VALUE IN HEALTH 22 (2019) 439–445

Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/jval

Methodology

Bayesian Networks for Risk Prediction Using Real-World Data: A Tool for Precision Medicine

Paul Arora, PhD [1,2,*], *Devon Boyne, MSc* [2,3], *Justin J. Slater, MSc* [2], *Alind Gupta, PhD* [2,4], *Darren R. Brenner, PhD* [5], *Marek J. Druzdzel, PhD* [6]

[1]*Division of Epidemiology, Dalla Lana School of Public Health, University of Toronto, Toronto, ON, Canada;* [2]*Lighthouse Outcomes, Toronto, ON, Canada;* [3]*University of Calgary Cumming School of Medicine, Calgary, AB, Canada;* [4]*Department of Molecular Genetics, University of Toronto, Toronto, ON, Canada;* [5]*Department of Oncology and Community Health Sciences, University of Calgary Cumming School of Medicine, Calgary, AB, Canada;* [6]*School of Computing and Information, University of Pittsburgh, Pittsburgh, PA, USA*

ABSTRACT

Objective: The fields of medicine and public health are undergoing a data revolution. An increasing availability of data has brought about a growing interest in machine-learning algorithms. Our objective is to present the reader with an introduction to a knowledge representation and machine-learning tool for risk estimation in medical science known as Bayesian networks (BNs). **Study Design:** In this article we review how BNs are compact and intuitive graphical representations of joint probability distributions (JPDs) that can be used to conduct causal reasoning and risk estimation analysis and offer several advantages over regression-based methods. We discuss how BNs represent a different approach to risk estimation in that they are graphical representations of JPDs that take the form of a network representing model random variables and the influences between them, respectively. **Methods:** We explore some of the challenges associated with traditional risk prediction methods and then describe BNs, their construction, application, and advantages in risk prediction based on examples in cancer and heart disease. **Results:** Risk modeling with BNs has advantages over regression-based approaches, and in this article we focus on three that are relevant to health outcomes research: (1) the generation of network structures in which relationships between variables can be easily communicated; (2) their ability to apply Bayes's theorem to conduct individual-level risk estimation; and (3) their easy transformation into decision models. **Conclusions:** Bayesian networks represent a powerful and flexible tool for the analysis of health economics and outcomes research data in the era of precision medicine.

Keywords: artificial intelligence, Bayesian networks, decision models, machine learning, precision medicine, real-world data, regression-based models, risk prediction, statistical methods

Background

The fields of clinical medicine and public health are undergoing a data revolution. Transformation of large volumes of medical records to an electronic format and the remarkable growth in the data collected by health registries and during clinical studies provide opportunities to make risk prediction and intervention selection more precise. This increasing availability of the so-called "big data" has brought about a growing interest in machine-learning algorithms for extracting knowledge from observations, typically conceptualized as datasets, and for constructing personalized risk prediction models. There are a wide variety of tools available for developing personalized risk prediction models, and a broader understanding of such tools may help researchers avoid common pitfalls of working with big data while improving model performance. Herein, we present the reader with an introduction to a knowledge representation and machine-learning tool for risk prediction known as Bayesian networks (BNs). We discuss some of the challenges associated with traditional risk prediction methods and then describe BNs, their construction, application, and advantages in risk prediction based on examples in cancer and heart disease. We also provide the reader with suggestions for the implementation of BNs and discuss software solutions for their application.

Methods of developing a risk prediction tool can be categorized into 2 approaches: (1) regression-based models and (2) machine-

* *Address correspondence to:* Paul Arora, PhD, Division of Epidemiology, Dalla Lana School of Public Health, University of Toronto, 155 College Street 6th floor, Toronto, ON M5T 1P8, Canada.

E-mail: paul.arora@utoronto.ca

1098-3015/$36.00 - see front matter

https://doi.org/10.1016/j.jval.2019.01.006

learning algorithms. Although regression-based models can be implemented in a Bayesian framework, historically in medicine, the tools most commonly used in clinical risk prediction analysis have been regression-based models (Cox, logistic, Poisson) applied in a frequentist framework.

Regression-based risk prediction models estimate a baseline risk, rate, or hazard and typically generate a linear combination of covariates using an algorithm that maximizes the likelihood of the outcome. In standard regression analysis, based on the frequentist approach, the analyst is presented with a table of regression coefficients, learned completely from data, corresponding to each covariate in the model. Using a linear combination of these covariates and the baseline value, an outcome (risk or rate of an event) is calculated. Two regression-based risk prediction scores commonly used in clinical medicine are the Framingham Risk Score for 10-year risk of cardiovascular disease[1] and the Kidney Failure Risk Score for 10-year risk of end-stage kidney failure.[2] Typically, in a clinical cohort setting, Cox proportional hazards regression analysis (CPHRA) is used to generate a risk equation to predict the risk of event based on the hazard function. The FRS and KFRS are able to predict outcomes in the original cohort on which these models were developed and have been validated in other settings. Advantages of regression-based approaches to risk modeling are their widespread use and flexibility in being able to model continuous, binary, count, and time outcomes. Three key disadvantages of using regression-based approaches are as follows: (1) they can model associations but not causal structure; (2) they operate under restrictive assumptions about the relationships among variables; (3) combining regression-based models is not straightforward. Furthermore, using regression-based models, each outcome of interest must be trained on its own individual model and a static set of data. Other problems with regression-based risk prediction models are reviewed elsewhere.[3]

Bayesian Networks

In comparison, BNs are compact and intuitive graphical representations of joint probability distributions (JPDs) that can be used to conduct causal reasoning and risk prediction analysis and offer several advantages over regression-based methods. Bayesian networks represent a different approach to risk prediction. They are graphical representations of JPDs that take the form of a network made up of nodes and edges representing model random variables and the influences between them, respectively. The JPD factorizes into conditional probability distributions associated with each node conditional on variables that directly influence it. The computational efficiency of a BN stems from its explicit representation of independencies, which results in a reduction of connectivity in the graph. This results in more compact networks and scales significantly better for large systems or networks because only a subset of all possible connections need be evaluated. A detailed description of BNs is outside the scope of this article and instead we focus on the important concepts, strengths, and challenges of applying BNs (Table 1). Detailed statistical information describing BNs can be found in several sources.[4–6]

BNs rely on the Bayesian approach to statistical inference. Frequentist inference assumes that the parameter of interest is an objective number representing a frequency in an infinite number of trials, whereas in the Bayesian approach the parameter reflects a measure of personal belief, that is, subjective knowledge.[7] Bayes's theorem provides a mechanism for updating knowledge when new evidence is collected. While regression models can be implemented in a Bayesian framework, BNs generate their estimates by calculations of the conditional probability distribution over a variable of interest given observations of other variables. This calculation is possible because a BN is a model of the JPD over all modeled variables, and information can spread from the observed variables to the variable of interest through the connections modeled explicitly in the graph.

The JPD represented by the BN is subjective and can be updated if new evidence is available using Bayes's theorem, allowing for incorporation of new data or expert opinion. This subjective character of knowledge is one of the advantages of Bayesian statistics for decision makers and clinicians. It allows for formal rigor in processing uncertainty in situations where little or no data are available but substantial expertise exists. For example, it is useful for patient risk stratification problems where patient treatment guidelines are not universal, are not available, or are rapidly changing (eg, peritoneal surface malignancy of colonic origin).[8] When substantial amounts of data exist, both approaches rely on data.

There are three approaches to the construction of BN structures: purely expert-elicited, purely automated or machine-learned, and a combined approach where prior expert knowledge is incorporated into the automated learning process. Expert-elicited models construct BNs that are based on expert opinion or subjective belief alone. This is very tedious, particularly when the models contain a large number of parameters, and so it is used

Table 1 – Comparison of regression approaches with BN (strengths and limitations)

	Strengths	Limitations
Regression-based models	• Widely used • Can be performed with nearly all statistical software packages	• Causal structure not modeled • Difficulty in handling nonlinear relationships • Generally, one model is needed per outcome under study
Bayesian networks	• Causal structure is explicitly represented; individual-level risk prediction • Can be learned from data, purely from expert knowledge (no data), or a combination of the two approaches • "What-if" scenarios can be modeled • BN can be extended into decision models by incorporating decision and utility nodes • Model multiple outcomes and exposures in a single model	• Not well known in health sciences • Specialized knowledge of Bayesian statistics required to understand process • Computationally expensive • Cannot model cyclical relationships such as feedback loops*

* "Dynamic" Bayesian networks provide a way of dealing with feedback loops.

primarily when no data are available. Purely machine-learned approaches include constraint- and score-based learning. Constraint-based approaches[9] use conditional independencies in the data to derive the model structure, whereas score-based systems[10] search for a model that maximizes the likelihood (or any other score) of the model given the data. The combined approach allows expert or user input to force known temporal relationships, direct relationships, and direct causal relationships to be part of a machine-learned proposed structure.

Risk prediction modeling with BNs has several advantages over regression-based approaches that address commonly encountered challenges in risk prediction, but we focus here on three: (1) they generate network structures such that the assumed underlying causal structure between variables can be visualized and readily disseminated; (2) they can be used to conduct what-if scenario analysis and individual-level risk prediction; and (3) they can be transformed into decision models in a relatively straightforward manner.

BNs are presented as directed acyclic graphs that are networks of nodes connected by edges in such a way that there is no way to trace a path starting at one node and follow a path, determined by edges, back to the same node (Fig. 1).[6] In this example, the arrows represent dependencies between the variables. A directed acyclic graph (DAG) of a JPD estimates the likelihood of an outcome *and* defines a hierarchy of conditional independence. A BN consists of two components: a qualitative component composed of the DAG structure and a quantitative component composed of a JPD that factorizes into a set of conditional probability tables governed by the structure of the DAG.[6] The qualitative structural component encodes the causal relationships among the variables; however, for models aimed only at risk prediction, these relationships do not have to be causal. The parameters encoded in the quantitative component quantify the relationships in the DAG. Unlike regression-based methods, the structure of the model, the pattern of dependencies between all variables in the model, is presented (whether learned from the data, expert opinion, or both). The transparent presentation of the role of risk predictors in the model facilitates visually appealing and intuitive communication of both the model and its results to decision makers. In Figure 1, we can see that three patient characteristics influence the likelihood of disease, and that the status of these patient characteristics can propagate through the network to influence test results for either test 1 or test 2. Using this example, we see that, unlike other machine-learning approaches such as artificial neural networks, users can clearly understand how a particular probability was determined and which variables are contributing proximally and distally to the outcomes of interest, that is, their relative position in the network.

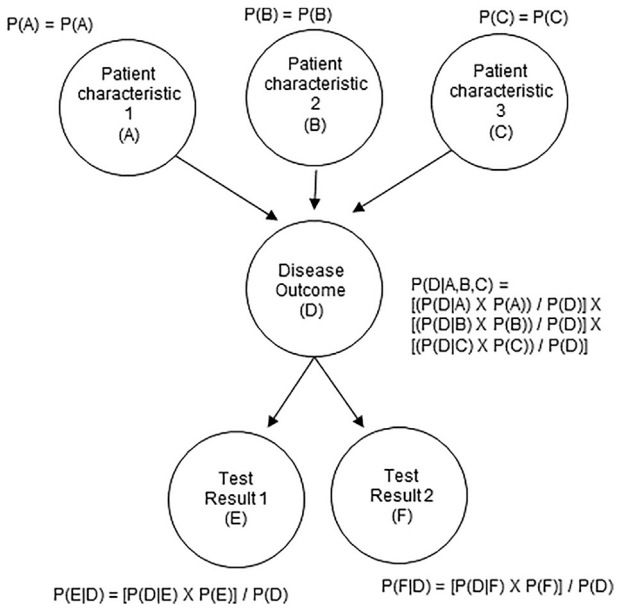

Fig. 1 – A simple annotated Bayesian network and accompanying conditional probabilities. Reasoning can be performed in either direction from effect-to-cause or cause-to-effect.

Individual Risk Prediction

BNs allow for ease of individual-level risk prediction via "what-if" analyses and effect-to-cause reasoning, also referred to as "abductive" or "diagnostic" reasoning. With respect to the "what-if" analysis, they are also possible with regression-based models, when these include causal information, that is, when the predictor variables are causally related to the predicted variables. This is, however, much more natural for Bayesian networks because influences are modeled by means of directed arcs. When each of the arcs models a causal relationship, the entire model, no matter how large and complex, will naturally predict the effects of causal manipulation, something that underlies the "what-if" analysis. The era of "precision medicine" will require the incorporation of more factors into risk prediction models as treatment options become more specialized, demanding individual risk prediction.[11] Individual risk prediction with regression-based approaches involves applying a baseline risk (cumulative hazard in the case of Cox models) and combining the coefficients of model covariates. This may not be an easy task if some values are missing or if a variable is continuous with a reference state equal to a mean value that can be difficult to interpret.[12] The ability to conduct "what-if" analysis is a result of the Bayesian nature of BNs. BNs use Bayes's theorem to perform inference not only from cause to effect, as in standard predictive models (eg, inferring from association between a disease and a symptom), but also from effect to cause (eg, calculating the probability of having disease after having observed a symptom). Reasoning from effect to cause is a special capability of BNs that facilitates diagnostic applications and improves decision-making support.[13] Prior probabilities are updated to posterior probabilities after having observed evidence for any number of variables in a network. Further, each unobserved variable can be ranked, in terms of sensitivity (or impact), to the specified target variable as part of a sensitivity analysis.[13] For example, in cancer care, one may wish to ask what the survival probability is, conditional on observed evidence thus far, for different treatment plans (which are unobserved). This is a hypothetical question that cannot generally be answered with discriminative machine-learning regression models that do not include causal information.[14]

For illustrative purposes, consider the fictitious BN in Figure 2A, originally presented by Lauritzen and Spiegelhalter.[15] This network represents the conditional dependence between characteristics relating to lung cancer and tuberculosis. With no prior information, we see that a patient's chance of having cancer or tuberculosis is 6%. Note that an observation of a single x-ray result or dyspnea does not distinguish between tuberculosis and lung cancer. Nevertheless, once we have examined the patient and have determined that the patient is a smoker, is experiencing dyspnea, and has recently visited Asia, this probability increases to 21% (Fig. 2B). We can also use this network to answer hypothetical questions such as "How would this probability have changed if the patient had never visited Asia?" We can input this new information into the network, updating the probability of tuberculosis to 16% (Fig. 2C). In each scenario, the unknown variables' distributions will be updated

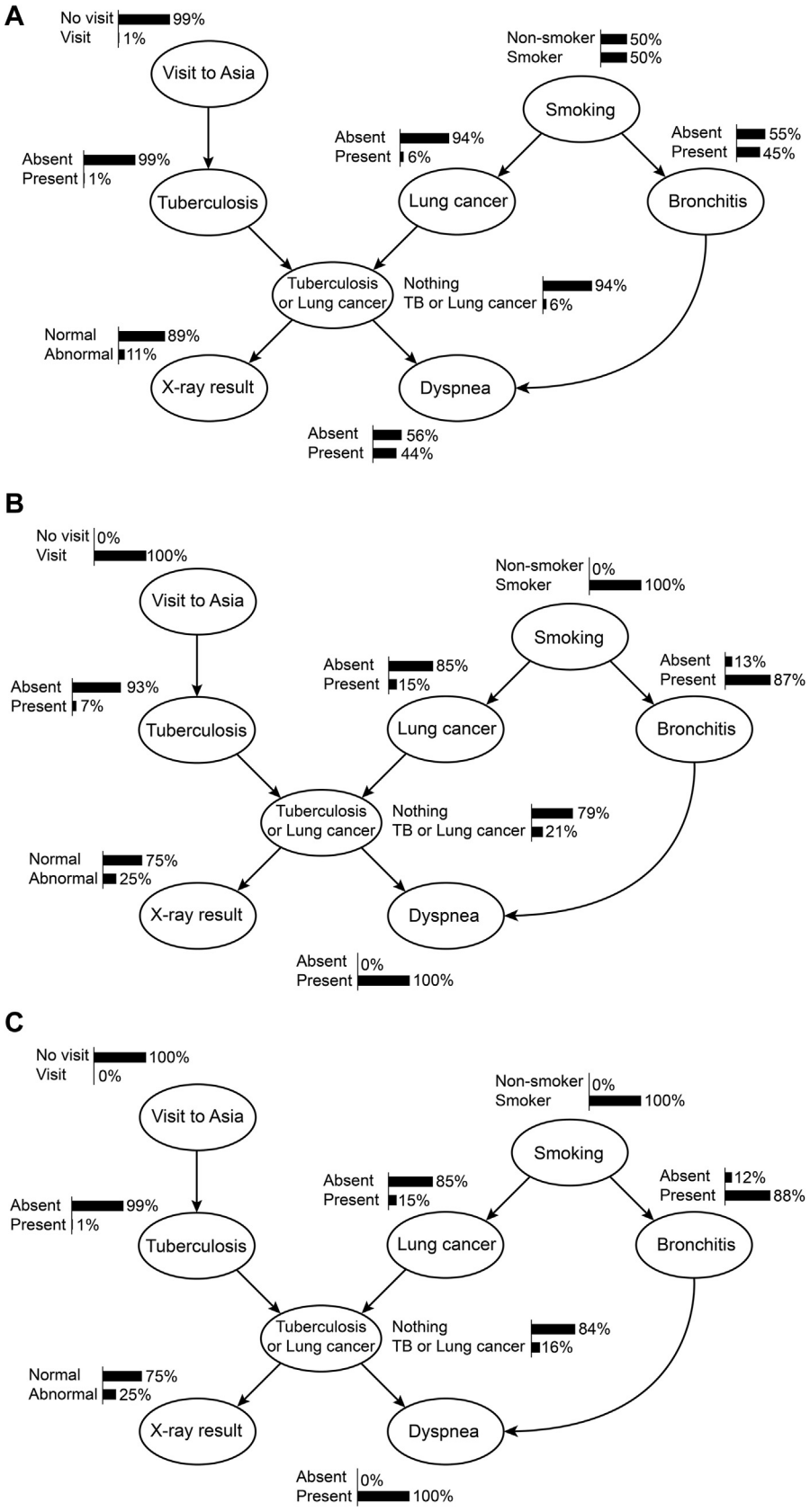

Fig. 2 – (A) A fictitious Bayesian Network for the prediction of tuberculosis or cancer. Modified from Lauritzen and Spiegelhalter.[15] (B) Example of individual risk prediction using BNs. Patient X presents as a smoker, experiencing dyspnea, and has recently visited Asia. Patient X's risk of tuberculosis or lung cancer is 21%. (C) Example of a "what-if" analysis using BNs. Suppose that we wanted to know what patient X's risk would have been if he or she had never visited Asia. We can change the "Visit to Asia" node and see that the risk of tuberculosis or lung cancer would have been 16%.

based on the known information. This demonstrates both the power and simplicity of BNs for risk prediction.

Decision Making Under Uncertainty

A BN can be enhanced, through the inclusion of decision and utility nodes, to a model for decision analysis, referred to as an "influence diagram." An example of such an enhancement is provided in Figure 3 with other examples in the literature.[16–18] Decision analysis is a modeling approach that compares decision alternatives and recommends the choice that maximizes the expected utility of the outcome to the decision maker.[17] Decision trees are the most commonly used tool to model decision problems; however, their complexity grows exponentially with problem size; thus, they are tractable only when they contain fewer than a handful of variables.[19] An influence diagram can represent the same problems as a decision tree, but its size grows only linearly with the problem size, allowing for modeling of more complex decision problems. This is important as clinicians and decision makers must assess the best treatments most likely to be effective for a patient, while also considering the trade-offs between possible benefits of therapy and potential loss of quality of life.[17] Maximizing the benefit to a patient is a difficult objective because of the growing number of interrelated factors associated with optimal utility. A clinical decision support system (CDSS) is software built around an influence diagram, designed to assist healthcare professionals with decision-making tasks. Influence diagrams containing utility nodes can be used to conduct cost-effectiveness or cost-utility analysis.[19] Value of information (VOI) analysis can be incorporated into BNs to streamline data collection and inform both the diagnosis and prognosis in an individual patient case, as well as future research priorities. Value of information analysis works by estimating the effect of observing a variable more precisely (reducing its uncertainty to zero) on some target variable of interest, on making a specific decision, or on cost (which can also be incorporated into BNs as utility nodes).[6]

Other Features

While we describe key strengths of the BN approach above, BNs have other advantages to standard regression-based techniques for risk prediction. Regression-based models for risk prediction must be re-estimated when new data are added and the risk scores recalculated. Further, each outcome must be evaluated with its own model. With BNs, multiple outcomes can be predicted from a single model, and therefore multiple questions can be answered from one model. BNs can be continuously updated with each new piece of data added. For example, if a BN is built using health system data, newly added data (each new patient record, for example) can continuously update the model parameters without the model having to be re-run or retrained. The implications of this are important. If trained and embedded in large administrative health databases, a BN risk prediction model can be up-to-date, locally relevant, and flexible to changes in the population, treatment, and testing over time. Further, a BN can improve over time as more cases are presented to it.

Problems in clinical medicine are often concerned with the development of a binary health state over time. These are typically modeled in linear regression–based approaches with time-varying covariates. "Dynamic" BNs have been applied to handle time-varying risk models by incorporating time steps and recursive models, where the state of the model in the previous time step informs the current state of the model and so on.[12,20] Furthermore, BNs are particularly adept at handling large sets of interrelated variables. In an article by Loghmanpour et al, BNs were created to predict mortality at five time points among patients receiving left ventricular assist devices.[21] The networks were constructed using three different machine-learning algorithms and evaluated the

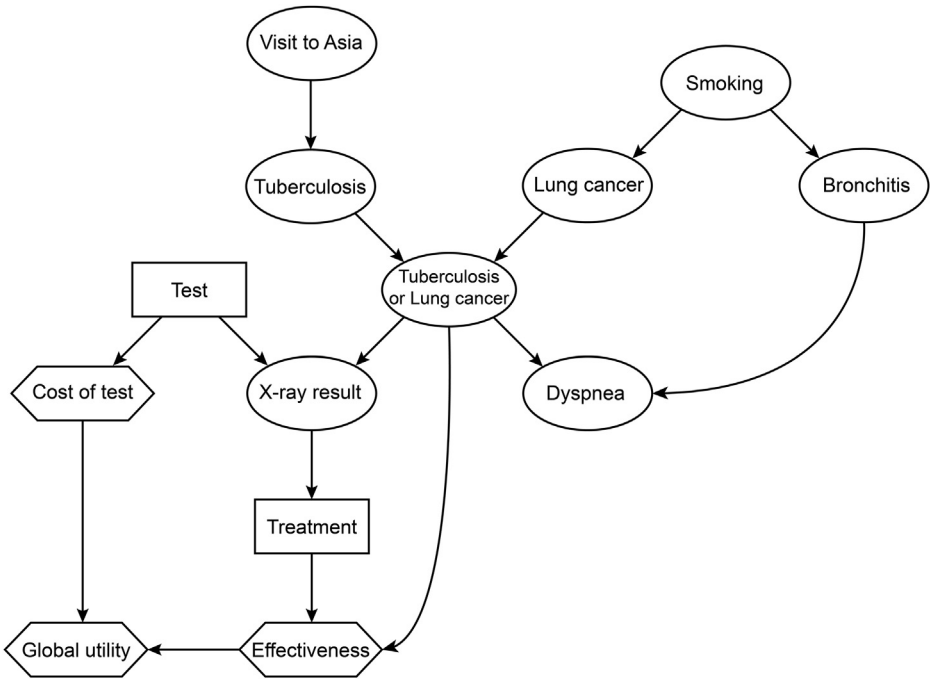

Fig. 3 – Translation of the Bayesian network from Figure 2 to a simple utility diagram for the purpose of illustration. Ovals represent chance nodes, rectangles represent decision nodes, and hexagons represent utility nodes. (Adapted from http://www.openmarkov.org/docs/tutorial/tutorial.html.)

dependencies between 226 preimplant variables recorded on 8050 subjects. The BN built by Loghmanpour et al modeled the risk of failure of left ventricular assist devices. In this study, their BN achieved accuracy of 90% and 83% 90-day and 1-year mortality, respectively, compared to 57% and 60% accuracy using the standard risk prediction tool (HeartMate II Risk Score). Other examples of BNs used to analyze large health datasets can be found in the recent literature.[8,14,22–24]

BNs are useful in the sense of being able to handle both linear and nonlinear relationships among variables, along with synergies between groups of variables and complex joint influence in general. This flexibility includes nonlinear relationships between exposures and the outcome and between the exposures themselves. Regression-based methods also allow for the incorporation of nonlinear relationships between exposures and outcomes, but this is limited to a single outcome per model and generally involves transformations of outcome variables or the introduction of polynomials, splines, or locally weighted smoothing, which may be difficult to communicate and interpret. Furthermore, while many researchers end up discretizing all variables for the sake of computational efficiency, this is not necessary; powerful software exists that does not require discretization and allows the capture of any relationship, including equations and continuous distributions.

Missing Data

Missing data can influence BNs into two broad ways; first, it can impact inference; second, it can impact learning of the causal structure. Missing data in real-world databases pose an important challenge for inference, and unfortunately it is very common. Fortunately, BNs handle missing data for inference well because reasoning with a BN does not require complete information on a patient to calculate risk for a response variable. The posterior probability distribution is calculated only based on the available covariates. This is different from linear regression-based models, where every risk factor is assumed to be known and either present or absent. Missing data can have a more deleterious effect with respect to learning the causal structure from data; however, here expert belief and published literature can be used to estimate the correct structure.

In the presence of data that are missing completely at random, or the missingness is dependent on some other variables, parameters for the BN can still be calculated.[25] This is typically done using the two-step algorithm known as the Expected Maximization (EM) algorithm. The first step imputes an expected value for the missing data points; the second step then uses the imputed dataset to calculate the maximum likelihood estimates of the model parameters. This process is then iterated until the algorithm converges. This method has been extended to structure learning, which has an added model selection step. This allows for the learning of model structure and parameters in the presence of missing data.[26] Although BNs possess several advantages over regression-based methods for risk prediction, they may not always be more accurate.[27] In general, although BNs have very good risk prediction performance, their main advantages over regression-based methods are their intuitiveness and elegance, their ability to represent the JPD as a network structure, their ability to handle what-if scenarios, their efficiency in dealing with missingness, and their ability to incorporate decision and utility nodes.

Future Directions

Why are BNs not more commonly used in medicine and health studies? One reason is that their methods arose from the fields of computer science and that most researchers in the medical sciences are trained in regression-based approaches. Second, and perhaps related to the previous point, guidelines on the application of Bayesian statistics are only recently becoming available in medicine and allied health fields.[28] The concepts of machine-learning and artificial intelligence are relatively new to researchers in the health space, and there may be some skepticism about their practical application. Nevertheless, there is a growing body of literature demonstrating the value of these approaches to problems in health and medicine. Third, BNs are computationally more expensive than classical regression approaches. Learning a BN from data can involve calculating hundreds of thousands of parameters, which has only become feasible in recent years. Finally, regression-based approaches generate concrete risk equations that a physician or decision maker can easily grasp and apply. BNs may face a knowledge translation challenge in that the network must be moved from software to a tool (web-based or otherwise) that a decision maker can use and manipulate. Several software options exist for this and BNs, and influence diagrams can be exported to graphical user interfaces for this purpose. BNs are flexible with respect to the types of data that can be used to build them. They can incorporate individual-level or aggregate data, expert opinion, and evidence synthesized from literature. Several software packages exist for reasoning with BNs (GeNIe, OpenMarkov, BayesiaLab, Hugin), including several free and open source packages for academic use and routines in SAS (HPBNET procedure) and R (gRain and bnlearn packages).

BNs provide a robust and flexible analytic approach to the challenge of complex health datasets. These complex health datasets pose specific analytic challenges because of missing data, large size, and complexity (of relationships not only between variables but also in the datasets themselves), changing populations, and nonlinear relationships between exposures and outcomes. BNs can continually be updated with new information and generate individual-level risk prediction that is up-to-date and locally relevant. The coming era of precision medicine will require novel approaches to risk prediction and decision analysis while maintaining a high degree of flexibility to accommodate developments in knowledge, new interventions, and database size and complexity. BN approaches facilitate this and should be explored further.

Acknowledgments

We thank Agnieszka Onisko for valuable discussions on Bayesian networks. We also thank Neha Sati for her assistance in preparing the article files and during the submission process.

REFERENCES

1. Wilson PW, D'Agostino RB, Levy D, et al. Prediction of coronary heart disease using risk factor categories. *Circulation*. 1998;97:1837–1847.
2. Tangri N, Stevens LA, Griffith J, et al. A predictive model for progression of chronic kidney disease to kidney failure. *JAMA*. 2011;305:1553–1559.
3. Collins GS, Omar O, Shanyinde M, et al. A systematic review finds prediction models for chronic kidney disease were poorly reported and often developed using inappropriate methods. *J Clin Epidemiol*. 2013;66:268–277.
4. Pearl J. *Bayesian Networks*. UCLA: Department of Statistics, UCLA; 2011. https://escholarship.org/uc/item/53n4f34m.
5. Pearl J. *Causality*. Cambridge: Cambridge University Press; 2009.
6. Kjærulff UBM, Anders L. *Bayesian Networks and Influence Diagrams: A Guide to Construction and Analysis*. Berlin: Springer Science & Business Media; 2012.
7. Martinez EZ, Achcar JA. Trends in epidemiology in the 21st century: time to adopt Bayesian methods. *Cad Saude Publica*. 2014;30:703–714.

8. Stojadinovic A, Nissan A, Eberhardt J, et al. Development of a Bayesian Belief Network Model for personalized prognostic risk assessment in colon carcinomatosis. *Am Surg*. 2011;77:221–230.
9. Spirtes P, Glymour CN, Scheines R. *Causation, Prediction, and Search*. 2nd ed. Cambridge, MA: MIT Press; 2000.
10. Cooper GF, Herskovits E. A Bayesian method for the induction of probabilistic networks from data. *Machine Learning*. 1992;9:309–347.
11. Collins FS, Varmus H. A new initiative on precision medicine. *N Engl J Med*. 2015;372:793–795.
12. Onisko A, Druzdzel MJ, Austin RM. How to interpret the results of medical time series data analysis: classical statistical approaches versus dynamic Bayesian network modeling. *J Pathol Inform*. 2016;7:50.
13. Coid JW, Ullrich S, Kallis C, et al. Improving risk management for violence in mental health services: a multimethods approach. *Prog Grants Appl Res*. 2016;4.
14. Sesen MB, Nicholson AE, Banares-Alcantara R, et al. Bayesian networks for clinical decision support in lung cancer care. *PLoS One*. 2013;8:e82349.
15. Lauritzen SL, Spiegelhalter DJ. Local computations with probabilities on graphical structures and their application to expert systems. *J R Stat Soc Ser B*. 1988;50:157–224.
16. Luque M, Díez FJ, Disdier C. Optimal sequence of tests for the mediastinal staging of non-small cell lung cancer. *BMC Med Inform Decis Making*. 2016;16:9.
17. Neapolitan R, Jiang X, Ladner DP, et al. A primer on Bayesian decision analysis with an application to a kidney transplant decision. *Transplantation*. 2016;100:489–496.
18. Seixas FL, Zadrozny B, Laks J, et al. A Bayesian network decision model for supporting the diagnosis of dementia, Alzheimer's disease and mild cognitive impairment. *Comput Biol Med*. 2014;51:140–158.
19. Arias M, Diez FJ. Cost-effectiveness analysis with influence diagrams. *Methods Inf Med*. 2015;54:353–358.
20. Austin RM, Onisko A, Druzdzel MJ. The Pittsburgh Cervical Cancer Screening Model: a risk assessment tool. *Arch Pathol Lab Med*. 2010;134:744–750.
21. Loghmanpour NA, Kanwar MK, Druzdzel MJ, et al. A new Bayesian network-based risk stratification model for prediction of short-term and long-term LVAD mortality. *ASAIO J*. 2015;61:313–323.
22. Fuster-Parra P, Tauler P, Bennasar-Veny M, et al. Bayesian network modeling: a case study of an epidemiologic system analysis of cardiovascular risk. *Comput Methods Prog Biomed*. 2016;126:128–142.
23. Kraisangka J, Druzdzel MJ, Benza RL. A risk calculator for the pulmonary arterial hypertension based on a Bayesian network. Proceedings of the 13th UAI Bayesian Modeling Applications Workshop (BMAW 2016) co-located with the 32nd Conference on Uncertainty in Artificial Intelligence (UAI 2016) New York City, NY, USA. June 25, 2016.
24. Multani P, Niemann U, Cypko MA, et al. Building a Bayesian network to understand the interplay of variables in an epidemiological population-based study, 2018 IEEE 31st International Symposium on Computer-Based Medical Systems (CBMS). Piscataway, NJ: IEEE; 2018.
25. Uusitalo L. Advantages and challenges of Bayesian networks in environmental modelling. *Ecol Model*. 2007;203:312–318.
26. Riggelsen C, Feelders A. Learning Bayesian network models from incomplete data using importance sampling. AISTATS, Citeseer, 2005.
27. Gao P, Zhou X, Wang ZN, et al. Which is a more accurate predictor in colorectal survival analysis? Nine data mining algorithms vs. the TNM staging system. *PLoS One*. 2012;7:e42015.
28. Center for Biologics Evaluation and Research. *Guidance for the Use of Bayesian Statistics in Medical Device Clinical Trials: Guidance for Industry and FDA Staff*. Rockville, MD: Food and Drug Administration; 2010.