thus qualitative and quantitative knowledge can be mixed [6]. Furthermore, uncertain knowledge can be modeled within a Bayesian network and missing data can be handled during the diagnosis process, which can successively be updated when more evidence becomes available [7].

Before we started to develop the automatic triage system, we also considered alternative evidential frameworks, such as evidence theory [11] and subjective logic [12], but decided together with the medical expert to use Bayesian networks based on four criteria: (1) the framework chosen needs to be able to express everything that is relevant for the task, (2) the design and inner workings of the framework should be easy to understand for the medical expert, (3) the framework should be considerably mature and (4) tools for developing the networks should be available.

In our project, as there is no data set available from that the Bayesian network could be automatically constructed and tested, it needs to be built by hand, whereby knowledge about the domain of diagnosing medical conditions is provided by a medical expert. [2] describes that the construction of a Bayesian network by hand usually involves five stages, which can be iterated during the construction process: (1) relevant variables need to be chosen; (2) relationships among the variables need to be identified; (3) logical and probabilistic constraints need to be identified and incorporated; (4) probability distributions need to be assessed; and (5) sensitivity analysis and evaluation of the network have to be performed.

Expert elicitation is an essential task in order to build the network and goes therefore hand in hand with the network construction. Following [13], expert elicitation is a five step process consisting of: (1) a decision has to be made how information will be used; (2) it has to be determined what information will be elicited from the expert; (3) the elicitation process needs to be designed; (4) the elicitation itself has to be performed; and (5) the elicited information needs to be translated (encoded) into quantities.

A specific problem when working with Bayesian networks is to elicit the prior and conditional probability values. [14] argue that even though probability theory is optimal for the task of decision making, it is often found to be impractical for people to use. On the other hand, qualitative approaches to deal with uncertainty, which appear to be more naturally usable by people, often lack in precision.

In order to elicit the prior and conditional probabilities for our project we developed a dictionary, which, as for example described in [14], can be specified to allow the expert to express his or her belief for or against a statement or claim in a so called argument. The argument is expressed in qualitative terms using qualifiers [14] that then are translated into probabilities. Several dictionaries have been described in the literature (e.g., [15]). However, for our task we needed to develop a suitable dictionary together with the expert, since it was important to the expert to know how the qualitative terms would translate into probabilities in order to fully understand what the qualitative terms stand for. It was also important that the

qualitative terms match, as much as possible, the way the expert intuitively thinks about probabilities of symptoms for a developing medical condition. Sometimes we had to reverse the reasoning, since the available information was in the form of  $P(\text{symptom}|\text{disease})$  rather than  $P(\text{disease}|\text{symptom})$ . The qualifiers and their associated probabilities used are defined as:

- $x$  is known to be false  $\rightarrow P(x) = 0$
- $x$  is very unlikely  $\rightarrow P(x) = 0.01$
- $x$  is unlikely  $\rightarrow P(x) = 0.1$
- $x$  has a negative indication  $\rightarrow P(x) = 0.25$
- $x$  is random  $\rightarrow P(x) = 0.5$
- $x$  has a positive indication  $\rightarrow P(x) = 0.75$
- $x$  is likely  $\rightarrow P(x) = 0.9$
- $x$  is very likely  $\rightarrow P(x) = 0.99$
- $x$  is known to be true  $\rightarrow P(x) = 1$

Note, that this dictionary is only applicable for specifying how probable a medical condition is, given the observable symptoms. To interpret the networks' outcome a different dictionary, which is specific for each individual network needs to be generated. This output dictionary specifies an upper and a lower threshold for the output probability for each risk class (e.g., the classes no risk and risk).

Further information needed from the expert was how the variables depend on each other, what the prior probabilities of the medical conditions and the observables are, what the conditional dependencies between the symptoms and the developing medical condition are, etc. A resulting network for hyperglycemia risk detection for a diabetic person is presented in Fig. 1, developed using GeNIe 2.0 [16]. This network has eight variables in total: food intake increase (FI), body weight gain (BW), soft drink intake increase (SD), gender (G), prostatic hypertrophy (PH), prolapsed bladder (PB), diuresis frequency increase (DF) and risk of hyperglycemia (RH). The latter one is the target variable which probability we are interested in. (A previous and invalidated version of this network can be found in [17].)

The target variable RH provides a probability value which in relation to the aforementioned thresholds indicates if the patient currently is at risk of developing/experiencing hyperglycemia. The lower threshold for risk of hyperglycemia for this particular network is  $P(RH) = 0.9$  (or true = 90 % for RH) which means that the value of RH for true = 95 %, indicates a risk of hyperglycemia. This result is based solely on the information that is available from the deployed input sources (FI, BW, SD, G, PH, PB, and DF). The network in Fig. 1 shows the case for a diabetic male patient (G male = 100 %) with no prostatic hypertrophy (PH false = 100 %) who has been observed to be drinking a lot of soft drinks (SD true = 100 %). An increased food intake or a body weight gain has not been observed and is therefore set to true = 75 % which is the base rate for either of these that have been derived from the practical experience of the medical expert.

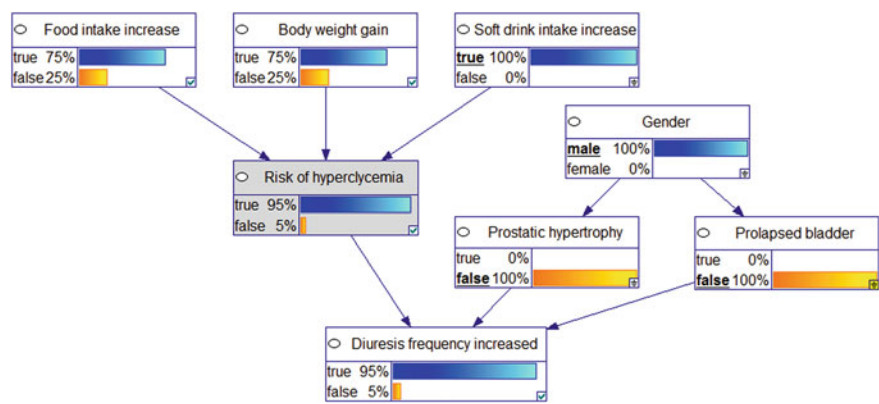


Fig. 1 Bayesian network for risk of hyperglycemia detection

The network in Fig. 1 is only one out of many possible Bayesian networks representing evaluation of risk of hyperglycemia. What variables are part of the network depend on (1) what information that can be provided by sensors and (2) how the experts, involved in the design of the network, perform the diagnosis based on the information from the available sensors. If more or different sensors are used, then the network's layout and the corresponding probability values must be refined. Further, it is important to realize (as previously mentioned) that the network is not performing a complete diagnosis. Instead, it only provides an indication of the risk that the patient is suffering from the effects of hyperglycemia. If the network indicates a risk, the patient will be asked by the AMI system to make sure that they are risking to suffer (or are suffering) from the effects of hyperglycemia by measuring the blood glucose level. The reason to avoid frequent direct measures of the glucose level is to increase the person's quality of life. People quickly tend to become annoyed when they are asked to interact with the system when they can see no obvious reason for it. As a consequence, they may react by generally ignoring the system's recommendations [18]. Therefore, reducing the frequency of and number of interactions whenever possible is important to ensure that the system is used appropriately.

### 3 Validating the Developed Model

Validity of Bayesian networks established through expert elicitation is, according to [19], usually tested by comparing the model's predictions to available data or by asking the expert to check whether the network's outcomes appear to be accurate. In our case, we need the expert to specify for each possible combination of evidence, what his diagnosis would be. For the Bayesian network to come to the same result means, that there exists a clear threshold for the probability that the patient

currently is at risk of hyperglycemia, that separates all evidence combinations into no risk and risk in the same way as the human expert. Given that each of the seven non-target variables can take one out of three values: true, false, respectively male or female for the gender node, or no evidence (n.e.), there are  $3^7 = 2187$  possible evidence combinations.

We can generate all combinations of symptoms automatically from the Bayesian network and at the same time calculate the resulting probability for RH (risk of hyperglycemia) for each of them. Table 1 shows an excerpt thereof.

**Table 1** Excerpt from the evidence combination table for risk of hyperglycemia detection

FI	BW	SD	G	PH	PB	DF	P (RH)	Risk
n. e.	n. e.	n. e.	n. e.	n. e.	n. e.	n. e.	0.85	False
True	n. e.	n. e.	n. e.	n. e.	n. e.	n. e.	0.89	False
True	True	n. e.	n. e.	n. e.	n. e.	n. e.	0.92	True
n. e.	n. e.	True	n. e.	n. e.	n. e.	n. e.	0.95	True
n. e.	n. e.	n. e.	n. e.	n. e.	n. e.	True	0.96	True
n. e.	n. e.	n. e.	Male	True	No	True	0.85	False
True	True	True	n. e.	n. e.	n. e.	n. e.	0.99	True
...	...	...	...	...	...	...	...	...

After that, it needs to be identified for which of these cases hyperglycemia actually is suspected, which is represented in the table's last column, denoted Risk. Some of the evidence combinations can be disregarded, as they make no sense. It is, for example, impossible for a patient to have both, prostatic hypertrophy and a prolapsed bladder. For all remaining cases, it needs to be decided if they represent a risk or no risk of hyperglycemia and thereby if the corresponding probability value should be below or above the threshold.

To alleviate this process, validation rules can be specified that cover several alternatives at once and for which the value for Risk then can be set automatically. For example, whenever  $DF = \text{true}$  and  $PH = \text{false}$  then  $Risk = \text{true}$ . This rule covers all cases where the patient suffers from an increase in diuresis frequency but does not have prostatic hypertrophy. Yet another way of formulating rules is to say, e.g. whenever two of the variables FI, BW, and SD are true then  $Risk = \text{true}$ . These validation rules support the process of partitioning the networks outcome into risk classes. They can be viewed as expressions of criteria for when a risk of the medical condition ought to be detected. These criteria may, however, be incomplete.

The next step is to identify if there is a threshold for  $P(RH)$  that clearly partitions all the possible cases into at least two classes one for  $Risk = \text{true}$  and one for  $Risk = \text{false}$ . For that, we need to identify the corresponding probability values for  $P(RH)$  in each of the partitions. For each partition, we calculate the interval from the lowest value for  $P(RH)$  for  $Risk = \text{false}$  to the highest value for  $P(RH)$  for  $Risk = \text{false}$  and respectively for  $P(RH)$  for  $Risk = \text{true}$ . If the resulting two intervals are non-overlapping, which is the case for the network presented in Fig. 1; we can identify a threshold between these intervals. Each value of  $P(RH)$  below this

threshold results in no risk of hyperglycemia ( $\text{Risk} = \text{false}$ ) and the system not intervening with the person and every value above or equal to the threshold results into risk of hyperglycemia ( $\text{Risk} = \text{true}$ ) and the system intervening with the person. If the intervals would overlap, the net is not fully valid to diagnose the risk of the disease without doubt. In that case, the network's design needs to be adjusted accordingly.

Additionally, [19] emphasizes that model validity should not only be checked regarding the model's outcome, but as well regarding the mechanism through that the outcome is obtained. They propose seven different types of validity that the net should be tested for: Nomological validity, face validity, content validity, concurrent validity, convergent validity, discriminant validity, and predictive validity.

As mentioned previously, Bayesian networks have been successfully used within medical diagnostic, which accounts for the nomological validity of our approach. The model's face validity is provided by the expert, who was involved in designing the net, and in analyzing the predictive validity of the net. Content validity is achieved by consulting the expert, rather than the literature. The expert decided what variables and what states of the variables need to be modeled with regard to building a net that models his or her own internal model for risk of hyperglycemia identification. At this stage, the network does not contain any reoccurring parts for that concurrent validity needs to be tested. Convergent and discriminant validity are achieved up to a certain degree through the fact that, as mentioned before, reasoning in medical diagnosis is usually done from symptoms to causes. The world is usually modeled in the way that causes are parent nodes of symptoms. How we achieve predictive validity has been already described above.

## 4 Discussion and Future Work

In this paper, we described the development of the Bayesian network for automatic detection of a person being at risk of a medical condition on the example of hyperglycemia in a diabetic patient. The purpose of the work presented here is to develop a general method for designing and validating risk detection networks. Deployment of more and different sensors might improve risk detection. The network is based on one expert's opinion only and it would therefore be interesting to investigate if a similar network that is based on the elicitation of several experts will show improved results. However, the next step in our project will be to test the network's results against the real world. To identify more risk classes, e.g., no risk, low risk, risk, high risk, very high risk, would be an additional improvement as the system is meant to monitor the patients and to encourage them to a healthier life style. When only a low risk of hyperglycemia is indicated, this could be used to prompt the patient to generally try to change an unhealthy habit that appears to be the reason for the risk being apparent. In order to do that, the system must know what the most likely reason for the diagnosis is. Therefore, explanation methods for Bayesian networks [20] could be applied.

Commonly used methods for information fusion can be roughly grouped into two groups (1) using precise probability e.g. based on Bayesian theory [21] that we have utilized in this approach and (2) using imprecise probability [22] e.g. different variants of evidence theory [11], or credal sets e.g. [23]. These two groups differ from each other regarding how evidence is modeled within the underlying evidential framework and how it is combined [24]. It would be interesting to compare the performance of imprecise frameworks for the same task.

**Acknowledgments** We would like to thank Giacomo Vespasiani, M.D. for his enthusiastic cooperation. The HELICOPTER project is funded by the Ambient Assisted Living Joint Program (AAL-2012-5-150) (<http://www.aal-europe.eu>).

## References

1. J. Mellin, H.J. Steinhauer, L. Boffi, D. Kristály, P. Ciampolini, N. Pierantozzi, B. Carlsson, C. Berg, M. de Pender, G. Vespasiani, An automated triage for ambient assisted living in HELICOPTER. J. Des. Test (under review)
2. P. Lucas, Bayesian networks in biomedicine and health care. Editor. Artif. Intell. Med. **30**, 201–214 (2004)
3. J. Pearl, Fusion, Propagation and Structuring in Belief Networks. Artif. Intell. **29**, 241–288 (1986)
4. P. Lucas, Bayesian networks in medicine: a model-based approach to medical decision making, in *Proceedings of the EUNITE Workshop on Intelligent Systems in Patient Care* (2001), pp. 73–97
5. S. Visscher, P. Lucas, K. Schurink, M. Bonten, Using a Bayesian-network model for the analysis of clinical time-series data, in *Artificial Intelligence in Medicine 2005*, ed. by S. Miksch et al. LNAI, vol. 358 (Springer, Berlin, 2005), pp. 48–52
6. S. Andreassen, C. Riekehr, B. Kristensen, H.C. Schønheyder, L. Leibovici, Using probabilistic and decision – theoretic methods in treatment and prognosis modelling. Artif. Intell. Med. **15**, 121–134 (1999)
7. D. Aronsky, P.J. Haug, Diagnosing community-acquired pneumonia with a Bayesian network, in *Proceedings of AMIA Symposium* (1998), pp. 632–636
8. I.A. Beinlich, H.J. Suermondt, R.M. Chavex, G.F. Cooper, The ALARM monitoring system: a case study with two probabilistic inference techniques for belief networks, in *Artificial Intelligence in Medicine*, ed. by J. Hunter, J. Cookson J, J. Wyatt. (Springer, 1989), pp. 247–256
9. O. Nee, A. Hein, Clinical decision support with guidelines and Bayesian networks, in *Decision Support Systems Advances*, ed. by G. Devlin. INTECH (2010), pp. 117–136
10. S. Russel, P. Norvig, *Artificial Intelligence: A modern Approach* (Pearson, Upper Saddle River, 2010)
11. G. Shafer, *A Mathematical Theory of Evidence* (Princeton University Press, Princeton, 1976)
12. A. Jøsang, Subjective logic, book draft (2013), [http://folk.uio.no/josang/papers/subjective\\_logic.pdf](http://folk.uio.no/josang/papers/subjective_logic.pdf)
13. T.G. Martin, M.A. Burgman, F. Fidler, P.M. Kuhnert, S. Low-Choy, M. McBride, K. Mengersen, Eliciting Expert Knowledge in Conservation Science. In: Conservation Biology. Society for Conservation Biology (2011)
14. J. Fox, D. Glasspool, J. Bury, in *Quantitative and Qualitative Approaches to Reasoning Under Uncertainty in Medical Decision Making*. ed. by S. Quagline, P. Barahona, S. Andreassen. AIME 2001 (Springer, Berlin, 2001), pp. 272–282



15. J. Fox, P. Krause, M. Elvang-Gøransson, Argumentation as a general framework for uncertain reasoning, in *proceedings of the ninth conference on uncertainty in artificial intelligence* (1993)
16. Decision Systems Laboratory, University of Pittsburgh. GeNIe 2.0, <http://genie.sis.pitt.edu/?ver=20048430>
17. J. Mellin, G. Vespasiani, M. Mustica, G. Matrella, M. Mordonini, C. Berg, I. Schoormans, B. Carlsson, J. Bak, Domain model (2014), <http://www.his.se/HELICOPTER-WP4-deliverables/>
18. P. Lyons, A.T. Cong, H.J. Steinhauer, S. Marsland, J. Dietrich, H.W. Guesgen, Exploring the responsibilities of single-inhabitant smart homes with use cases. *J. Ambient Intell. Smart Environ.* **2**(3), 211–232 (2010)
19. J. Pitchford, K. Mengersen, A proposed validation framework for expert elicited Bayesian networks. *Expert Syst. Appl.* **40**(1), 162–167 (2013)
20. C. Lacave, F.J. Díez, A review of explanation methods for Bayesian networks. *J. Knowl. Eng. Rev.* **17**, 107–127 (2002)
21. J.M. Bernardo, A.F.M. Smith, *Bayesian Theory* (Wiley, Hoboken, 2000)
22. P. Walley, *Statistical Reasoning with Imprecise Probabilities* (Chapman and Hall, London, 1991)
23. A. Karlsson, R. Johansson, S.F. Andler, Characterization and empirical evaluation of bayesian and credal combination operators. *J. Adv. Inf. Fusion* **6**(2), 150–166 (2011)
24. A. Karlsson, Evaluating credal set theory as a belief framework in high-level information fusion for automated decision-making. PhD thesis, Örebro University, School of Science and Technology (2010)

Ambient Intelligence - Software and Applications  
6th International Symposium on Ambient Intelligence  
(ISAmI 2015)

Mohamed, A.; Novais, P.; Pereira, A.; Villarrubia, G.;  
Fernández-Caballero, A. (Eds.)

2015, XIII, 274 p. 90 illus., Softcover

ISBN: 978-3-319-19694-7