



## Exploiting causal functional relationships in Bayesian network modelling for personalised healthcare



Marina Velikova<sup>a,\*</sup>, Josien Terwisscha van Scheltinga<sup>b</sup>, Peter J.F. Lucas<sup>a</sup>, Marc Spaanderman<sup>c</sup>

<sup>a</sup> Institute for Computing and Information Sciences, Radboud University Nijmegen, The Netherlands

<sup>b</sup> Department of Obstetrics and Gynaecology, Radboud University Nijmegen Medical Centre, The Netherlands

<sup>c</sup> Department of Obstetrics and Gynaecology, Maastricht University Medical Center, The Netherlands

### ARTICLE INFO

#### Article history:

Available online 10 April 2013

#### Keywords:

Temporal Bayesian networks  
Causal independence models  
Clinical decision support  
Pregnancy disorders  
eHealth  
Home monitoring

### ABSTRACT

Bridging the gap between the theory of Bayesian networks and solving an actual problem is still a big challenge and this is in particular true for medical problems, where such a gap is clearly evident. We argue that Bayesian networks offer appropriate technology for the successful modelling of medical problems, including the personalisation of healthcare. Personalisation is an important aspect of remote disease management systems. It involves the forecasting of progression of a disease based on the interpretation of patient data by a disease model. A natural foundation for disease models is physiological knowledge, as such knowledge facilitates building clinically understandable models. This paper proposes ways to represent such knowledge as part of engineering principles employed in building clinically practical probabilistic models. The methodology has been used to construct a temporal Bayesian network model for preeclampsia – a pregnancy-related disorder. The model is the first of its kind and an integral part of a mobile home-monitoring system intended for use in daily pregnancy care. We conducted an evaluation study with actual patient data to obtain insight into the model's performance and suitability. The results obtained are encouraging and show the potential of exploiting physiological knowledge for personalised decision-support systems.

© 2013 Elsevier Inc. All rights reserved.

## 1. Introduction

As making decisions in medicine is fraught with difficulty because of the significant uncertainties involved, researchers have tried to represent that uncertainty explicitly. In particular Bayesian networks appear to offer a natural and intuitive formal foundation for uncertainty models that are part of clinical decision-support systems [20]. However, despite the progress made in the research, there are still very few systems based on Bayesian networks that are actually used in daily clinical practice. A possible explanation for this is that medical doctors are reluctant to adopt new technology unless its advantages are crystal clear. One new area that meets such requirement is the *personalisation* of healthcare. Personalisation is understood as involving the forecasting of the progression of a disease based on the interpretation of patient data by a probabilistic model. Without intelligent systems that are able to fuse information coming from different sources, with much of this information being uncertain, personalisation will not be possible. Bayesian networks in particular appear to offer the right capabilities to represent and manipulate this uncertain knowledge for personalisation. The predictions obtained in this way inform the patient and doctor on whether or not the disease is under control. Forecasting concerns reasoning about the state of a system that evolves over time, and therefore temporal data and knowledge are of interest here.

\* Corresponding author.

E-mail address: [marinav@cs.ru.nl](mailto:marinav@cs.ru.nl) (M. Velikova).

It is often the case with a severe disorder in medicine that the number of patients with this disorder is limited, whereas the whole group of patients under surveillance for the disorder is much larger. Under such conditions it is hard to develop a probabilistic model based on data alone. Exploiting biological knowledge about the way the relevant organ systems function, i.e. *physiological* knowledge, if available, may then guide the modelling process.

In contrast to earlier work exploiting such knowledge, in our research we are aiming at the development of a generic Bayesian network methodology for building probabilistic clinical models to predict disease progression. As the resulting models must be suitable for daily clinical use, employing sound engineering principles is of crucial importance. A novel contribution is that the introduced principles support finding the right balance between incorporating simple, easy-to-collect measurements and risk factors, on the one hand, and physiological knowledge to structure these measurements, on the other hand. The resulting models are also suitable for personalisation of disease management.

Thus, modelling organ functioning at a certain level of abstraction is taken as the principle that integrates the interactions between disease outcome, observables, such as signs, symptoms and lab tests, actions or treatments, and risk factors. This choice has been motivated by the way clinical knowledge is presented in medical textbooks and in clinical education. The method was used in developing a Bayesian network for predicting the development of a hypertension-related pregnancy syndrome, called *preeclampsia*, with personalisation of the associated treatment as a goal. The model built is the first of its kind and an integral part of a smart home-monitoring system for pregnancy care that provides decision support to the patient and the caregiver with the aim to offer timely management of the life-threatening complication of preeclampsia.

The paper is organised as follows. In Section 2 we review related research in the area of modelling within the medical domain with an emphasis on the development of patient-specific models. Basic Bayesian network theory and the proposed methodology for modelling disease progression – the main scientific contribution of this research – are described in Section 3. In the same section we also discuss issues concerning evaluation of the methodology and propose a number of criteria for that. The application of the methodology to the development of a mobile system for pregnancy care is described in Section 4. What we have achieved in this research is brought in perspective in Section 5.

## 2. Related work

Early work mostly focused on using Bayesian network as aids for the diagnosis of disease. Well-known examples include the MUNIN [2] and Pathfinder [12,11] models. Special diagnostic reasoning methods for test selection in medicine have also been developed [1,45]. Various approaches for simplifying the network structure and the specification of probability distributions when building diagnostic models are discussed in [25]. In [31] the authors propose techniques for automatic construction of dynamic influence diagrams from a set of causal rules in a knowledge base, with an application to diagnosis of acute abdominal pain.

Later more emphasis was placed on treatment selection and making a prognosis, e.g. [8,17–19]. Markov decision processes and dynamic influence diagrams are examples of temporal probabilistic models that have been used in [35] for selection of treatment strategies. In [36] a noisy-threshold model has been proposed including 11 attributes measured at admission and playing the role of causes for predicting carcinoid heart disease; the model has been learned and trained on a small dataset of 54 patients. The use of dynamic Bayesian networks as the basis for the construction of prognostic models has been explored in [34] with the particular application to prognosis of low-grade carcinoid tumours; evaluation results regarding three patients were also presented. Another Bayesian-network-based decision-support system, called TREAT, aiming at predicting bacteremia in patients is described in [28]. A randomised clinical trial showed improvement in the percentage of appropriate empirical antibiotic treatments when using the system [29]. A study that dealt with the problem of predicting the susceptibility to future antibiotic treatments using a Bayesian network is presented in [47], which also discusses the potential of including the method in a larger decision-support system. To facilitate explaining the optimal policies of decision-making problems solved by means of influence diagrams, a method based on parsimonious lists of alternative decisions is proposed in [5]; the practical usefulness of the method is investigated with respect to the treatment of non-Hodgkin lymphoma of the stomach. A construction method of a prognostic Bayesian network model using clinical data and supervised learning models, such as decision trees, is proposed in [42]. An application to the domain of cardiac surgery, where hospital mortality is used as outcome variable, illustrates the use of the methodology. Another recent paper that studied the use of probability theory in the context of rule-based computer-assisted diagnosis is presented [6].

There is some other, earlier, work on exploiting physiological knowledge in building Bayesian networks [26,4,32,16]. Other investigations have focused on the development of patient-specific models based not only on probabilistic graphical models [33], but also on causal hierarchical structures representing physiological knowledge [27], fuzzy cognitive maps for modelling cause-effect relationships in the medical domain [21] and probabilistic soft logic specifying medical domain knowledge [3].

Despite the large body of research, only a very small number of applications have successfully made the step towards clinical deployment. With the emerging trend towards provision of personalised healthcare, there is a clear need for the development of novel approaches to the construction of patient-specific models that can cope with the often limited availability of clinical data, capture physiological causal relationships in a systematic manner and allow easy deployment. The methodology presented next is a step in this direction.

### 3. Building Bayesian networks from medical principles

The human body consists of an interacting system of organs, each of them with a multitude of functions. An example is the circulatory system which includes the heart, blood vessels and blood as organs concerned with the transportation of nutrients, gases, blood cells and hormones through the body. Another example is the excretory system with as a major organ the kidneys where one of the functions is the removal of waste products from the blood. An understanding of organ functioning and malfunctioning, i.e. physiological knowledge, is used by clinicians for the diagnosis, prognosis and treatment of a disorder. Hence, in developing patient-specific models one also needs to explore the basic causal mechanisms and interactions known from physiology. This is the principle we adopted in the temporal Bayesian network methodology presented here for acquiring the structure and probabilistic parameters. However, we first review the basics of Bayesian networks underlying the proposed methodology.

#### 3.1. Preliminaries

A Bayesian network is defined as a pair  $BN = (G, P)$ , where  $G$  is an acyclic directed graph, ADG for short,  $G = (V, E)$ , with nodes  $V$  and directed edges or arcs  $E \subseteq V \times V$ , and  $P$  is a joint probability distribution of the random variables  $X$ . There exists a 1–1 correspondence between the nodes in  $V$  and the random variables in  $X$ ; the directed edges  $E$  express (conditional) dependence and independence relationships between the variables. We say that  $G$  is an  $I$ -map of  $P$  if any independence represented in  $G$ , denoted by  $A \perp\!\!\!\perp_C B \mid C$ , is also satisfied by  $P$ , i.e.

$$A \perp\!\!\!\perp_C B \mid C \implies X_A \perp\!\!\!\perp_P X_B \mid X_C,$$

where  $A, B$  and  $C$  are mutually disjoint sets of nodes of the ADG  $G$  and  $X_A, X_B$  and  $X_C$  are the corresponding sets of random variables, indexed by  $A, B$  and  $C$ . The graphical part of a Bayesian network is by definition an  $I$ -map of the associated joint probability distribution. A Bayesian network offers a compact representation of the joint probability distribution  $P$  in terms of local conditional probability distributions, CPDs for short, by taking into account the conditional independences represented by the ADG as follows:

$$P(X_V) = \prod_{v \in V} P(X_v \mid X_{\text{par}(v)}),$$

where  $P(X_v \mid X_{\text{par}(v)})$  is a family of CPDs, where each value of the set of variables  $X_{\text{par}(v)}$  corresponding to the parents of the node  $v$ ,  $\text{par}(v)$ , yields a CPD.

Determining the CPDs from expert knowledge or by learning from data is a cumbersome task when one has to deal with a large number of variables involved in complex relationships – often the case for real applications. From the Bayesian network definition it is clear that the size of a probability table, including all the CPDs of a node, is exponential in the number of parents. To provide an efficient way to specify interactions among variables in a compact fashion, one may resort to using logistic regression if data are available, or alternatively, if no or insufficient data are available one may use the notion of causal independence [10]. Causal independence arises when multiple causes (parent nodes) lead to a common effect (child node) through interaction of independent uncertain processes. This type of models allows decomposing a probability distribution in a systematic way in terms of Boolean interactions among local parameters. Such a decomposition makes it easier and tractable to deal with problems involving a large number of causes.

The definition of the notion of causal independence given here follows the one from [43]. The general structure of a causal-independence model is shown in Fig. 1. It expresses the idea that causes  $C_1, \dots, C_n$  influence a given common effect  $E$  through intermediate variables  $I_1, \dots, I_n$ ; the intermediate variable  $I_k$  is considered to be the contribution of the cause variable  $C_k$  to the common effect  $E$ . The interaction function  $b$  represents in which way the intermediate effects  $I_k$ , and indirectly also the causes  $C_k$ , interact. This function  $b$  is defined in such way that when a relationship between the  $I_k$ 's and

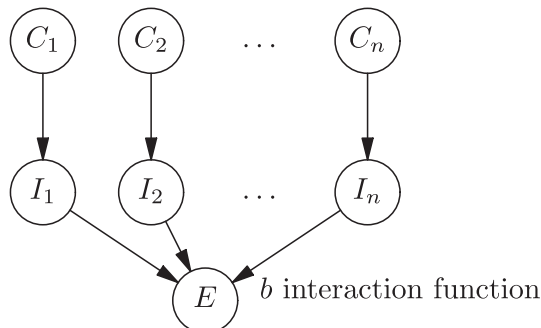


Fig. 1. Causal-independence model:  $C_k$  are cause variables,  $I_k$  are intermediate, hidden nodes, and  $E$  stands for the effect variable.

$E = 1$  is satisfied, then it holds that  $b(I_1, \dots, I_n) = 1$  (*true*); otherwise, it holds that  $b(I_1, \dots, I_n) = 0$  (*false*). The graph structure expresses that the effect variable  $E$  is conditionally independent of each cause  $C_k$  given the associated variable  $I_k$ . It is assumed that absent causes do not contribute to the effect, i.e.  $P(I_k = 1 \mid C_k = 0) = 0$ . There is other recent research where causal independence models have been explored in designing medical Bayesian networks [37,41,38,43].

Bayesian networks also offer powerful methods to represent and reason about uncertain state changes over time. The best known representatives are *dynamic Bayesian networks* – temporal models where the same variables of interest, describing both the state of the system, observables, conditions and actions that may change the state at different points of time. A usual assumption underlying these models is that: (i) the future state is conditionally independent of the past state given the present state (first-order Markov property), and (ii) the probabilistic temporal relations between adjacent states do not change over time (time invariance or stationarity condition). This way, a dynamic Bayesian network becomes a compact process representation that can be employed for forecasting [46]. However, for some problems, the real system changes in functionality because of change in the parameters or structure, and then the assumption for time invariance is incorrect. A typical example from medicine is the patient's physiology that may change due to a disease process. In the following, we present a temporal Bayesian network model where the probabilistic parameters of the Bayesian network are *not* assumed to be time invariant and this is a necessary assumption in that domain.

### 3.2. A causal functional model for syndrome progression

In medicine, a *syndrome* is defined as a fixed set of signs and symptoms. The common cause of these signs or symptoms may be unknown, but the underlying mechanisms giving rise to the signs and symptoms may be known. If there is also a known common cause of the signs and symptoms, one refers to the syndrome as a 'disease'. Below, we will use the term 'syndrome' in a wide sense, encompassing that of 'disease'. Symptoms are subjective evidence reported by the patient, e.g. coughing, fatigue, headache, whereas signs are objective evidence obtained via tests and measuring devices, e.g. increased body temperature as measured by a thermometer and high blood pressure via a sphygmomanometer. Such evidence may point to the dysfunctioning of one or more body organs  $X, Y, \dots$  and are sometimes distinguished in major and minor in establishing the diagnosis of a syndrome, in an attempt to give them a different weight. All signs, symptoms and lab test results have associated uncertainty.

The relationship between a syndrome and the associated evidence is established via the functioning of organ- $X$ . This can be modelled by the graph structure  $\text{SYNDROME} \leftarrow \text{ORGAN-}X\text{-FUNCTION} \rightarrow \{\text{SIGNS, SYMPTOMS}\}$ . Here,  $\text{SYNDROME}$  is the variable of interest, having two states: *absent* and *present*. The hidden, non-measurable variable  $\text{ORGAN-}X\text{-FUNCTION}$  has in the simplest case also two states: *normal*, i.e. normal physiology, and *abnormal*, i.e. pathophysiology. The observable variables  $\text{SIGNS}$  and  $\text{SYMPTOMS}$ , on the other hand, may be multi-valued or have a continuous range. The explicit modelling of organ function can be exploited in developing prognostic models. Such a model would allow detecting early warning signs of the syndrome through worsened organ function before the signs, symptoms and lab data for diagnosing the syndrome are actually observed.

While a disease and signs and symptoms may be correlated, they need not have a direct cause-effect relationship; it is possible that a separate underlying problem or *risk factor* explains the correlation. For example, the patient's characteristics such as preexisting diseases, age, gender and genetics may affect the functioning of the organs related to the disease. Apart from the risk factors, *treatment* may also affect the functioning of an organ, for example by removing risk factors or by preventing complications. These cause-effect relationships can be modelled by the following graph structure:  $\text{RISKFACTOR} \rightarrow \text{ORGAN-}X\text{-FUNCTION} \leftarrow \text{TREATMENT}$ . The variables  $\text{RISKFACTOR}$  and  $\text{TREATMENT}$  are categorical and have two or more states. Their prior probability is often determined using population statistics, e.g. for the population of pregnant women the chance of having diabetes prior to pregnancy is known to be around 1.1%, and 90% of women do not get any treatment at pregnancy week 12.

It is often the case that the impact of individual risk factors is known from epidemiological research although their combined effect is not. This makes it difficult, if not impossible, to estimate the combined probabilistic effect of various factors together on organ dysfunction. This problem can be tackled by applying causal independence modelling as discussed in Section 3.1. More precisely, for each risk factor  $A, B, \dots, M$  we add the intermediate (hidden, non-observable) variables  $I_k$ ,  $k \in \{A, B, \dots, M\}$ , to model the interaction between the risk factors and the organ functioning. To explicitly model the risk of organ  $X$ 's dysfunction, another hidden variable is introduced:  $\text{ORGAN-}X\text{-RISK}$ . Since, in the medical domain the risk factors often interact by adding up to the risk, the logical OR seems the appropriate operator for representing this interaction with regard to  $\text{ORGAN-}X\text{-RISK}$ ; this implies that if at least one of the risk factors is present then the risk of dysfunction of organ  $X$  is true. In addition, the impact strength of each risk factor on  $\text{ORGAN-}X\text{-RISK}$  is reflected in the probability distribution  $P(I_j \mid \text{RISKFACTOR}_j)$ . Given the values of two risk factors  $\text{RISKFACTOR}_j = a$  and  $\text{RISKFACTOR}_k = b$  and the knowledge that the former has more likely impact on  $\text{ORGAN-}X\text{-RISK}$  than the latter, this is modelled by the following qualitative constraint:

$$P(I_j = 1 \mid \text{RISKFACTOR}_j = a) \geq P(I_k = 1 \mid \text{RISKFACTOR}_k = b).$$

The representation of these relationships is depicted in Fig. 2(a), where the arrows indicate the cause-effect direction. This representation corresponds to a static view of organ function, e.g. at the moment when measurements are made or

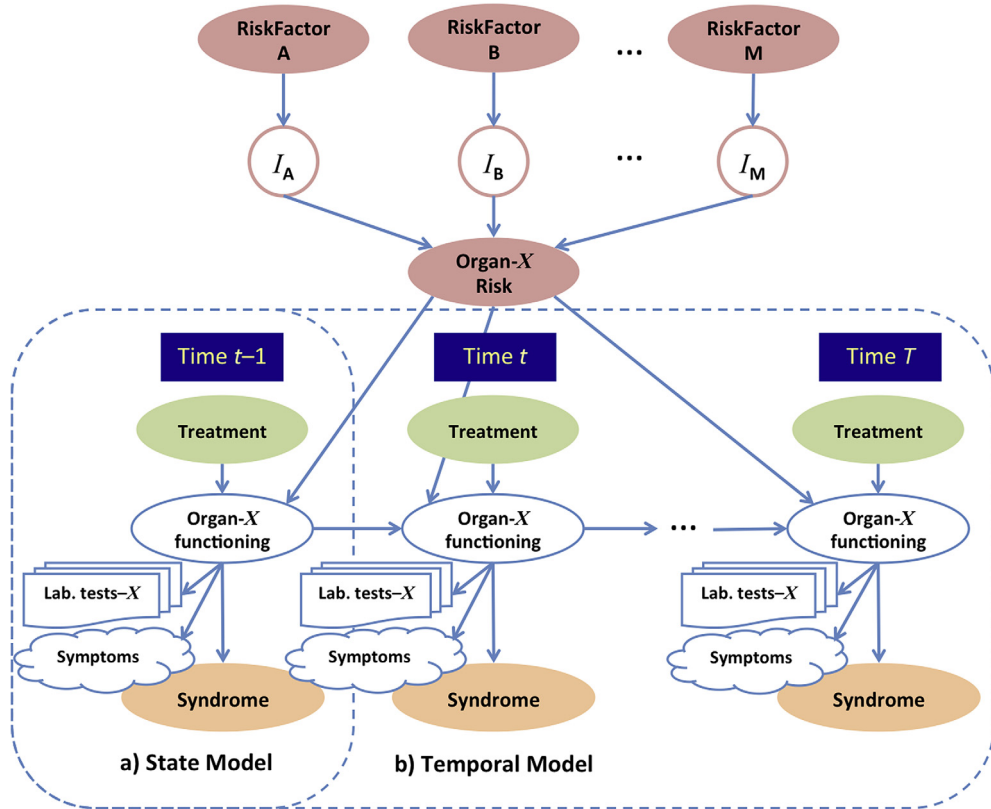


Fig. 2. Causal functional model that expresses the temporal development of a syndrome: (a) Static state model ignoring time; (b) model showing temporal progression.

symptoms are detected. While such a “snapshot” of the body system can be useful for diagnostic purposes, it is definitely insufficient for prognostic tasks, which concern predicting the organ’s function over time.

Clinical monitoring of the patient’s health status usually consists of checkups done at different time points  $t = 1, \dots, T$ . Taking into account the functioning of organ X at previous checkups allows the medical doctor to conclude whether or not changes in function have occurred. Based on this information it is possible to establish a diagnosis at time  $t$ . Furthermore, with the history and current health status, one can also predict future organ functioning and progression of a syndrome – make a prognosis – and adjust the treatment. The number of time points to be considered in the model will depend on the clinical problem at hand and on the availability of observations. The chain of temporal dependencies of organ functioning between the subsequent checkups is depicted in Fig. 2(b).

### 3.3. Evaluation issues

As the proposed methodology explores available medical domain knowledge extensively, this is expected to yield at least some confidence in the correctness of the model. However, the model’s performance with respect to the clinical task, such as diagnosis and prognosis, for which it was built still needs to be determined. In some cases, patient data with the target outcome known are available for that purpose. However, in medicine very often the ground-truth is unknown and one is obliged to rely on expert judgements. We introduce evaluation methods that support accounting for the dynamics and uncertainty of the prediction of syndrome progression.

Accuracy is a typical performance measure of the diagnostic capabilities of a model, where the extent to which the model can correctly diagnose a syndrome is compared to the clinical or pathological ground-truth. Given the continuous-valued probability estimates of interest, the Receiver Operating Characteristic (ROC) curve and the Area Under the Curve (AUC) are commonly used to determine the balance between true positive (TP) and false positive (FP) rates. In an ROC curve the true positive rate is plotted as a function of the false positive rate for different cut-off points. In contrast to the standard evaluation procedure, where the positive and negative cases are defined without reference to time, in the current methodology time of diagnosis has to be considered. In other words, a patient case is considered *positive* only at the point of diagnosis and before that considered *negative*.

Although diagnosis is of crucial clinical importance, forecasting and prognosis are significant for prevention and treatment selection as they can predict the development of a syndrome before its manifestation. However, evaluating the prognostic

capabilities of a model is nontrivial as there is often no clear definition of the patient's condition prior to the occurrence of a syndrome. In this study we consider a number of methods that can provide insight into this task:

- M<sub>1</sub>** Early detection of an increased risk of the development of a syndrome. For each positive patient one can check whether or not the model is capable of detecting an increased risk prior to the diagnosis and, if so, at which time point. To decide when the risk is considered "increased" a cut-off FP rate can be chosen.
- M<sub>2</sub>** Distinction between high- and low-risk patients. Given the temporal nature of the kind of problems we wish to tackle, one can compute the probability for developing a syndrome at time  $t'$ ,  $t' > t$ , given the risk factors and the measurements made up to time point  $t$ . The effect of treatment during the time period  $[t + 1, t']$  can be eliminated by supplying *no treatment* as evidence in the model. The resulting probabilities  $P_{t'}$  (SYNDROME = yes) are distinguished into categories, and these are used to construct histograms for both patients with and without the disease to check whether a clear distinction between high- and low-risk patients can be made. If so, appropriate monitoring and treatment programs can be applied.

Finally, the model's data fitting capabilities can be tested by using the *Bayesian information reward* metric [14], which measures for the variables of interest how different the model's posterior distribution is from the expected prior. It is especially suited for non-uniform priors, as it is often the case in medical applications, and it is defined by:

$$IR_B = \frac{\sum_i I_i}{n}, \quad (1)$$

where  $n$  is the number of cases,  $I_i = I_i^+$  for the true class and  $I_i = I_i^-$  otherwise, with:

$$I_i^+ = \log \frac{p_i^*}{p_i} \quad \text{and} \quad I_i^- = \log \frac{1 - p_i^*}{1 - p_i},$$

where  $p_i^*$  and  $p_i$  are the posterior probability computed from the model and the prior probability for case  $i$ , respectively. The range values of  $IR_B$  are finitely bounded in the positive direction as priors are never zero and  $-\infty$  when the probability for the true value is 0. Furthermore,  $IR_B = 0$  indicates that  $p_i^* = p_i$ . A typical choice for  $p_i$  is the fraction of positive cases in the dataset. Intuitively, by taking prior probabilities into account  $IR_B$  rewards learners not only for making the right decisions, but also for getting the probability of an event correct. The learner shows the ability to account for evidence by updating its distribution; perfect calibration maximises the reward [14].

## 4. Application to pregnancy care

### 4.1. Background

Approximately 15% of first-time pregnant women develop high blood pressure and approximately half of them develop associated problems, such as kidney damage and subsequent proteinuria (protein in the urine; normally urine contains almost no protein as it should not pass the glomerular membranes of the kidney's nephrons), leading to the syndrome of *preeclampsia*. Preeclampsia is the most important cause of death among pregnant women in the Western world and it is also a leading cause of foetal complications, which include low birth weight and stillbirth. As a multisystem disorder, preeclampsia affects the placenta, and it can affect the mother's kidney, liver, heart and brain. When preeclampsia causes epileptic seizures, the condition is known as *eclampsia*. As a pregnancy-related condition the only way to cure preeclampsia is to deliver the baby.

The diagnosis of preeclampsia is established when both gestational hypertension and proteinuria are present. However, the practice of establishing a diagnosis is not that simple, as hypertension and proteinuria are correlated, which is not reflected in the clinical definition, and the challenge is to predict whether the patient *will develop* preeclampsia rather than whether the patient already has it. Gestational hypertension is diagnosed when blood pressure exceeds 140 mmHg systolic or 90 mmHg diastolic on more than two readings six hours apart after 20 weeks pregnancy (gestation). Proteinuria can be diagnosed using the protein-to-creatinine ratio – the amount of protein in the urine sample compared with the amount of creatinine, a waste product of normal muscle breakdown. If the test exceeds 30 mg/mmol there is preeclampsia. Beginning stages of kidney damage, however, can be diagnosed at protein leakage level between 3.4 and 30 mg/mmol, which can help getting timely treatment and preventing development of complications. Early anti-hypertensive treatment in the subclinical, mostly moderately hypertensive phase, reduces the risk of preeclampsia. Furthermore, it is known that various maternal factors enhance the risk for developing preeclampsia, such as older age, a family or personal history of preeclampsia, chronic hypertension, kidney disease, diabetes and multiple pregnancy [22,9]. For women who have given birth (parous women) without a history of preeclampsia the risk of developing preeclampsia is lower than for women who have never given birth (nulliparous women).

As mentioned above, the timely diagnosis of preeclampsia is not an easy task. On the one hand the stage in pregnancy when high blood pressure appears is variable and associated problems can develop within a few days. Furthermore, since the



Fig. 3. Scheme of the eMomCare system's components.

final diagnosis of gestational hypertension can only be made in retrospect, a clinician may be forced to treat some women with gestational hypertension as if they had preeclampsia. On the other hand, if a woman has underlying cardiovascular or renal disease, and thus preexisting chronic hypertension or proteinuria, the diagnosis of preeclampsia may not become clear until the disease becomes severe. This requires frequent outpatient checkups, where the health condition of a pregnant woman is monitored by diagnostic tests such as measuring maternal blood pressure, heart rate, haemoglobin, urine analysis and by observing the foetal condition. Careful checkups are carried out when the mother experience warning symptoms such as headaches or rapid weight increase, that are associated with hypertensive complications. These outpatient checkups require the patient to come to the hospital on a regular basis, thus leading to high pressure on the patient and the caregivers.

#### 4.2. eMomCare: a personalised home-monitoring system for pregnancy care

Current technological developments allow large amounts of patient data to be collected at home and automatically sent to the health-care team. This offers a number of advantages: (i) it yields test results that are closer to the real physiological values, as the effect of the medical doctor on the blood pressure, the so-called white-coat effect, is eliminated; (ii) there will be no need for the patient to frequently visit the hospital and she can be actively involved in her own medical care; (iii) the work pressure on obstetric care and healthcare costs can be possibly reduced.

This form of home monitoring requires an intelligent system, because it is desirable that both doctor and patient get insight into the patients health status and need for care without additional efforts, such as frequent telephone contacts. Of course, appropriate security must also be ensured. We took a step in this direction by developing a personalised smart care-assistant for the pregnant woman based on state-of-the-art smartphone technology, so that part of the clinical decision-making process can be moved to anywhere where the patient resides. The architecture of the smart care-assistant includes a number of components that carry out the following functions, as shown in Fig. 3:

- Collection of patient and sensor data. This is done by means of questionnaires, automatic reading of measurement equipment such as electronic blood pressure meter via Bluetooth and automatic analysis of urine strips using the phone's camera and image processing techniques.
- Automatic interpretation of both patient and sensor data *within the smartphone itself* by a specially-designed preeclampsia model. The strong cause-effect relationships in describing physiology and the inherent uncertainty in the medical domain justify our choice of Bayesian networks as a modelling technique. The model can provide feedback, explain the results obtained and recommend actions to the patient and the care team regarding the progression, or lack thereof, of the syndrome.
- Communication of the results, both textually and visually, to the care team and the patient. The data should be stored in a hospital database for further inspection by the caregivers.

**Table 1**  
Factors with their values and abbreviated names as used in the Bayesian network model.

Type	Factor	Abbreviation	Values/ranges	
Risk factors	Antiphospholipid syndrome	APS SYNDROME	No, yes	
	Parity and History of preeclampsia	PARITY-HISTORYPE	Nulliparous, parous=yes, parous=no	
	Chronic hypertension	CHRONIC HT	No, yes	
	Renal disease	RENALDISEASE	No, yes	
	Diabetes	DIABETES	No, yes	
	Drugs for Chronic hypertension	TREATMENT-CHT	No, yes	
	Drugs for Renal disease	TREATMENT-RD	No, yes	
	Drugs for Diabetes	TREATMENT-DB	No, yes	
	Family history of preeclampsia	FH-PE	No, fatherPEpreg, mother-sister	
	Family history of hypertension	FH-HT	No, yes	
	Family history of diabetes	FH-DIAB	No, yes	
	Multiple pregnancy	MULTI-PREGNANCY	No, yes	
	Obesity	OBESITY	Normal, overweight, obese	
	Maternal age	AGE	< 20, 21–25, . . . , > 40	
	Smoking	SMOKING	No, yes	
	Signs	Systolic blood pressure (mmHg)	SBP	< 109, 110–119, . . . , 160–169, > 170
		Diastolic blood pressure (mmHg)	DBP	< 59, 60–69, . . . , 100–109, > 110
Hemoglobin (mmol/L)		Hb	6.2, 6.3, . . . , 9.3	
Creatinine ( $\mu\text{mol/L}$ )		CREAT	< 45, . . . , 118–121, > 122	
Protein (Albumin)–Creatinine ratio		PACR	0–0.03, 0.04–0.06, . . . , 4.5–5, > 5	
Extern.	Drugs taken by the patient	TREATMENT	No, Anti-HT, Other, Anti-HT+Other	
Hidden	Vascular risk	VASCRISE	False, true	
	Vascular function	VASCFUNC	Hypotens., normal, hypertens., severe-hypertens.	
	Renal function	RENALFUNC	Ok, nok	
Syndr.	Preeclampsia	PE	No, yes	

The actual development of the Bayesian network model was the central achievement in the eMomCare system development, because such a model does not exist yet. Previous studies for predicting the development of preeclampsia are mostly based on pregnancy-specific markers using logistic regression whose parameter estimation is highly dependent on the sample data used [30,13,15,23]. The static nature of these models can be used for predicting preeclampsia onset given a set of observations at a certain moment of time but they cannot model the temporal trends of the syndrome, thus failing to account for the previous and future patient status. Furthermore, most of the measurements used in the models require collection in a hospital setting, or they are expensive. To alleviate these burdens, we propose a temporal Bayesian network model for preeclampsia, which is manually built using expert knowledge and literature studies, and it is based on simple, easy-to-collect measurements and a set of risk factors. In our previous studies we reported about some preliminary results [39,40]. For the research described in this paper, we extended the model to incorporate all relevant clinical knowledge and we evaluated its performance thoroughly based on the methodology described in Section 3.

#### 4.3. A temporal Bayesian network model for predicting preeclampsia

*Network structure.* We start by providing all relevant variables with their possible values, listed in Table 1. We included common risk factors, which contribute most to the prior risk for preeclampsia according to the gynaecologists and as also reported in [9]. In particular, these factors affect the functioning of the vascular system and hence, they determine the risk of vascular pathology, to which for brevity we refer to as *vascular risk* (the variable VASCRISE in the model).

According to the description of preeclampsia (PE) in Section 4.1, renal and cardiovascular pathology are intimately linked to the syndrome. In our modelling scheme the renal function is represented as a binary variable RENALFUNC with possible values *ok* and *nok* (not ok), indicating normality or abnormality, respectively. In the previous version of the model described in [40], vascular function was also represented as a binary variable. However, the model showed unsatisfactory performance and inability to distinguish well between various blood-pressure conditions. Therefore we refined the domain of the variable VASCFUNC by distinguishing four state values. Several laboratory tests are performed at every checkup to determine the current health status, including the renal and vascular function, of the patient. The most common and easily made measurements are systolic and diastolic blood pressure, informing about vascular function, and haemoglobin, serum creatinine and protein to creatinine ratio in urine reflecting renal function. In addition, the results of these measurements are explained by the presence of certain risk factors; for example, chronic hypertension affects the blood pressure, renal disease impacts the creatinine level and protein-creatinine ratio, and smoking usually increases the level of haemoglobin. These relationships are captured by arcs from the risk factor to these tests. Any treatment at the time of checkup was also incorporated as a causal factor in the renal and vascular function.

To capture the temporal evolution of both renal and vascular function, following the representation in Fig. 2, we created a temporal model by adding links between the respective functional status of successive 10 checkups at 12, 16, 20, 24,



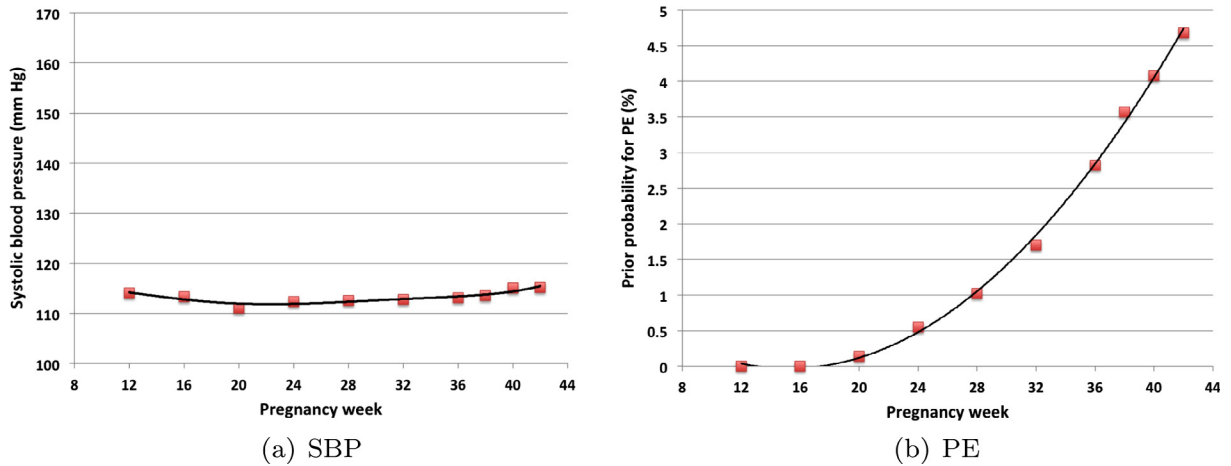


Fig. 4. Expected prior distributions for SBP (a) and the probability of PE (b) over the whole pregnancy in the model (square points) and by domain knowledge (the continuous line).

28, 32, 36, 38, 40 and 42 weeks of pregnancy. These are currently the most common and relevant checkup points during pregnancy. As a result, the combined status of the renal and vascular function describe the development of the syndrome of preeclampsia at each medical checkup captured by the causal  $v$ -structure.

*Network parameters.* We next proceeded with determining the prior and conditional probabilities of the variables in the Bayesian network. With respect to the risk factors this task was easy due to the availability of expert knowledge and reported incidence rates concerning the population of pregnant women (see Section 4.1). To define the CPT of VascRisk for all possible combinations of risk factor values would require a table with 165888 entries, which is practically impossible. Therefore to represent in a compact fashion this probability distribution, we used the causal independence models with the logical OR as an interaction function, as described in Section 3.2. Note that the variables related to the three disorders – chronic hypertension, diabetes and renal disease – influence the vascular risk through a combined effect, represented by one hidden variable for each of them. This causal independence representation of the CPT of VascRisk required the estimation of only 39 independent parameters.

Most lab tests in medicine yield results on a continuous scale; we discretised them into a number of ranges using the expert knowledge of the gynaecologists involved in this research; see the last column in Table 1. Except for PACR, whose expected distribution is fairly constant in the first and second trimester and increases in the third one [44], for the remaining measurements clinical studies indicate that they show a slight decrease in the second trimester and then again increase in the third one [24]. This is reflected in the prior probabilities for tests in the model; see for example the prior distribution of SBP as defined in the model and by domain knowledge, shown in Fig. 4(a). Similarly, we defined the probability distribution  $P(\text{PE}|\text{VASCFUNC}, \text{RENALFUNC})$  based on domain knowledge provided by the gynaecologists stating that the probability for developing preeclampsia increases exponentially towards the end pregnancy, which is due to the increased burden for the woman as the pregnancy progresses (see Fig. 4(b) for the prior probability distribution of PE). Note that at 12 and 16 week PE is officially not diagnosed so the prior probabilities are nearly zero. However, by explicitly modelling these temporal states we facilitate early detection of a decline in the patient's health.

The final model structure consists of 115 nodes and 206 edges. Fig. 5 depicts part of the model, including a subset of the variables, their dependencies and prior probabilities for the first two time points (12 and 16 week of pregnancy). For notational convenience in the remainder of the paper we refer to the final model as the 'PEModel'.

#### 4.4. Evaluation

The evaluation of the proposed model was done using patient data, as described in the next section; the procedure outlined in Section 3.3 was used to account for the dynamics and the uncertainty in the progression of preeclampsia.

*Data.* The dataset concerned pregnancy data collected and pre-processed as part of earlier clinical studies from the period of 2008–2011 at Radboud University Nijmegen Medical Centre, the Netherlands. It includes data spanning the entire pregnancy period of 205 high-risk patients of whom 15 ( $\sim 7.3\%$ ) were diagnosed with preeclampsia (PE). Both the diagnosis of PE and delivery of the child occurred at different stages in pregnancy for preeclamptic and non-preeclamptic patients. We assigned the diagnosis and delivery time to the closest of the 10 checkup points in the model, resulting in the following outcome distribution:

Pregnancy week	28	32	36	38	40	42	Total
Number of diagnosed PE patients	3	2	4	1	4	1	15
Number of non-PE patients who delivered	3	7	6	75	77	22	190
Total number of cases	205	199	190	180	104	23	

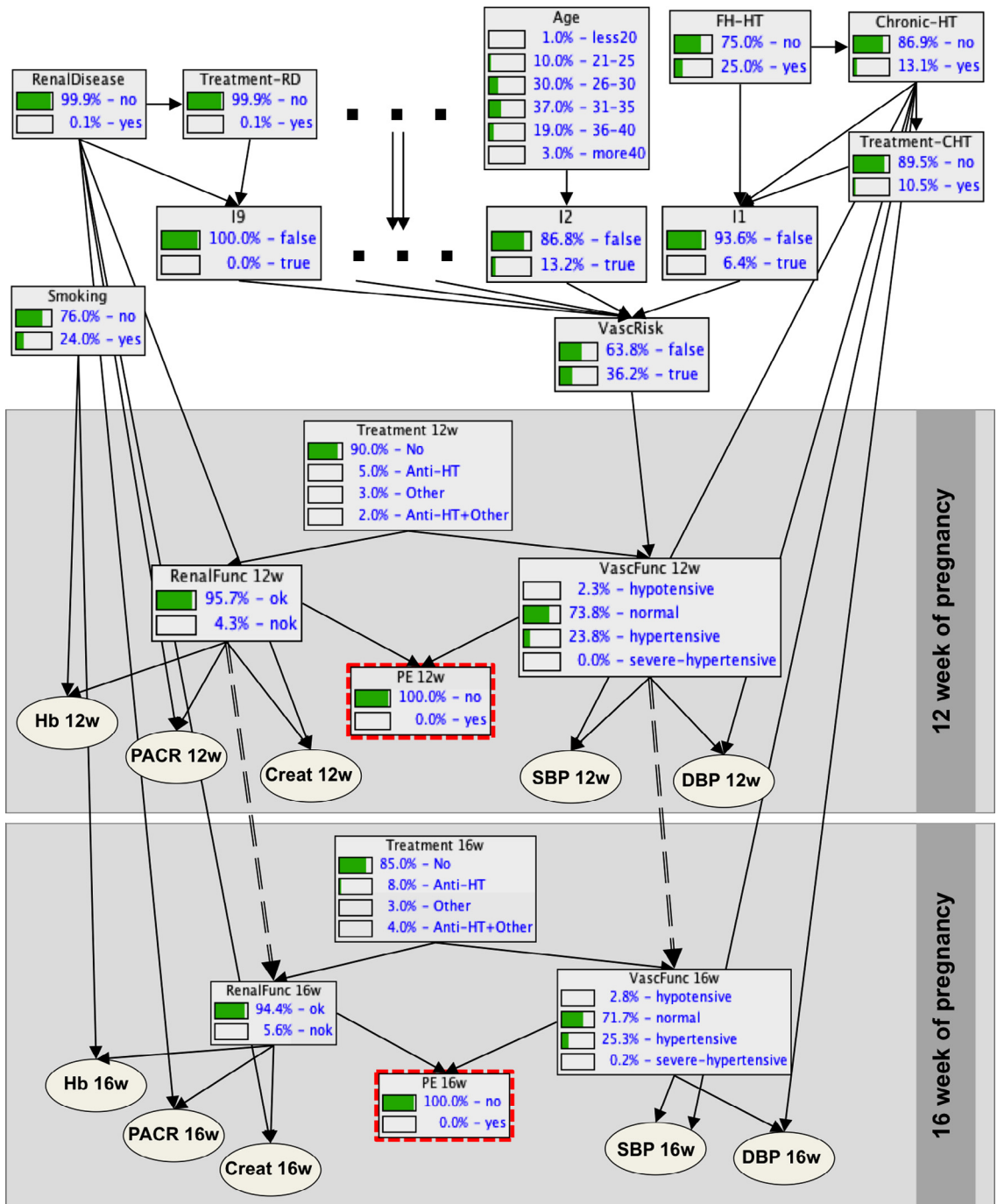


Fig. 5. Part of the temporal Bayesian network model for preeclampsia with prior probabilities. The double dashed arcs represent temporal dependencies between the organ functioning at subsequent time points.

**Table 2**  
Diagnostic accuracy of the PEModel in terms of AUC per week of PE diagnosis.

	Week of PE diagnosis					
	28	32	36	38	40	42
AUC	0.996	0.801	0.868	0.994	0.856	1.00

**Table 3**  
Week of clinical diagnosis vs. week of the PEModel prognosis for each preeclamptic patient at 10% false positive rate.

Patient no	Week of clinical diagnosis	Week of PEModel prognosis	Patient no	Week of clinical diagnosis	Week of PEModel prognosis
1	28	20	9	36	36
2	28	20	10	38	36
3	28	28	11	40	24
4	32	28	12	40	32
5	32	32	13	40	32
6	36	20	14	40	40
7	36	32	15	42	42
8	36	36			

Note that the patient who was diagnosed with PE at week 42 of pregnancy has delivered at term; if the delivery had occurred earlier the patient would have been classified as *non*-preeclamptic. The gynaecologists suggested that it is important to be able to predict the development of early onset PE in patients.

For this dataset only a small set of risk factors was fully recorded, corresponding to the following variables in the PEModel: CHT, TREATMENT-CHT, PARITYHISTPE, OBESITY, AGE and SMOKING. In addition, there were not much data on lab tests. In present pregnancy care in the Netherlands, only blood pressure is measured on regular basis and only when this is high, additional tests such as for PACR, creatinine and haemoglobin are done. The resulting missing values did affect the model's performance (see below for results). Furthermore, the collected laboratory data were previously aggregated by averaging over a number of repeated measurements within a short period of time around the main checkup week. This resulted sometimes in lower measurement results, preventing detection of abnormality, thus affecting again the evaluation results obtained from the model.

**Results.** In Table 2 we report the AUCs obtained for the weeks at which PE was diagnosed in at least one pregnant woman. The results show that the PEModel is able to make a clear distinction between preeclamptic and non-preeclamptic patients. Although for the preeclamptic patients the model gives a high probability of PE, for some of the non-preeclamptic patients this was also the case, which might be explained by the fact that the dataset includes only high-risk patients for whom pregnancy complications are likely. Note that no AUCs have been reported for the pregnancy period before week 28 as then patients are non-preeclamptic by definition.

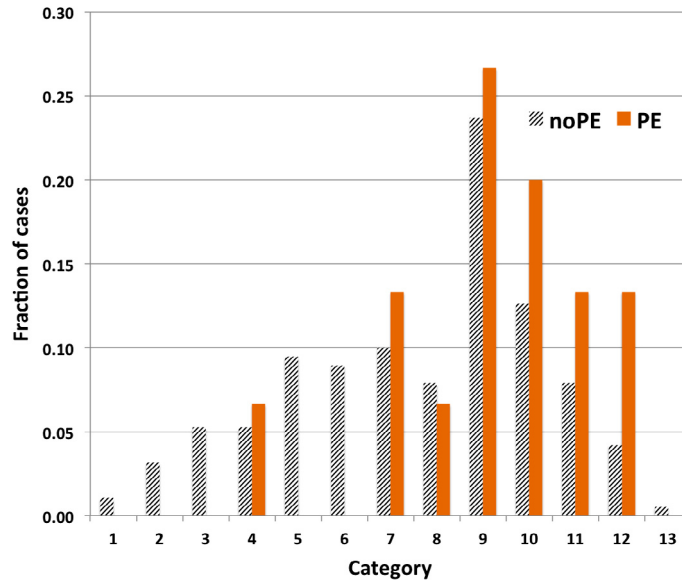
To prevent occurrence of preeclampsia, signs of its development have to be detected as early as possible. Using criterion  $M_1$  from Section 3.3, Table 3 reveals the week at which the PEModel detects an increased risk of preeclampsia for each preeclamptic patient at a 10% FP rate. For 60% of the preeclamptic patients we observe the capability of the PEModel to reliably predict PE at least 4 weeks before the diagnosis was actually made. For the remaining patients this early prediction was not possible due to inaccurate or missing test results. For patient 3 only blood pressure measurements were recorded for the weeks 12–24 and they were low, which prevented detecting potential problems. Similarly, for patient 14 only blood pressure measurements were taken and up to week 36 they were lower than 133/82 mmHg – less than the hypertension diagnostic value of 140/90. For patient 5, the measured PACR was 0 up to the week of diagnosis and blood pressure was lower than 132/82. Patient 8, who had a history of preeclampsia, was obese, in the age range of 36–40 years and she began her pregnancy with a very high blood pressure of 176/96, PACR of 0.15; up to the time of diagnosis these test results remained the same or slightly lower due to the effect of treatment. Thus, the patient remained high-risk up to the time of diagnosis. For patient 9 the blood pressure was also observed at constant level of 132/85 or even lower in the second trimester and PACR was 0 or not measured at some of the checkups, implying that proteinuria was not detected.

Another clinical problem closely related to the early detection of preeclampsia is the distinction between high- and low-risk patients in the initial stage of pregnancy following the procedure described by criterion  $M_2$  in Section 3.3. We computed the probabilities at week 42 (end of pregnancy)  $P_{42wk}(PE = yes)$  given the patient's risk factors, the tests done at 12, 16 and 20 weeks, under the assumption that no treatment was given from week 24 until the end of pregnancy. The resulting probabilities were further distinguished into categories; see Table 4. In Fig. 6 we present for both patient groups (PE and noPE) the histograms of probability intervals. Using the cut-off point of 7.3%, which is the expected percentage of PE patients in this sample of patients, the model is able to correctly identify 93% (14 out 15) of PE patients as being high-risk patients and thus requiring intensive monitoring, vs. a 33% correct classification of non-PE as low-risk. The lower percentage for the non-PE patients may be explained again by the high-risk profile of the patients in the dataset. The only PE patient with low predicted probability at week 42 – category 4 – is patient 3 from Table 3 who only had (low) blood pressure values, not sufficient to detect the development of PE.

Insight into the prognostic capabilities of the PEModel can also be obtained by a comparison of the absolute mean probability  $P(PE = yes)$  between both PE and non-PE patient groups and two subgroups – with and without measured PACR (withPACR and withoutPACR) up to the week when preeclampsia is first diagnosed; see Table 5. The results show that on

**Table 4**  
Splitting up the probability of preeclampsia into categories.

	Category					
	1	2	3	4	5	...
Range for $P(\text{PE} = \text{yes})$ (in %)	< 5.0	5.0–5.5	5.5–6.0	6.0–6.5	6.5–7.0	...



**Fig. 6.** Histograms of different categories of  $P(\text{PE} = \text{yes})$  at week 42 given the patient's risk factors, the measurements made at 12, 16 and 20 week with no treatment from week 24 until the end of pregnancy.

**Table 5**  
Mean and standard deviation (STD. DEV.) of absolute  $P(\text{PE} = \text{yes})$  obtained from PEModel per patient group and two subgroups – with and without measured PACR.

Group	12 week			16 week			20 week			24 week		
	# patients	Mean	STD. DEV.	# patients	Mean	STD. DEV.	# patients	Mean	STD. DEV.	# patients	Mean	STD. DEV.
PE	15	0.00115	0.00215	15	0.024	0.029	15	1.03	0.87	15	2.64	1.8
withPACR	10	0.00146	0.00262	2	0.039	0.055	3	1.22	1.04	9	2.98	2.1
withoutPACR	5	0.00054	0.00030	13	0.022	0.027	12	0.99	0.87	6	2.13	1.4
noPE	190	0.00048	0.00028	190	0.014	0.018	190	0.69	0.58	190	1.71	0.9
withPACR	152	0.00049	0.00027	33	0.014	0.023	39	0.77	0.76	84	1.67	0.8
withoutPACR	38	0.00042	0.00031	157	0.014	0.017	151	0.67	0.53	106	1.74	1.0

average the PEModel produces a higher mean for  $P(\text{PE} = \text{yes})$  for the preeclamptic patients than for the non-preeclamptic ones. For the preeclamptic patients with measured PACR, the PEModel clearly yields a higher mean probability than for those without PACR. For the non-preeclamptic patients rather the opposite trend is observed. These results have an important clinical implication – measuring PACR early in pregnancy can help detecting worsening of the renal function, a sign for potential development of preeclampsia. Appropriate treatment can then be started at a proper, early time.

The model's fit to the dataset was tested using the Bayesian information reward metric from Eq. (1) computed per pregnancy week for PE and non-PE patients and compared to the model's prior; see Table 6. The prior based on the dataset (fraction of preeclamptic cases per week) is also reported and despite the small size of the dataset, it is interesting to observe that this prior is close, except for week 38, to the model's prior determined only on the basis of domain knowledge and without using any data. The  $IR_B$  results for the preeclamptic patients (fourth row) show that the model's posterior probabilities are well distinguished from the prior with high average reward of 0.66 over the six weeks. The low negative values for  $IR_B$  for the non-PE patients show that the estimated probabilities are slightly higher than the prior, thus the model is "punished".

Given the high-risk profile of the patients in the dataset, we expected that the prior distribution of the treatment based on the patient data would differ from the model's prior based on epidemiological information. In Table 7 we present per pregnancy week the model prior probability for "TREATMENT = no" against the proportion of pregnant women who did not receive treatment at that moment. The results show that more than 60% of the women in the dataset are under treatment already at the beginning of pregnancy, which is different from the model's prior of up to 20%. This clearly indicates that the data is not representative for the entire population of pregnant women. However, usually a treatment variable is instantiated,

**Table 6**

Bayesian information reward computed per week with respect to the model's prior probabilities.

Week of pregnancy	28	32	36	38	40	42
Prior $P(\text{PE} = \text{yes})$ (model)	0.010	0.017	0.028	0.036	0.041	0.047
Prior $P(\text{PE} = \text{yes})$ (dataset)	0.015	0.010	0.021	0.006	0.038	0.043
$IR_B$ (diagnosed PE patients)	0.869	0.471	0.699	0.618	0.581	0.694
$IR_B$ (non-PE patients)	-0.008	-0.012	-0.019	-0.021	-0.026	-0.024

**Table 7**

Prior probabilities for TREATMENT = No based on the model and the dataset.

Pregnancy week	12	16	20	24	28	32	36	38	40	42
Prior $P(\text{TREATMENT} = \text{No})$ (model)	0.90	0.85	0.845	0.835	0.827	0.812	0.80	0.795	0.79	0.79
Prior $P(\text{TREATMENT} = \text{No})$ (dataset)	0.38	0.36	0.34	0.30	0.28	0.27	0.25	0.27	0.31	0.26



(a) Status (b) Status: Measurement analysis (c) Status: Prognosis for PE risk

**Fig. 7.** Snapshots of the system's user interface based on the temporal Bayesian network on an Android smartphone.

and then what matters are only to what extent the probabilistic influences from treatment to the other variables reflect the situation for the entire group of patients.

4.5. Embedding the model in the eMomCare system

The Bayesian network is part of the eHealth system eMomeCare, implemented using Google's open source Android operating system. The Bayesian network model is consulted using the publicly available java-based Bayesian network package EBayes [7], which provides platform-independent engine for performing inference given evidence in the network. A java-version of the network was subsequently embedded in the Android application. We are still investigating a suitable design for the user interface. Fig. 7 shows the current main screens – current health status and advice, measurement analysis and a prognostic chart for the development of preeclampsia until the end of pregnancy – based on the results obtained from the PEModel. The relevant input data–risk factors and measurements – are provided via specially-designed screens (tabs 'Clinic' and 'Lab') in the mobile app. Blood pressure and urine analysis are the most important measurements and they can be performed at home – the former via a digital monitor that is capable of transmitting the data via Bluetooth to the smartphone and the latter via urine reagent strips that are imaged and the colour is automatically analysed in the device. The patient's status is computed based on the input data and the temporal Bayesian network model. The result of the inference process on the smartphone is obtained within 10 seconds, despite the large size of the network.

5. Conclusions and lessons learned

In this research we have described a generic Bayesian network methodology for building probabilistic disease progression models based on clinical principles. We showed that exploiting physiological knowledge in developing a clinical Bayesian network can guide the building process. This was illustrated by the description of the construction of a temporal Bayesian

network model for preeclampsia. This model is part of a real-world home-monitoring system for personalised pregnancy care. Insights into the performance of the model were obtained through an experimental study with actual patient data.

The system's capability to offer personalised advice to patients required a model that was able to provide feedback about the patient's current and expected future health status. This meant that the model had to be based on general clinical principles, while capturing the relevant functional relationships of the disease to allow personalisation of the advice. There were also requirements on the system with respect to its capability of generating an explanation of the model's predictions. The fact that sufficient relevant knowledge could be obtained from clinical experts and clinical studies, whereas clinical data was only scarcely available, motivated us to opt for manual construction of the model rather than to learn the model from data. To our knowledge, this is the first clinical model built for use in personalisation of treatment in the home environment using sensors and smartphones.

Collecting a dataset of sufficient size and quality appeared to be a major problem in our research. The original version of the dataset used in this study contained data of 297 pregnant women, but for 92 of them no outcome, i.e. whether or not preeclampsia had developed, had been recorded. These patients were discarded for the evaluation. The same problem we encountered in another cohort of 534 patients from whom only 2 patients were recorded as being preeclamptic, whereas many more should have been included based on the prior probability of getting preeclampsia. According to clinical experts, these problems were due to the fact that the data were just collected as part of daily practice and not for research purposes. We have now set up an improved data collection process with the aim of collecting sufficient data to refine the network parameter and for further evaluation of the model. In this respect, the advantages of a home-monitoring system can be exploited, as in principle it gives access to a large collection of clinical and laboratory data, simply by involvement of sufficient numbers of patients in the research.

Another critical and nontrivial task in our research was the evaluation of the developed Bayesian-network model. Adopting the traditional distinction between positive and negative cases was not sufficient to test the model's performance because preeclamptic patients (positive cases) are considered negative until the moment of diagnosis. So, for example, using standard accuracy measures to check whether or not the model can correctly classify patients based on the output at first trimester does not make much sense – at that time all the patients are non-preeclamptic. So, diagnostic accuracy was determined only for the time when positive patients were available. Guided by the gynaecologists involved in the project, we were also able to get insight in the prognostic capabilities of the model using the model's outcome at early pregnancy. In this respect the model showed encouraging results by detecting an increased risk for preeclampsia for the majority of preeclamptic patients between 4 and 16 weeks in advance.

Furthermore, in applications where personalisation is a critical issue, the overall model evaluation should be based not only on overall performance but also on per-patient analysis. When performing the latter we observed that for preeclamptic patients with missing or non-compatible data the model performed worse, by underestimating, for example, the probability of preeclampsia. Another problem is that non-preeclamptic patients may also experience worsening of their health status, not necessarily leading to preeclampsia. In the current model, such patients get a higher probability for preeclampsia, which also deteriorates the model's performance. Of course, from a clinical point of view, this is not necessarily a problem as the costs for false positives are generally lower than those for false negatives but unnecessary pressure on the patient and caregiver should also be minimised.

A generic temporal, causal framework for clinical disease models, such as proposed in this paper, certainly made model development easier. A detailed analysis of the absolute probabilities for the first and second trimester computed from the model showed a clear trend implying that not only blood-pressure measurement, as currently done, but also carrying out a urine analysis on a regular basis may help detecting preeclamptic patients at earlier stages of pregnancy. When this is confirmed by more extensive studies, taking into account health-care costs, it is likely that the model leads to changes to present-day clinical practice.

Our future work concerns studying the effect of treatment per pregnancy week on the model's performance in relation to actual patients, and whether different treatment scenarios can help prevent worsening of the patient's condition. Furthermore, we prepare a prospective pilot study, where the mobile application with the embedded temporal Bayesian network model will be used in parallel to the standard clinical care. This will allow us to compare the recommendations given by the model concerning, for example, the number of checkups and contacts with the clinician against present care standards.

## Acknowledgement

The work has been funded by the STITPRO Foundation, The Netherlands.

## References

- [1] S. Andreassen, Planning of therapy and tests in causal probabilistic networks, *Artificial Intelligence in Medicine* 4 (3) (1992) 227–241.
- [2] S. Andreassen, M. Woldbye, B. Falck, S. Andersen, MUNIN – a causal probabilistic network for interpretation of electromyographic findings, in: *Proceedings of the 10th International Joint Conference on Artificial Intelligence*, Milan, Italy, 1987, pp. 366–372.
- [3] S.H. Bach, M. Broecheler, S. Kok, L. Getoor, Decision-driven models with probabilistic soft logic, in: *NIPS Workshop on Predictive Models in Personalized Medicine*, 2010.

- [4] I. Beinlich, H. Suermondt, R. Chavez, G. Cooper, The ALARM monitoring system: a case study with two probabilistic inference techniques for belief networks, KSL-Report 88-84, Section on Medical Informatics, Stanford University, 1988.
- [5] C. Bielza, J.F. del Pozo, P.J.F. Lucas, Explanation of clinical decisions through the extraction of regularity patterns, *Journal of Decision Support Systems* 44 (2008) 397–408.
- [6] A. Ciabattini, D.P. Muiñ, T. Vetterlein, M. El-Zekey, Formal approaches to rule-based systems in medicine: the case of CADIAG-2, *International Journal of Approximate Reasoning* 54(1) (2013) 132–148.
- [7] F.G. Cozman, 1999. EBayes: embedded Bayesian networks. <<http://www.cs.cmu.edu/javabayes/EBayes/index.html>>.
- [8] P. Dagum, A. Galper, Forecasting sleep apnea with dynamic models, in: *Proceedings of UAI'94*, 1994, pp. 64–71.
- [9] K. Duckitt, D. Harrington, Risk factors for pre-eclampsia at antenatal booking: systematic review of controlled studies, *BMJ* 330 (7491) (2005) 565–571.
- [10] D. Heckerman, J.S. Breese, Causal independence for probability assessment and inference using Bayesian networks, *IEEE Transactions on SMC – A* 26 (6) (1996) 826–831.
- [11] D.E. Heckerman, E.J. Horvitz, B.N. Nathwani, Toward normative expert systems: part II – probability-based representations for efficient knowledge acquisition and inference, *Methods of Information in Medicine* 31 (1992) 106–116.
- [12] D.E. Heckerman, B.N. Nathwani, Toward normative expert systems: part I – the Pathfinder project, *Methods of Information in Medicine* 31 (1992) 90–105.
- [13] I. Herraiz, J. Arbues, I. Camano, E. Gomez-Montes, A. Graneras, A. Galindo, Application of a first-trimester prediction model for pre-eclampsia based on uterine arteries and maternal history in high-risk pregnancies, *Prenatal Diagnosis* 29 (12) (2009) 1123–1129.
- [14] K.B. Kevin, A.E. Nicholson, *Bayesian Artificial Intelligence*, first ed., Chapman and Hall/CRC, 2003.
- [15] A.A. Khalil, N.J. Cowans, K. Spencer, S. Goichman, H. Meiri, K. Harrington, First trimester maternal serum placental protein 13 for the prediction of pre-eclampsia in women with a priori high risk, *Prenatal Diagnosis* 29 (8) (2009) 781–789.
- [16] M. Korver, P.J.F. Lucas, Converting a rule-based expert system into a belief network, *Medical Informatics* 18 (3) (1993) 219–241.
- [17] P.J.F. Lucas, Knowledge acquisition for decision-theoretic expert systems, *AISSB Quarterly* 94 (1996) 23–33.
- [18] P.J.F. Lucas, H. Boot, B. Taal, Computer-based decision-support in the management of primary gastric non-hodgkin lymphoma, *Methods of Information in Medicine* 37 (1998) 206–219.
- [19] P.J.F. Lucas, N. de Bruijn, K. Schurink, I. Hoepelman, A probabilistic and decision-theoretic approach to the management of infectious disease at the ICU, *Artificial Intelligence in Medicine* 19 (3) (2000) 251–279.
- [20] P.J.F. Lucas, L.C. van der Gaag, A. Abu-Hanna, Bayesian networks in biomedicine and health-care, *Artificial Intelligence in Medicine* 30 (2004) 201–214.
- [21] L.J. Mazlack, Causal modeling approximations in the medical domain. In: *Proceeding of 2011 IEEE International Conference on Fuzzy Systems*, 2011, pp. 1822–1829.
- [22] R.B. Ness, J.M. Roberts, Epidemiology of pregnancy-related hypertension, in: M.D. Lindheimer, J.M. Roberts, F.G. Cunningham (Eds.), *Chesley's Hypertensive Disorders in Pregnancy*, Academic Press, 2009 (Chapter 3).
- [23] M. Nijdam, K.J.M. Janssen, K.G.M. Moons, D.E. Grobbee, J.A. van der Post, M.L. Bots, A. Franx, Prediction model for hypertension in pregnancy in nulliparous women using information obtained at the first antenatal visit, *Journal of Hypertension* 28 (1) (2010) 119–126.
- [24] N. Ochsenein-Köblle, M. Roos, T. Gasser, R. Huch, A. Huch, R. Zimmermann, Cross sectional study of automated blood pressure measurements throughout pregnancy, *International Journal of Obstetrics & Gynaecology* 111 (4) (2004) 319–325.
- [25] K. Olesen, S. Andreassen, Specification of models in large expert systems based on causal probabilistic networks, *Artificial Intelligence in Medicine* 5 (3) (1993) 269–281.
- [26] K. Olesen, U. Kjaerulff, F. Jensen, F. Jensen, B. Falck, S. Andreassen, S. Andersen, A MUNIN network for the median nerve – a case study on loops, *Applied Artificial Intelligence* 3 (2–3) (1989) 385–403.
- [27] R.S. Patil, P. Szolovits, W.B. Schwartz, Causal understanding of patient illness in medical diagnosis, in: *Proceedings of the 7th International Joint Conference on Artificial Intelligence*, vol. 2, 1981, pp. 893–899.
- [28] M. Paul, S. Andreassen, A. Nielsen, E. Tacconelli, N. Almanasreh, A. Fraser, D. Yahav, R. Ram, L. Leibovici, Prediction of bacteremia using TREAT, a computerized decision-support system, *Clinical Infectious Diseases* 42 (2006) 1274–1282.
- [29] M. Paul, S. Andreassen, E. Tacconelli, A. Nielsen, N. Almanasreh, U. Frank, R. Cauda, L. Leibovici, Improving empirical antibiotic treatment using TREAT, a computerized decision support system: cluster randomized trial, *Journal of Antimicrobial Chemotherapy* 58 (2006) 1238–1245.
- [30] L. Poon, N.A. Kametas, N. Maiz, R. Akolekar, K.H. Nicolaides, First-trimester prediction of hypertensive disorders in pregnancy, *Hypertension* 53 (5) (2009) 812–818.
- [31] G. Provan, J. Clarke, Dynamic network construction and updating techniques for the diagnosis of abdominal pain, *Transactions on Pattern Analysis and Machine Intelligence* 15 (3) (1993) 299–307.
- [32] G.W. Rudedge, S.K. Andersen, J.X. Polaschek, L.M. Fagan, A belief network model for interpretation of ICU data, in: *Proceedings of the Annual Symposium on Computer Application in Medical Care*, 1990, pp. 785–789.
- [33] G.D. Sanders, R.A. Harris, M.A. Hlatky, D.K. Owens, Prevention of sudden cardiac death: a probabilistic model for decision support, in: *Proceedings of the Annual Symposium on Computer Application in Medical Care*, 1995, pp. 258–262.
- [34] M. van Gerven, B. Taal, P.J.F. Lucas, Dynamic Bayesian networks as prognostic models for clinical patient management, *Journal of Biomedical Informatics* 41 (2008) 515–529.
- [35] M. van Gerven, F. Díez, B. Taal, P.J.F. Lucas, Selecting treatment strategies with dynamic limited-memory influence diagrams, *Artificial Intelligence in Medicine* 40 (3) (2007) 171–186.
- [36] M.A.J. van Gerven, R. Jurgelenaitė, B.G. Taal, T. Heskes, P.J.F. Lucas, Predicting carcinoid heart disease with the noisy-threshold classifier, *Artificial Intelligence in Medicine* 40 (2007) 45–55.
- [37] M.A.J. van Gerven, P.J.F. Lucas, T.P. van der Weide, A generic qualitative characterization of independence of causal influence, *International Journal of Approximate Reasoning* 48 (2008) 214–236.
- [38] M. Velikova, P.J.F. Lucas, M. Samulski, N. Karssemeijer, A probabilistic framework for image information fusion with an application to mammographic analysis, *Medical Image Analysis* 16 (2012) 865–875.
- [39] M. Velikova, P.J.F. Lucas, M. Spaanderman, e-MomCare: a personalised home-monitoring system for pregnancy disorders, *Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering* 69 (2011) 267–274.
- [40] M. Velikova, P.J.F. Lucas, M. Spaanderman, A predictive Bayesian network model for home management of preeclampsia, in: *Lecture Notes in Artificial Intelligence*, vol. 6747, 2011, pp. 179–183.
- [41] M. Velikova, M. Samulski, P.J.F. Lucas, N. Karssemeijer, Improved mammographic cad performance using multi-view information: a Bayesian network framework, *Physics in Medicine and Biology* 54 (2009) 1131–1147.
- [42] M. Verduijn, N. Peek, P.M.J. Rossee, E. de Jonge, B.A.J.M. de Mol, Prognostic Bayesian networks, *Journal of Biomedical Informatics* 40 (6) (2007) 609–630.
- [43] S. Visscher, P.J.F. Lucas, C.A.M. Schurink, M.J.M. Bonten, Modelling treatment effects in a clinical Bayesian network using Boolean threshold functions, *Artificial Intelligence in Medicine* 46 (2009) 251–266.
- [44] J. Waugh, S. Bell, M. Kilby, P. Lambert, C. Blackwell, A. Shennan, A. Halligan, Urinary microalbumin/creatinine ratios: reference range in uncomplicated pregnancy, *Clinical Science* 104 (2) (2003) 103–107.
- [45] W. Wiegierinck, H. Kappen, E. ter Braak, W. ter Burg, M.O.Y. Nijman, J. Neijt, Approximate inference for medical diagnosis, *Pattern Recognition Letters* 20 (1999) 1231–1239.
- [46] Y. Xiang, J. Smith, J. Kroes, Multiagent Bayesian forecasting of structural time-invariant dynamic systems with graphical models, *International Journal of Approximate Reasoning* 52 (2011) 960–977.
- [47] A. Zalounina, M. Paul, L. Leibovici, S. Andreassen, A stochastic model of susceptibility to antibiotic therapy – the effects of cross-resistance and treatment history, *Artificial Intelligence in Medicine* 40 (1) (2007) 57–63.