Available online at www.sciencedirect.com

ScienceDirect

Journal of Biomedical Informatics 40 (2007) 609–618

Journal of Biomedical Informatics

www.elsevier.com/locate/yjbin

Prognostic Bayesian networks I: Rationale, learning procedure, and clinical use

Marion Verduijn [a,e,*], Niels Peek [a], Peter M.J. Rosseel [d], Evert de Jonge [b], Bas A.J.M. de Mol [c,e]

[a] Department of Medical Informatics, Academic Medical Center (AMC), P.O. box 22700, 1100 DE Amsterdam, The Netherlands
[b] Department of Intensive Care Medicine, AMC, Amsterdam, The Netherlands
[c] Department of Cardio-thoracic Surgery, AMC, Amsterdam, The Netherlands
[d] Department of Anesthesia and Intensive Care, Amphia Hospital, Breda, The Netherlands
[e] Department of Biomedical Engineering, University of Technology, Eindhoven, The Netherlands

Received 3 August 2006
Available online 25 July 2007

Abstract

Prognostic models are tools to predict the future outcome of disease and disease treatment, one of the fundamental tasks in clinical medicine. This article presents the prognostic Bayesian network (PBN) as a new type of prognostic model that builds on the Bayesian network methodology, and implements a dynamic, process-oriented view on prognosis. A PBN describes the mutual relationships between variables that come into play during subsequent stages of a care process and a clinical outcome. A dedicated procedure for inducing these networks from clinical data is presented. In this procedure, the network is composed of a collection of local supervised learning models that are recursively learned from the data. The procedure optimizes performance of the network's primary task, outcome prediction, and handles the fact that patients may drop out of the process in earlier stages. Furthermore, the article describes how PBNs can be applied to solve a number of information problems that are related to medical prognosis.

Keywords: Prognosis; Health care process; Bayesian network; Network learning procedure

1. Introduction

Prognostic models have become important instruments in medicine. Given a set of patient specific parameters, they predict the future occurrence of a medical event or outcome. Example events are the occurrence of specific diseases (e.g., cardiovascular diseases and cancer) and death. The models are used for prediction purposes at levels that range from individual patients (where their predictions help doctors and patients to make treatment choices) to patient groups (where they support health care managers in planning and allocating resources) and patient populations (where they provide for case-mix adjustment) [1,2].

Prognostic models are usually induced from historical data by applying supervised data analysis methods such as multivariate logistic regression analysis or tree induction. This approach has three limitations. First, supervised data analysis methods apply attribute selection before inducing a model, often removing many attributes that are deemed relevant for prognosis by users of the model (e.g., clinicians). Second, the resulting models regard prognosis to be a one-time activity at a predefined time. In reality, however, expectations with respect to a patient's future may regularly change as new information becomes available during a disease or treatment process. And third, the models impose fixed roles of predictor (independent variable, input) and outcome variable (dependent variable, output) to the attributes involved. This approach ignores

* Corresponding author. Address: Department of Medical Informatics, Academic Medical Center, P.O. box 22700, 1100 DE Amsterdam, The Netherlands. Fax: +31 20 6919840.
E-mail address: m.verduijn@amc.uva.nl (M. Verduijn).

1532-0464/$ - see front matter © 2007 Elsevier Inc. All rights reserved.
doi:10.1016/j.jbi.2007.07.003

the dynamic nature of care processes, where today's outcome helps to predict what will happen tomorrow.

This article introduces a new type of prognostic model based on the Bayesian network methodology [3], that overcome these limitations. Since the introduction of Bayesian networks in the 1980s, a large number of applications have been developed in different medical domains. Most of the applications aim to support diagnosis, e.g., [4–7] and therapy selection, e.g., [8–10]. Prognostic applications of Bayesian networks form a rather new development [11], and are relatively rare [12–15]. The *prognostic Bayesian network* (PBN) provides a structured representation of a health care process by modeling the mutual relationships among variables that come into play in the subsequent stages of the care process and the outcome. As a result, the PBN allows for making predictions at various times during a health care process, each time using all the available information of the patient concerned. Furthermore, prognostic statements are not limited to outcome variables, but can be obtained for all variables that occur beyond the time of prediction.

This article presents the rationale of PBNs and a dedicated procedure to learn a PBN from local supervised learning models, and describes the functionality of PBNs in clinical practice. In a companion article, an application of the learning procedure in the domain of cardiac surgery is described [16].

The article is organized as follows. In Section 2, the PBN is placed in the field of prognostic models. Section 3 presents the procedure for PBN learning from data. In Section 4, we describe prognostic uses of PBNs in clinical practice. We conclude the article with a discussion and conclusions in Section 5.

2. Representation and functionality of prognostic models

Prognostic models describe the relationship between predictor and outcome variables. The standard methodology to obtain an objective description of this relationship is building predictive models from a set of observed patient data and outcomes [17,18]. Generally, the first step in the process is to choose a time of prediction, such as hospital admission. All patient data that are available at this time are then taken into account for model development. Subsequently, variables that are found to have predictive value for the outcome are selected for inclusion of the model (feature selection). The relation between the predictors and the outcome variable is described by the function $Y = f(\mathbf{X})$ using supervised learning methods (e.g., logistic regression), where Y is the outcome variable and $\mathbf{X}$ are the predictors. We refer to the resulting prognostic models as *traditional models* [19–21].

The methodology described above is illustrated in Fig. 1a. The figure shows a prediction problem in a health care process that can be regarded as a template of a care process in which a medical intervention is performed; the intervention is preceded by a stage of diagnosis and treatment selection, and followed by a stage of recovery. The problem is prediction of the outcome hospital mortality with five variables as available predictors. The variables are observed at different times in the care process, and are interrelated. The prediction time is predefined as 'prior to the intervention'. Therefore, the predictors that are observed before the intervention are taken into account and later predictors are excluded from the modeling process. Using a standard supervised learning method, the variables that describe a patient's condition before the

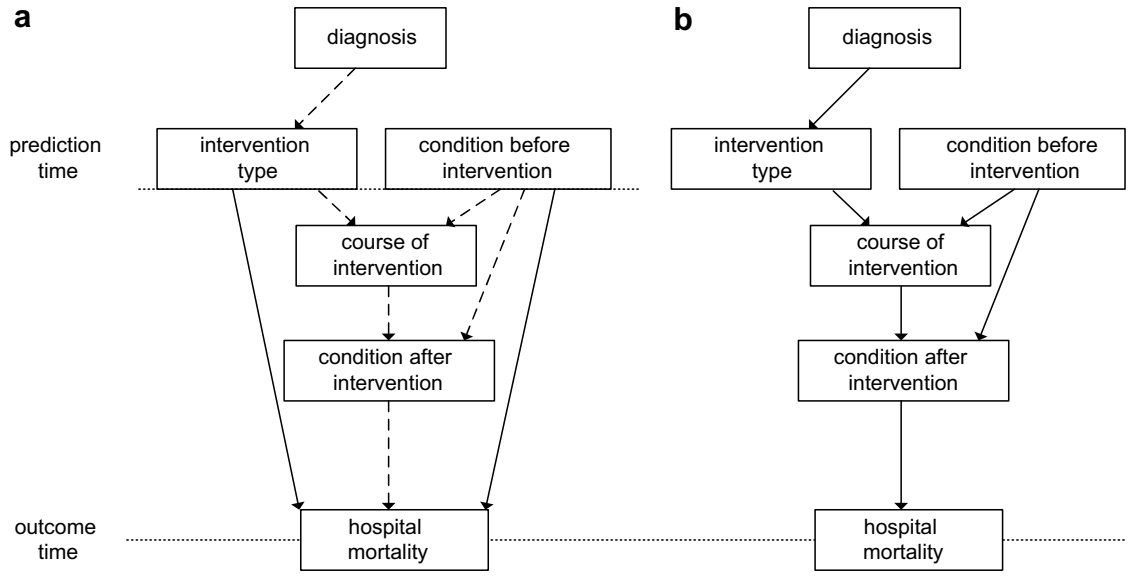

Fig. 1. Modeling a prediction problem of hospital mortality with five variables as available predictors in (a) a traditional model and (b) a prognostic Bayesian network as structured model; the solid arcs represent relationships that are described in the models, and the dotted arcs represent relationships that remain obscured in the traditional model.

intervention and the intervention type are then selected and their relation with hospital death is described in a predictive model. Although a patient's diagnosis has predictive value for the outcome, this variable is ignore and not included in the model; it is shielded from the outcome by the intervention type due to the strong relationship between these variables. Furthermore, the model does not reveal that the relationship between the intervention type and the outcome actually passes through the variables that describe the course of intervention and a patient's condition afterwards. The dotted arcs in Fig. 1a represent relationships that remain obscured, while the solid arcs represent relationships that are described in the model.

This approach of predictive modeling has in our view some shortcomings, as a result of which the traditional model has limited functionality. First, prediction is assumed as a one-time activity at a predefined time; the model can not be used to update the prognostic expectations based on data that become available as the process progresses. Second, the model does not reflect that the predictors are related to the outcome variable through a process of intermediate variables by excluding all variables beyond the prediction time from the modeling process. Third, the feature selection step can be misleading and not intuitive for clinicians, because not all variables that have predictive value are generally included in the model. In case of collinearity among two predictive variables usually only one of them is included, while the other variable is left out; which variable is included may depend on chance [18].

To overcome the shortcomings of traditional predictive modeling, researchers have examine new approaches, such as spline regression analysis, artificial neural networks, and genetic algorithms [22]. These methods, however, are mainly aimed to overcome shortcomings with respect to assumptions of linearity and additivity that may not hold for a modeling problem.

In this article, we propose to model the mutual relationships among variables that come into play in a health care process and the outcome as a Bayesian network to solve the above-mentioned shortcomings of the traditional modeling approach. Fig. 1b shows the PBN structure for the above prediction problem. The direction of arcs in the network structure represents the flow of time. The PBN has no predefined prediction time, and imposes no fixed roles of predictor and outcome variable to the variables involved. As such, the PBN implements a process-oriented view on prognosis which can be examined at any time during the health care process. The methodology that underlies the PBN also allows the analysis of scenarios that lead to disease outcomes.

The health care processes modeled in PBNs are composed of a sequence of substantially different phases, and have no recurring character such as a Markov process [23]. The observed variables are mainly phase-specific and not repeatedly measured during the process. So, although time is an important factor, the data are not suitable to be modeled as a dynamic or temporal Bayesian network [24], as used for prognostic modeling of repeated measurements in [25].

3. Learning a prognostic Bayesian network from local models

In the past decade, several algorithms for learning Bayesian networks from data have been developed, e.g., [26–30], and implemented in different software tools.[1] Applying these algorithms Bayesian network learning is considered an unsupervised learning task. No variable is considered to be more important than any other variables, and the network structure is built up by recursively adding arcs between pairs of variables that appear most strongly correlated in the data. Furthermore, dedicated learning algorithms have been developed for Bayesian network classifiers [31]. These algorithms optimize the networks for their intended use, classification of a predefined variable [32,33]. Similar, a final outcome variable exists in PBNs, whose accurate prediction is of principal importance, and preference must be given to the prediction task during the construction of the model.

The algorithms for learning Bayesian network (classifiers) assume that all variables are meaningful for each case in the data set (i.e., the network is learned from a 'flat table'). This assumption fails for PBN learning due to the fact that not all patients who enter the care process being modeled actually pass through all stages of the entire process, as patients may die during early stages of care or end therapy. Variables that are observed in the later stages of the care process are irrelevant for these patients. We refer to this phenomenon as *patient dropout*. This section presents a dedicated procedure to induce a PBN from local supervised learning models. The procedure exploits the temporal structure of the health care process being modeled, optimizes the performance of the network's primary task, outcome prediction, and adequately handles patient dropout.

3.1. The learning procedure

First, we introduce some notation. Let $\mathbf{X} = \{X_1, \ldots, X_m\}$ denote a set of random variables. Let X_m denote the outcome variable of the process described by $\mathbf{X}$; X_m is therefore also denoted by Y. We use $G = (\mathbf{X}, A)$ to denote the graphical part of the Bayesian network, where $A \subseteq \mathbf{X} \times \mathbf{X}$ is set of ordered pairs that represent arcs. The procedure assumes all continuous variables to be discretized prior to network learning. To ensure that the flow of time is captured in the network structure, the procedure requires a temporal sequence depending on the time and order that the variables are observed. Let $s(X_i) = t$ denote the tempo-

[1] For an overview of available software tools for Bayesian networks see: http://www.cs.ubc.ca/~murphyk/Software/BNT/bnsoft.html.

ral stratum of variable X_i, where t is the index of the stratum of this variable ($1 \leqslant t \leqslant T$); the outcome variable is in the highest stratum, $s(Y) = T$.

The learning procedure is based on the following correspondence. Building the graphical part of a Bayesian network boils down to selecting, for each variable X_i, a set S_{X_i} of 'nearby' variables that separate X_i from all other variables. The set S_{X_i} is called the *Markov blanket* of variable X_i; given this set, X_i should be conditionally independent of all other variables (in the probability distribution that generated the data). Finding the Markov blanket S_{X_i} corresponds to selecting the best predictive feature subset for variable X_i in the data, a typical supervised machine learning problem. So, we can build a Bayesian network by selecting the best predictive feature subset in our data for each variable that is to be included in the network, and transform these feature subsets into Markov blankets by drawing the corresponding arcs in the graph.

The transformation of a collection of feature subsets into a graphical representation is not trivial, though. In PBNs, we require the direction of arcs to be consistent with the flow of time in the medical process. We therefore exploit the temporal structure on the variables as defined in terms of the temporal strata during the learning process. We start network learning with an empty graph (no arcs), consisting only of nodes that represent the predictor variables and one node to represent the outcome variable, and perform feature subset selection in a top-down approach, starting with the outcome variable of the process. For this variable, a feature subset is selected and a predictive model is built from the data using a supervised learning algorithm, such as generalized linear regression analysis and tree induction. As the outcome variable is known to be a sink node in the graph, all selected features for this variable can be represented as parent nodes. Subsequently, for each variable that occurs in this subset of selected features, the unknown part of the feature subset (i.e., the parent nodes) is selected and a predictive model is built. This feature subset selection and local model building is recursively applied until a feature subset has been assessed for each variable in the network. The set of selected features is used as the set of parents of the variable, and represented as such with incoming arcs in a graph, while the local predictive model is used to represent the conditional probability distribution of the variable given its parents in the network. Using this procedure, we arrive at a directed acyclic graph as graphical part of the Bayesian network, and a collection of local predictive models as the numerical part. They jointly constitute the PBN.

We now describe the learning procedure in more detail. The learning procedure includes five steps. Step III and Step V are related to network learning in case of patient dropout; these steps are therefore described in Section 3.3. Initially, we assume that the phenomenon of patient dropout does not occur, so that all patients pass through the entire care process.

3.1.1. Step I

The learning procedure starts with the empty graph $G = (\mathbf{X}, \varnothing)$. In the first iteration of the procedure, a predictive model for outcome Y with predictive features from the set $\{X_i \in \mathbf{X} | X_i \neq X_m\}$ is induced from the data to assess the set of parents and a local model for Y in the Bayesian network. Let S_Y denote the set of features that have been included in the model. Arcs are added to graph G from the selected features in set S_Y to the outcome Y; these features thus become parent nodes of Y. The predictive model is used as the local conditional probability model for Y in the network.

3.1.2. Step II

The learning procedure proceeds by recursively applying this step to all variables in the network, starting with the selected features in the set S_Y. For that purpose, the selected features in set S_Y are enqueued in a priority queue, denoted by Q. The 10-fold cross validated information gain ΔI for the outcome Y is used as priority value. The estimated information gain ΔI is defined as

$$\Delta I = H(P(Y = \mathbb{T})) - \frac{1}{n}\sum_{j=1}^{n} H(P(Y = \mathbb{T} | X_i = x_{i,j})), \qquad (1)$$

where $H(p) = -p\log_2 p$, n is the number of observations in the learning set, and $P(Y = \mathbb{T} | X_i = x_{i,j})$ is the conditional probability that $Y = \mathbb{T}$ given the observed value of variable X_i for observation j in this set [34].

In the second iteration of the learning procedure, variable X_i with the highest (univariate) predictive value for outcome Y is dequeued from priority queue Q. A set of parents is assessed for variable X_i by selecting a feature subset from its potential predictors, and their relation is modeled using the supervised learning algorithm. A potential predictive feature for variable X_i is each other variable $X_j, X_i \neq X_j$, that is not in a higher temporal stratum than X_i, $\sigma(X_j) \leqslant \sigma(X_i)$, and is no descendant of X_i in the current graph. Let the set of all descendants of variable X_i in the current graph G, including X_i itself, be denoted by $\sigma^*_G(X_i)$. The set of potential features for variable X_i is then $R_{X_i} = \{X_j \in \mathbf{X} | (\sigma(X_j) \leqslant \sigma(X_i), X_j \notin \sigma^*_G(X_i))\}$. Let $S_{X_i} \subseteq R_{X_i}$ denote the set of features that are selected for variable X_i. Arcs are added in the graph from the selected features in set S_{X_i} to the variable X_i to designate these features as parent nodes of variable X_i. Subsequently, the selected features in the set S_{X_i} are enqueued in priority queue Q, if they had not been enqueued before. This procedure is repeated until the queue is empty.

3.1.3. Step IV

At this point in the learning procedure, there may exist some variables that were never selected in any feature subset and therefore remain as free nodes in the graph. There are two explanations for this. First, the variable are independent of any feature in the network, or second, they are conditional independent of later process and outcome

variables given other variables in the network. The second explanation can be illustrated with the following example. Suppose there is a variable X_1 in stratum t and the variables X_2 and Y in stratum $t+1$. If $Y \perp\!\!\!\perp X_2 | X_1$, variable X_2 will not be included in the network using the above procedure, despite the fact that $X_2 \not\perp\!\!\!\perp Y$. The reason for this is that after selection of variable X_1 for outcome Y, the learning procedure will proceed with feature subset selection for X_1; variable X_2, however, is no potential predictor for X_1, as it is in a higher stratum and will be excluded from the learning process. This example is depicted in Fig. 2.

We aim to model these relations in the network; the variables that are independent of any other variable are excluded from the network, though. To solve this problem, the procedure is concluded with inducing the local network structure for these variables using the following strategy. All unselected variables are enqueued in the priority queue Q with the information gain ΔI for the outcome Y as priority value, and again the above procedure is repeated until the queue is empty. All nodes that remain as free nodes in the graph after these iterations are excluded from the network.

3.2. Representing patient dropout in the network

To correctly capture the phenomenon of patient dropout in a PBN, patient dropout in the different strata must be separated in our representation. We therefore add the variables $Y_1, \ldots, Y_T$ to the network. For each $t = 1, \ldots, T$, Y_t represents the event that the patient drops out of the process in stratum t. Furthermore, we define the global outcome variable Y in terms of them:

$$Y = \begin{cases} \mathtt{T}, & \text{if } Y_1 = \mathtt{T} \text{ or} \ldots \text{or } Y_T = \mathtt{T}, \\ \mathtt{F}, & \text{otherwise.} \end{cases}$$

We will refer to the variables $Y_1, \ldots, Y_T$ as *subsidiary outcomes*, or *sub-outcomes* for short. They become the parent nodes of the global outcome Y in the network.

In this representation, simple deterministic relationships exist between the sub-outcome Y_t and each variable in higher temporal strata including the subsequent sub-outcomes. When category '$\mathtt{I}$' denotes irrelevancy of the variable in question, it formally holds that

$$P(X_i = \mathtt{I} | Y_t = \mathtt{T}) = 1, \tag{2}$$

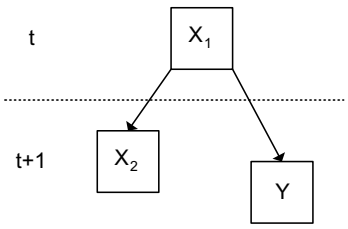

Fig. 2. Conditional independency relationship of outcome Y and variable X_2 given variable X_1, where X_1 is in stratum t and X_2 and Y in stratum $t+1$.

for each variable X_i with $s(X_i) > t$ including the sub-outcomes $Y_{t+1}, \ldots, Y_T$ independent of any other variable.

We propose to include these deterministic relationships in the representation as follows. For each $t = 1, \ldots, T-1$, an arc is added from Y_t to each variable X_i in stratum $t+1$ including the subsidiary outcome Y_{t+1}. This arc represents the deterministic relationship

$$P(X_i = \mathtt{I} | Y_t = \mathtt{T} \text{ or } Y_t = \mathtt{I}) = 1. \tag{3}$$

The deterministic relationships between Y_t and the variables in higher strata is recursively passed through the deterministic relationship between the sub-outcome Y_t and Y_{t+1}.

We propose to learn all predictive relationships from the data using the modified learning procedure that we describe below, and subsequently, to model the above-mentioned deterministic relationships in the resulting network.

3.3. Network learning with handling patient dropout

We modified the network learning procedure to learn the probabilistic relationships among variables from data while accounting for patient dropout, and included two additional steps in the procedure. The modified learning procedure assumes a temporally ordered set of strata on the predictor and subsidiary outcome variables.

The modified learning procedure starts with the final sub-outcome Y_T in the initial iteration. Data from patients who drop out prior to stratum T cannot play a role in data analyses for variables in stratum T; the variables are irrelevant for these patients. Therefore, feature subset selection and local model building for the sub-outcome Y_T and all variables in the corresponding stratum are based on a subgroup of patients that survived prior phases of care. This strategy holds for each Y_t, and the variables that are observed in the corresponding stratum. It follows that the data of all patients are used for the analyses of the first sub-outcome Y_1 and all variables that are in the corresponding stratum. In the iteration for each predictor variable, the subsidiary outcome in the corresponding stratum is excluded from the set of potential predictive features.

3.3.1. Step III

After selecting all feature subsets for the variables that appear in the priority queue for the sub-outcome Y_T and its predictive features as described in Step I and II, the procedure of feature subset selection and local model building is subsequently applied to the subsidiary outcomes $Y_1, \ldots, Y_{T-1}$, and their predictive features that have not been enqueued in prior iterations, starting with sub-outcome Y_{T-1} and concluding with the sub-outcome Y_1. This third step precedes the earlier presented Step IV of the procedure.

3.3.2. Step V

To complete the network, the deterministic relations as described in Eq. (3) are modeled in the network by adding, for each $1 \leqslant t \leqslant T$, arcs from the subsidiary outcome Y_t to

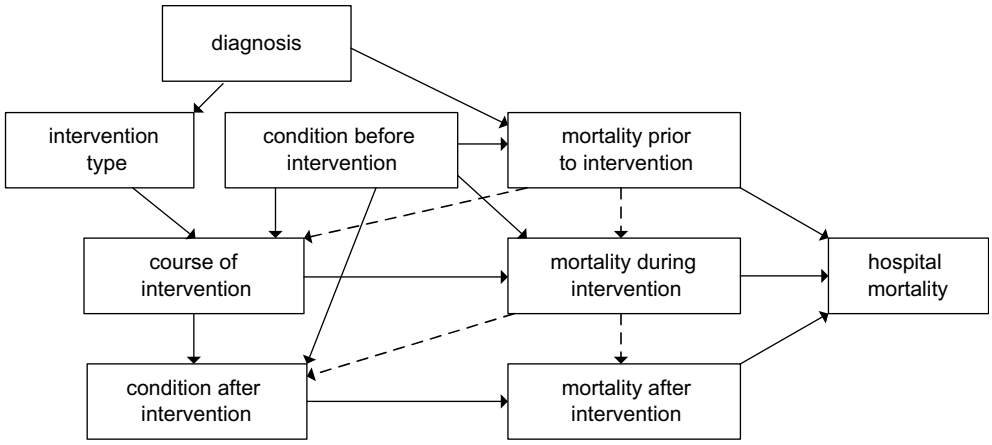

Fig. 3. Representation of patient dropout in the network structure of the prediction problem as modeled in Fig. 1b; the dotted arcs represent the deterministic relationships between each subsidiary outcome and the (sub-outcome) variables in the subsequent temporal stratum.

each variable in the temporal stratum $t+1$ including the subsidiary outcome Y_{t+1}, and extending the corresponding local conditional probability models.

Fig. 3 shows a PBN structure of the prediction problem from Fig. 1b representing patient dropout due to death in the different stages of a medical care process.

Modeling care processes in Bayesian networks involves the problem of patient dropout. In our description of representing patient dropout in a PBN through subsidiary outcomes and the modified learning procedure, we assumed a subsidiary outcome to be defined for each temporal stratum. In practice, it may not be always possible or meaningful to define a subsidiary outcome for each separate stratum. In that case, a subsidiary outcome is defined for a number of consecutive strata.

4. Clinical use of PBNs

PBNs can be applied in practice to solve a number of information problems that are related to medical prognosis.

4.1. Prognosis

The primary application of PBNs is *prognosis*, i.e., estimating the distribution of variables that represent future events. These events may pertain to conditions that occur during the process in question (*process variables*), or to endpoints of that process (*outcome variables*). The predictions can be used for decision making and resource allocation in individual cases. Furthermore, they can be used for case-mix adjustment and benchmarking in groups or populations [2]. In this case, only patient data should be included in the network, that are observed prior to the medical procedure to be evaluated. In the application of prognosis, the proposed model is thus closely related to the traditional prognostic model, although most traditional models provide limited prognostic information, as they predict a single outcome variable at a predefined prediction time.

4.2. Quick prognostic assessment

Sometimes it is not possible to collect all the information of a case at hand, while a prediction would still be useful. In an emergency setting, for instance, one may not know whether a patient is diabetic or not. Bayesian networks can perform probabilistic inference with any number of observed variables; this property allows us to make predictions with PBNs with incomplete information. As more information becomes available, the prognosis can be updated. In case of few patient data, the estimated probabilities tend to the global average of the patient population, while the estimations become more patient specific as more information is included in the model.

4.3. Prognosis updating

A patient's prognosis may change as the health care process evolves and more information becomes available. The Bayesian methodology that underlies PBNs allows us to implement a dynamic notion of prognosis, by employing probability updating based on this new information. The PBN thus provides clinicians who are involved in later phases of the process with predictions that are adjusted for the course of the preceding phases, for instance a complicated surgical intervention. In addition to the adjusted risk estimations, the change in estimated probabilities with earlier prediction times, for instance quantified in terms of risk ratios, contains important information about risk progress.

4.4. Prognostic scenario analysis

Instead of considering the prognoses for future events (e.g., complications and outcome) separately, it is often more natural to take their connection into account and consider *prognostic scenarios* of related events that are about to take place. For instance, a patient may face the prospect of severe complications and prolonged hospital-

ization when difficulties arise during surgery, or mild complications and a short hospital stay otherwise. Because of the statistical dependencies between prognostic variables, such scenarios cannot be assessed by determining the most likely values for each of the prognostic variables separately. Instead, the k most probable configurations from the Cartesian product of all possible values of these variables must be determined. Several algorithms have been developed for performing this type of probabilistic inference with Bayesian networks [35,36]. This inference with PBNs can be used to assess the k most likely clinical scenarios for a given patient or patient group.

4.5. What-if scenario analysis

The occurrence of clinical events during a health care process (e.g., a particular complication) generally changes the expectations for future parts of the process. Combining the types of probabilistic inference of Bayesian networks that we employed in the previous use cases allows us to analyze clinical *what-if scenarios* for a given patient or patient group, and to identify critical events to account for in decision making and treatment. In a what-if scenario analysis, the user is asked to specify a future event (i.e., variable-value pair) to focus on in the simulation. The PBN subsequently supplies the risk profile and the most likely scenarios that are related to the occurrence of this event. This use case illustrates the operation of the PBN as a simulation tool.

4.6. Risk factor analysis

The occurrence of unfavorable events (e.g., (post-)operative complications) and negative outcomes induces clinical questions concerning the variables that are important predictors of these events, in which stage the predictors are observed, and whether they can be influenced by the clinical staff. Risk factor analysis takes the event of interest as starting point, simulates the preceding variables for the occurrence and nonoccurrence of the event, and quantifies the predictive value of the variables for the event in terms of risk ratios. The ratio has the following form in this analysis:

$$RR(X') = \frac{P(X' = x' | X = x, \xi)}{P(X' = x' | X \neq x, \xi)}, \quad (4)$$

where X' is a process variable that precedes the event under consideration X (e.g., mortality) and ξ is given background knowledge of the patient (group) under consideration; a high value for this risk ratio indicates X' as an important risk factor for the event X in the patient group that is considered.

The six use cases illustrate various prognostic tasks for which PBNs can be applied. These tasks can be accomplished by performing 'conventional' probabilistic queries on the PBN, but they generally require that multiple queries be performed and the results be aggregated. To support the use of PBNs in medical practice, we propose the PBN to be embedded in a three-tiered architecture in which the PBN as domain layer is supplemented with a task layer, that holds a number of procedures to perform the prognostic tasks of PBNs, and a user interface as presentation layer.

5. Discussion and conclusions

This article presents the PBN as a new type of prognostic model that builds on the Bayesian network methodology and introduces a dedicated procedure for PBN learning from local supervised learning models. The health care processes that are modeled in PBNs are composed of a sequence of substantially different stages, during which patients may drop out of the process. The learning procedure explicitly accounts for the PBN's primary task, prediction, and of characteristics of the medical process being modeled in the network, including the phenomenon of patient dropout.

One way to consider the task of learning a Bayesian network structure is that we must assess an appropriate Markov blanket for each variable. The proposed learning procedure is based on the notion that assessing such a Markov blanket of a variable corresponds to selecting the best predictive feature subset for this variable in the data. For the tasks of feature subset selection and model building, any supervised learning algorithm that meets the following requirements can be plugged in. First, assuming that all network variables are discrete, the algorithm should be able to handle class variables with more than two outcome categories. Furthermore, the algorithm should provide estimated conditional class probability distributions. In addition, effective feature selection should be performed to avoid dense networks. The methodology for building classification and regression trees [37], for instance, meets these requirements; moreover, it has been shown empirically that tree methods are well able to identify Markov blankets from data [38].

The local models are used to represent the conditional probability distribution of each variable given its parents in the network. When using local models, the number of parameters that are required to encode the conditional probability distribution is lower than in a tabular representation, which results in more robust estimations of the distributions. In the work of Friedman and Goldszmidt [39] and Chickering et al. [40], tree models and a generalization thereof, decision graphs, were earlier proposed for compact representation of the local conditional probability distributions, and it was shown how such representations can be exploited by K2-type methods [27] for learning Bayesian networks from data. In contrast to our learning procedure, the local models are employed to reduce the variance in the scoring function as used in the K2-type methods.

In the above-mentioned studies [39,40], the method of global search to maximize the likelihood remains intact. In our learning procedure, however, the network is induced

from data by a local search strategy. As the main task of PBNs is outcome prediction, this local search strategy starts with the outcome variable of the process being modeled, and assigns a special role to this variable throughout network learning. The search as performed in our procedure is therefore biased, and does not necessarily maximize the global likelihood. In this search strategy, we deployed a supervised learning method to build a predictive model for each network variable; the models are subsequently combined to obtain the global network. The use of the supervised learning method is therefore 2-fold in our procedure: (a) for compact representation of the conditional probability distributions, and (b) for inducing local predictive models from data.

The learning procedure assumes a temporally ordered set of attribute strata defined by the time and order that they are observed, with the outcome variable in the highest stratum. The outcome variable is used in the initial step of the procedure, and the temporal strata are used to achieve that the direction of arcs in the resulting network represents the flow of time. Nevertheless, the procedure can be applied if just an outcome variable is available, but no ordering on the predictor variables exists. Absence of such an ordering, however, entails increasing the variance in the structure of the resulting networks, as the strata impose limitations on the possible topologies of the network and is therefore a benefit when learning from data. If no outcome variable is available, there is a variant of the learning procedure conceivable in which a feature subset is selected for each network variable, whereupon the collection of feature subsets is transformed into a graphical representation. Which strategy is suitable to be used for this latter step is still an open question and an interesting subject for further investigation.

The phenomenon of patient dropout is represented by subsidiary outcome variables in the network. Patient dropout due to the occurrence of the outcome event of the PBN including the occurrence of more serious variants of this event can be modeled in this representation. Examples of these events are the occurrence of complications and death. Patients may also drop out of a care process due to reasons that are independent of the outcome event, e.g., they may change hospitals. The current representation of patient dropout is not sufficient to represent this type of patient dropout in the network, and extension of the representation of patient dropout is an important topic for future work.

With employing Bayesian networks for prognostic purposes in this article, we did not intend to exploit the entire potential of this methodology. This includes for instance our assumption of all continuous variables to be discretized prior to network learning. In the literature on Bayesian networks, strategies have been presented for variable discretization during network learning [41], as well as for inclusion of continuous variables by estimating a parametric distribution [42]. Another interesting subject that could be exploit for PBNs is network learning with hidden variables [43].

This article also provides an explicit description of prognostic tasks that can be supported with PBNs. The six use cases were defined within the domain of cardiac surgery together with three clinical experts (PR, EdJ, BdM). In our view, these use cases are relevant in many medical procedures. The set of use cases may be incomplete, though, as some additional functionality could be defined when the proposed type of model is applied to other clinical domains.

One may argue that the tasks that we defined for PBNs could be fulfilled by a collection of traditional models that have been developed for different future (outcome) variables and different prediction times and sets of covariates. Such a collection could then be used for (quick) prognostic assessment and prognostic updating. However, the number of traditional models that is needed to equal the flexibility of a Bayesian network in performing these tasks is exponential in the number of covariates. For a single outcome variable, there exist $2^n - 1$ different nonempty sets of n covariates. This means that an equal number of different models would be needed to predict and update one outcome variable with equal flexibility as a PBN. Moreover, the tasks of prognostic scenario analysis and what-if scenario analysis (use cases 4 and 5) can not be performed by a collection of traditional models.

We presented the simulation of what-if scenarios as a functionality of PBNs. It is worth to note that in this analysis, the simulation of the causal effect of an event or its underlying clinical decision on the further course of the process is biased when observational data are used for network learning, instead of data from randomized controlled studies. In general, the analysis of causal effects is complicated due to the problem of *counterfactuals* [44]. That is, for each patient in which an event occurred, the outcome is unknown that would have been observed if the event did not occur, as well as the outcome that would have followed the occurrence of an event in patients in which the event did not occur. Randomized controlled studies enable researchers to compute unbiased estimates of causal effects, as these studies ensure exchangeability of patient groups [45]. In observational studies, however, the analysis is biased due to the lack of this exchangeability. Simulation of what-if scenarios using networks based on observational data can therefore only be used for an exploratory comparison of the differences between two clinical courses, and not for simulation of the effect of an event or its underlying clinical decision. Modeling of counterfactuals in graphical models has been described in [46].

In conclusion, this article introduces PBNs as a new type of prognostic model that builds on the Bayesian network methodology. It presents a dedicated procedure for PBN learning from local tree models. The procedure accounts for the prognostic task of PBNs, and for characteristics of the medical process being modeled in the network, including the phenomenon of patient dropout. Furthermore, a number of clinical uses of PBNs are explicitly described. As such, we adapted the Bayesian network for

prognostic application to support the clinical use of it. The PBN extends the functionality of the traditional prognostic model.

Acknowledgments

The authors thank Ameen Abu-Hanna and Frans Voorbraak for their valuable comments in the early stage of the study. Niels Peek receives a grant from the Netherlands Organization of Scientific Research (NWO) under project No. 634.000.020.

References

[1] Wyatt J, Altman DG. Prognostic models: clinically useful or quickly forgotten? BMJ 1995;311:1539–41.
[2] Abu-Hanna A, Lucas PJF. Prognostic models in medicine. Methods Inf Med 2001;40:1–5.
[3] Pearl J. Probabilistic reasoning in intelligent systems: networks of plausible inference. San Mateo, CA: Morgan Kaufmann; 1988.
[4] Andreassen S, Woldbye M, Falck B, Andersen SK. MUNIN—A causal probabilistic network for interpretation of electromyographic findings. In: Proceedings of the tenth international joint conference on artificial intelligence; 1987. p. 366–372.
[5] Shwe MA, Middleton B, Heckerman DE, Henrion M, Horwitz EJ, Lehmann HP, et al. Probabilistic diagnosis using reformulation of the INTERNIST-1/QMR knowledge base. The probabilistic model and inference algorithm. Methods Inf Med 1991;30:241–55.
[6] Long WJ, Naimi S, Criscitiello MG. Development of a knowledge base for diagnostic reasoning in cardiology. Comput Biomed Res 1992;25:292–311.
[7] Heckerman DE, Horvitz EJ, Nathani BN. Towards normative expert systems. I. The Pathfinder project. Methods Inf Med 1992;31:90–105.
[8] Bellazzi R, Berzuini C, Quaglini S, Spiegelhalter D, Leaning M. Cytotoxic chemotherapy monitoring using stochastic simulation on graphical models. In: Stefanelli M, Hasman A, Fieschi M, editors. Proceedings of the third conference on artificial intelligence in medicine. Berlin: Springer-Verlag; 1991. p. 227–38.
[9] Lucas PJF, de Bruijn NC, Schurink K, Hoepelman IM. A probabilistic and decision—theoretic approach to the management of infectious disease at the ICU. Artif Intell Med 2000;19:251–79.
[10] van der Gaag LC, Renooij S, Witteman CL, Aleman BM, Taal BG. Probabilities for a probabilistic network: a case study in oesophageal cancer. Artif Intell Med 2002;25:123–48.
[11] Lucas PJF, van der Gaag LC, Abu-Hanna A. Bayesian networks in biomedicine and health care. Artif Intell Med 2004;30:201–14.
[12] Lucas PJF, Boot H, Taal BG. Computer-based decision support in the management of primary gastric non-Hodgkin lymphoma. Methods Inf Med 1998;37:206–19.
[13] Sierra B, Larrañaga P. Predicting survival in malignant skin melanoma using Bayesian networks automatically induced by genetic algorithms. An empirical comparison between different approaches. Artif Intell Med 1998;14:215–30.
[14] Andreassen S, Riekehr C, Kristensen B, Schønheyder HC, Leibovici L. Using probabilistic and decision–theoretic methods in treatment and prognosis modeling. Artif Intell Med 1999;15:121–34.
[15] Sakellaropoulos GC, Nikiforidis GC. Development of a Bayesian network for the prognosis of head injuries using graphical model selection techniques. Methods Inf Med 1999;38:37–42.
[16] Verduijn M, Rosseel PMJ, Peek N, de Jonge E, de Mol BAJM. Prognostic Bayesian networks II: an application in the domain of cardiac surgery. J Biomed Inform. 2007;40:619–30.
[17] Hosmer DW, Lemeshow S. Applied logistic regression. 2nd ed. New York: John Wiley & Sons; 2000.
[18] Harrell Jr FE. Regression modeling strategies. Berlin: Springer; 2001.
[19] Knaus WA, Wagner DP, Draper EA, Zimmerman JE, Berger M, Bastos PG, et al. The APACHE III prognostic system: risk prediction of hospital mortality for critically ill hospitalized adults. Chest 1991;100(6):1619–36.
[20] Le Gall J, Lemeshow S, Saulnier F. A new Simplified Acute Physiology Score (SAPS II) based on a European/North American multicenter study. J Am Med Assoc 1993;270:2957–63.
[21] Nashef SAM, Roques F, Michel P, Gauducheau E, Lemeshow S, Salomon R. European system for cardiac operative risk evaluation (EuroSCORE). Eur J Cardiothorac Surg 1999;16:9–13.
[22] Hastie T, Tibshirani R, Friedman J. The elements of statistical learning. Berlin: Springer; 2001.
[23] Beck JR, Pauker SG. The Markov process in medical prognosis. Med Decis Making 1993;3:419–58.
[24] Dagum P, Galper A. Time series prediction using belief network models. Int J Hum Comput Stud 1995;42:617–32.
[25] Kayaalp MM. Learning dynamic Bayesian network structures from data. University of Pittsburgh; 2003.
[26] Buntine WL. A guide to the literature on learning probabilistic networks from data. IEEE Trans Knowledge Data Eng 1996;8:195–210.
[27] Cooper GF, Herskovits E. A Bayesian method for the induction of probabilistic networks from data. Mach Learn 1992;9:309–47.
[28] Lam W, Bacchus F. Learning Bayesian belief networks. An approach based on the MDL principle. Comput Intell 1994;10:269–93.
[29] Heckerman DE, Geiger D, Chickering DM. Learning Bayesian networks: the combination of knowledge and statistical data. Mach Learn 1995;20:197–243.
[30] Cheng J, Greiner R, Kelly J, Bell D, Liu W. Learning Bayesian networks from data: an information-theory based approach. Artif Intell 2002:195–210.
[31] Friedman N, Geiger D, Goldszmidt M. Bayesian network classifiers. Mach Learn 1997;29:131–63.
[32] Cheng J, Greiner R. Comparing Bayesian network classifiers. In: Proceedings of the fifteenth conference on uncertainty in artificial intelligence (UAI'99). Morgan Kaufmann; 1999. p. 101–7.
[33] Grossman D, Domingos P. Learning Bayesian network classifiers by maximizing conditional likelihood. In: Proceedings of the 21st international conference on machine learning. ACM Press; 2004. p. 361–8.
[34] Mitchell TM. Machine learning. New York: McGraw-Hill; 1997.
[35] de Campos LM, Gámez JA, Moral S. Partial abductive inference in Bayesian belief networks—an evolutionary computation approach by using problem-specific genetic operators. IEEE Trans Evol Comput 2002;6:105–31.
[36] Nilsson D. An efficient algorithm for finding the *M* most probable configurationsin probabilistic expert systems. Stat Comput 1998;8:159–73.
[37] Breiman L, Friedman JH, Olshen RA, Stone CJ. Classification and regression trees. Monterey: Wadsworth & Brooks; 1984.
[38] Frey L, Fisher D, Tsamardinos I, Aliferis CF, Statnikov A. Identifying Markov blankets with decision tree induction. In: Proceedings of the third IEEE international conference on data mining; 2003. p. 59–66.
[39] Friedman N, Goldszmidt M. Learning Bayesian network with local structure. In: Proceedings of the twelfth conference on uncertainty in artificial intelligence. San Francisco, CA: Morgan Kaufmann; 1996. p. 252–62.
[40] Chickering DM, Heckerman D, Meek C. A Bayesian approach to learning Bayesian networks with local structure. In: Proceedings of the thirteenth conference on uncertainty in artificial intelligence; 1997. p. 80–89.
[41] Friedman N, Goldszmidt M. Discretizing continuous attributes while learning Bayesian networks. In: Proceedings of the thirteenth international conference on machine learning. San Francisco, CA: Morgan Kaufman; 1996. p. 157–65.

[42] Friedman N, Goldszmidt M, Lee TJ. Bayesian network classification with continuous attributes: getting the best of both discretization and parametric fitting. In: Proceedings of the fifteenth international conference on machine learning. San Francisco, CA: Morgan Kaufman; 1998. p. 179–87.
[43] Friedman N. Learning belief networks in the presence of missing values and hidden variables. In: Proceedings of the fourteenth international conference on machine learning. San Francisco, CA: Morgan Kaufman; 1997. p. 125–33.
[44] Lewis D. Counterfactuals. Oxford: Basil Blackwell; 1973.
[45] Hernán MA. A definition of causal effect for epidemiological research. J Epidemiol Community Health 2004;58:265–71.
[46] Pearl J. Causality: models, reasoning, and inference. Cambridge: Cambridge University Press; 2000.