

User-centered Development of a Clinical Decision Support System

Anna Kleinau¹, Alex Mo³, Juliane Müller-Sielaff¹, Johanna M. A. Pijnenborg⁴, Peter J. F. Lucas³ and Steffen Oeltze-Jafra^{1,2}

¹Department of Neurology, Otto von Guericke University Magdeburg, Germany

²Center for Behavioral Brain Sciences (CBBS), Magdeburg, Germany

³DMB Department, University of Twente, Enschede, The Netherlands

⁴Department of Obstetrics and Gynaecology, Radboud University Medical Center, Nijmegen, The Netherlands

Abstract

Scientific progress is offering increasingly better ways to tailor a patient's treatment to the patient's needs, i.e., better support for optimal clinical decision-making can be offered. Choosing the appropriate treatment for a patient depends on numerous factors, including pathology results, tumor stage, genetic, and molecular characteristics. Bayesian networks are a type of probabilistic artificial intelligence, which in principle would be suitable to support complex clinical decision-making. However, most clinicians do not have experience with these networks. This paper describes an approach of developing a clinical decision support system based on Bayesian networks, that does not require insight knowledge about the underlying computational model for its use. It is developed as a therapy-oriented approach with a focus on usability and explainability. The approach features the computation and presentation of individualized treatment recommendations, comparison of treatments and patient cases, as well as explanations and visualizations providing additional information on the current patient case.

Keywords


Bayesian Networks, Clinical Decision Support Systems, Human Computer Interaction

1. Introduction

The last few decades, clinical management of disease in patients has become increasingly complicated. There are more diagnostic tests and more treatment options available than ever before and knowledge about particular disorders has further deepened. In the case of *personalized* treatment, the specific nature of the disease, the patient (using increasingly often disease- and patient-specific genetic markers), and the patient's environment are taken into account. The aim is to offer *optimal* disease management for the individual patient in the face of recent scientific evidence. Often this clinical trend is referred to as *personalized medicine* [1].

In the past, proper clinical management of a patient's disease was mainly the responsibility of the individual medical specialist, who made decisions based on clinical experience and expertise. In the modern era, scientifically trusted clinical knowledge on the management of specific diseases is gathered by organizations of clinical specialists. These organizations develop


AIxIA 2021 SMARTERCARE Workshop, November 29, 2021, Milan, IT

 0000-0002-3415-6316 (A. Kleinau); 0000-0002-8279-0901 (J. Müller-Sielaff); 0000-0002-6138-1236

(J. M. A. Pijnenborg)



© 2021 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

 CEUR Workshop Proceedings (CEUR-WS.org)

clinical guidelines, textual descriptions of the optimal management of specific disorders, that support the selection of appropriate diagnostic or treatment actions, based on signs, symptoms, and laboratory test results. Through the inclusion of prediction equations among much text, guidelines may offer some flexibility to personalize advice. However, most of a guideline is just extensive, non-interactive text.

An alternative to clinical guidelines are *clinical decision support systems* (CDSS). These computer-based systems are interactive and support clinical decision-making based on evidence regarding an individual patient, employing one or more prediction models that capture knowledge of one or more disorders. In particular, Bayesian networks are considered to be promising formalisms to capture clinical knowledge, as they can represent the associated uncertainty readily, and support the representation of conditional and causal knowledge in a natural fashion. Whereas Bayesian networks appear to offer potential as clinical decision models, it is still necessary to design and develop the functionality of the software used to consult a given network to help in decision-making. Although software for computing probabilities from Bayesian networks is available (e.g. Hugin [2] and Ergo [3]), the functionality of such software is not geared to the needs of a clinician who typically wishes to use a Bayesian network model in the context of a clinical consultation without possessing a full understanding of the inner workings of the computational model.

The present paper discusses the requirements, design, and a validation study of a CDSS based on Bayesian network models. CDSS can take many different forms and functions. Among others, they can support the clinician in finding the right diagnosis or treatment of a patient's disorder; providing intelligent alarm (e.g. avoid drug adverse effects) is an application that would help in improving patient safety [4]. The actual work presented in this paper focuses on a single disorder, *endometrial cancer*, and an associated Bayesian network, the ENDORISK Bayesian network, that was developed recently [5]. All the work was done in close collaboration with clinicians, in particular with gynaecologist-oncologists with specialized knowledge of endometrial cancer.

2. Related Research

2.1. Current Software Solutions

To allow computational models to be used in clinical practice, they need to be integrated into daily work. Usually this means some kind of integration with the electronic patient record system (EPRS), but as there are many commercial EPRS on the market, rendering this is a major undertaking. It is therefore attractive to focus on web-based software solutions, as these are nowadays easily accessible from any computer operating system. An existing example is the web-based software tool *Evidencio*¹, which offers centralized access to computational prediction equations for particular disorders. Its aim is to bridge the gap between the scientific literature, where these equations are described, and their use by providing access to working models.

Software specifically constructed for Bayesian networks as the computational model is often all-in-one software that supports both the creation of Bayesian networks as well as inference.

¹<https://www.evidencio.com>

Software such as Hugin [2], Netica², Ergo [3], and GeNIe [6] provide functionality to construct Bayesian networks, as well as inference methods and a user interface to consult the network. However, they are mostly developed domain-agnostic and do not have features specific for clinical decision support.

CDSS approaches specifically constructed for Bayesian networks as knowledge bases are sparse. A broad approach is chosen by Zagorecki et al.³ with their symptom checker app *Symptomate*⁴. Through a simple interface, the user can anonymously enter various symptoms and get a diagnosis as a result. The system works with a broad range of diseases. It is supported by a large Bayesian network and does not assume much medical knowledge of the user.

Müller et al. [7] proposed a CDSS approach that was an important inspiration for the here presented approach. Their work highlights the importance of transparent, explainable recommendations in CDSS and presents an approach to score the relevance of evidence items for a recommendation. Additionally, an interactive software based on visualizations of these relevance scoring is presented. The software was developed in the context of laryngeal cancer therapy but can work with different Bayesian networks. The *ENDORISK* network was used to demonstrate the generalizability of the approach.

2.2. Explainability of Bayesian Networks

Explaining the reasoning behind inferences from Bayesian networks is a complex task but crucial to generate trust. Bayesian networks work with probabilities and conditional probabilities. The nodes in a network are structured causally, but influence each other in all kinds of directions.

Yap et al. [8] developed a text-based approach that explains a node using the most important nodes in its Markov blanket. If some of those nodes are inferred, the user can get explanations for them using recursion. A disadvantage of the approach is that the algorithm does not memorize the nodes that were already used as explanation for the current node.

Timmer et al. [9] solved this problem in their argumentation-based approach by saving for each node a forbidden set of nodes containing all nodes that cannot be used anymore in an explanation.

Shi et al. [10] proposed decision graphs based on just the evidence nodes that reason equivalently to the Bayesian network. The chosen path is used to display just the most important evidence nodes for the current decision.

Müller et al. [7] presented a relevance-based approach that assigns a relevance value to each evidence item and subsequently imposes an ordering of items by importance. This approach is especially close to the reasoning process of clinicians in clinical decision-making. This relevance computation is used in this work.

²<https://norsys.com/netica.html>

³<https://norsys.com/netica.html>

⁴<https://www.symptomate.com>

3. Background

3.1. Bayesian Networks

A *Bayesian network* $B = (G, P)$ is a directed acyclic graph $G_B = (V, E)$, with V a set of *nodes* and $E \subseteq V \times V$, a set of *directed edges*, with an associated joint probability distribution P_B . Each node $i \in V$ of the network is linked to a random variable X_i , and there is a one-to-one correspondence between the set of nodes V and variables $X: V \leftrightarrow X$. Each variable $X_i \in X$ can take different states x_i . A variable X_i has an associated table stating a family of conditional probability distributions: how probable each state x_i of the variable X_i is depending on the states of its *parents*, $X_{\text{pa}(i)}$, i.e., $\text{pa}(i)$ are the nodes from which the node i has incoming edges [11]. Alternatively, it is said that the node i is the *child* of each of the nodes in the parents $\text{pa}(i)$.

The graph G_B states conditional dependencies between the variables of X through directed edges and conditional independencies through omitted edges. Given a realization of the set of variables X , the joint probability distribution P_B (1) can then be calculated as the product of the local probability distributions of all variables X_i given the realisation of their parent variables $X_{\text{pa}(i)}$ [12]:

$$P_B(X_1, X_2, \dots, X_n) = \prod_{i=1}^n P_B(X_i | X_{\text{pa}(i)}) \quad (1)$$

One concept used in this paper is the concept of a Markov blanket. A *Markov blanket* is defined as the set of the parents, children, and other parents of the children of a variable X_i [9]. This set contains all nodes that can have a direct impact on the node. If all nodes of the set are observed, then the variable X_i is independent of the other nodes in the network (conditioned on the Markov blanket) and its distribution can be derived from the nodes of that set alone.

3.2. ENDORISK

The primary use case for the research described here was to provide a flexible consultation software for making the ENDORISK Bayesian network on endometrial cancer [5] accessible in a usable way for clinicians without any knowledge of Bayesian networks. It focuses on preoperative risk stratification in endometrial cancer therapy. Approximately 10% of endometrial cancer patients present with lymph node metastasis at diagnosis. In most cases these metastasis are not recognized by current imaging modalities, and require lymph node dissection as gold standard. Preoperatively identifying these patients allows for proper surgical management and adjuvant therapy tailored to their needs. On the other side, identifying patients with a low risk of lymph node metastasis prevents increased surgical-related morbidity and unnecessary adjuvant therapy. Thus, individualized treatment improves the quality of care for all patients. To facilitate stratification, a Bayesian network was developed that predicts lymph node metastasis and 5-year survival by using preoperative biomarkers.

The network was developed on the basis of data obtained from a cohort study of 763 patients who had been surgically treated for endometrial cancer. The network was validated on two more external cohorts of together 830 endometrial cancer patients, which yielded a performance of the Area Under the Curve (AUC) of over 0.82.

4. Methodology

4.1. Development Process

The project started with a requirement analysis, to study the users needs and to define the project goals. In the following prototyping phase, different visual and functional designs of the website were created and compared. Subsequently, the best approach was implemented. During the project, regular group meetings and multiple evaluation sessions with expert gynecologists and medical informaticians were held, with a final evaluation with gynecologists near the end of the project.

4.2. Requirement Analysis

Our primary user group are clinicians. Consequently, our approach should not focus on explaining the medical background of a given model. The main challenge was to design a user interface that requires little to no knowledge of Bayesian networks. One important aspect considered was that Bayesian networks are based on (conditional) probabilities just as clinical decision making. The risk of getting a disease has always to be weighed against the risks and benefits of treatment, which is what doctors do. However, it has been observed that clinicians have difficulty in working with concrete probabilities in the context of clinical decision making [13], which also should be taken into account. In contrast to other Bayesian networks in the clinical field, the *ENDORISK* network was created for patients who already have a diagnosis. Thus, the approach should accordingly be *therapy-oriented*. To support re-usability, the system should support the ability to use different therapy-oriented Bayesian networks. Finally, the clinician is responsible for recommending and explaining the best therapy options to the patient. The interface was thus designed to support clinicians and, accordingly, has to provide explanations of its recommendations, supporting the creation of trust and understanding.

5. Result: DoctorBN

5.1. Architecture and Start Interface

The user interface of the resulted “DoctorBN”⁵ system is web-based. The home page of the website is used for network selection and offers general user information through a FAQ (Fig. 1). The user can choose between selecting one of the predefined networks or uploading a new one. Networks uploaded by the user are saved locally and the user can request to include them in the public network database. After the user selects a network, the *main view* is displayed (Fig. 2) together with a short tutorial. Then, the user can freely interact with the different views of the software.

5.2. Data Input, Output and Privacy

The first step of a typical workflow is that the user inserts the patient information. Patient data can be loaded from a CSV file or inserted using the input menus. The data is structured in three

⁵<https://doctorbn.herokuapp.com>

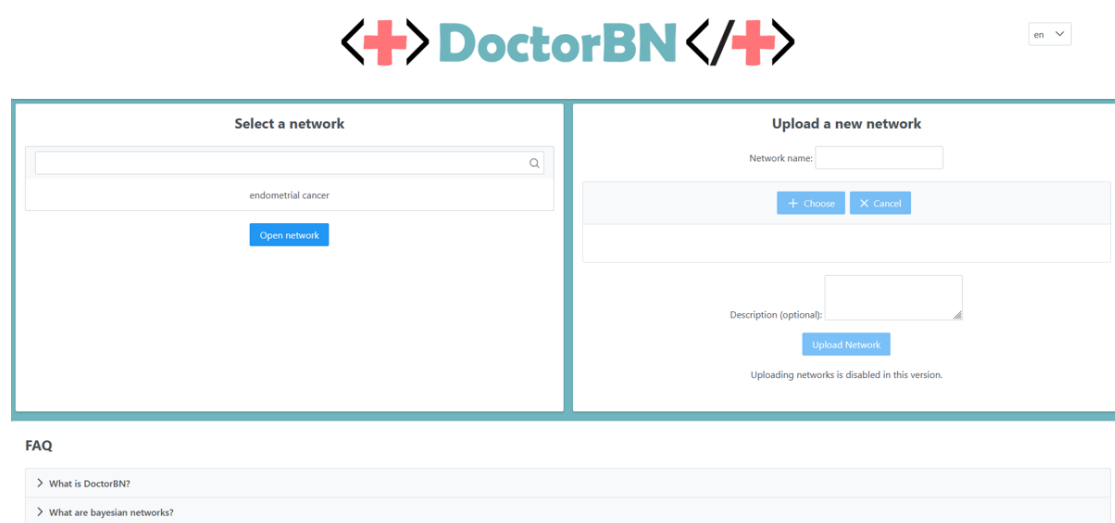


Figure 1: Network selection page: The user can select a given network or upload a new one from their file system. Additionally, a FAQ is provided that introduces the project.

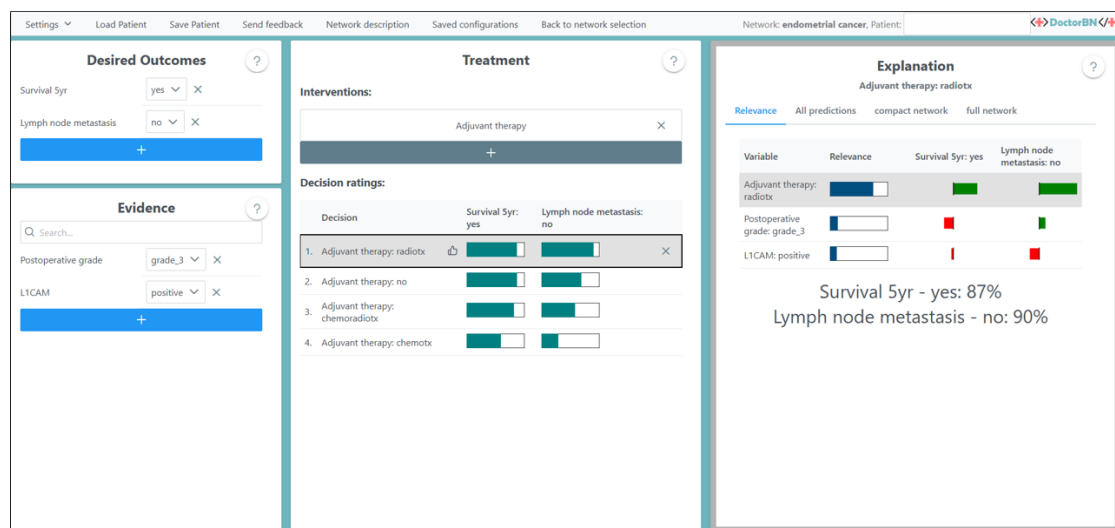


Figure 2: The website consists of multiple views. Left: Input windows for patient information. Center: Treatment view with treatments and treatment recommendations. Right: Explanation view with different tabs.

categories:

- *Evidence*: Everything that is known and cannot be changed about the patient, e.g. symptoms, demographics, etc.
- *Desired Outcomes*: the desired results for the patient case, typically that the patient wants

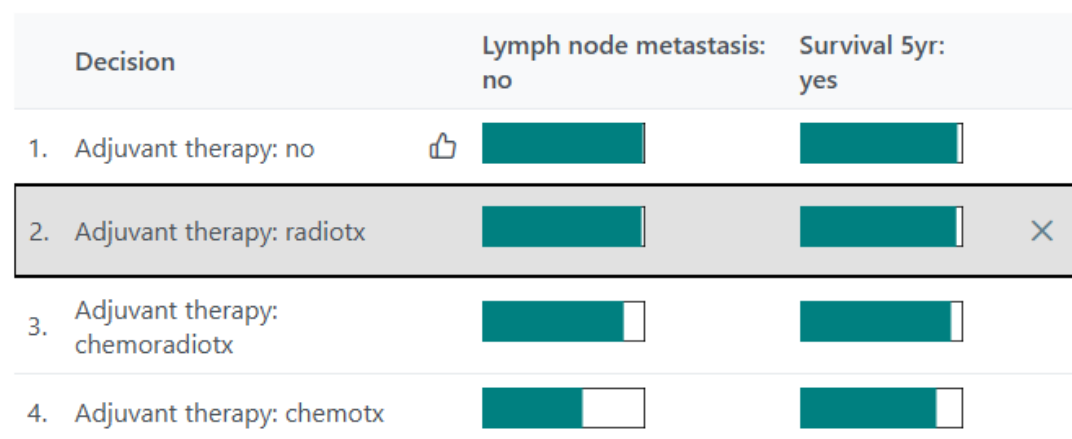


Figure 3: The Treatment Comparison view displays the different treatment recommendations ordered by their likeliness to fulfill all desired outcomes. If a treatment is selected for additional information, it is highlighted in grey.

to survive or have no lymph node metastases.

- *Interventions:* All variables that can be changed to achieve the desired outcomes, e.g. treatment.

5.3. Treatment Comparison

With the patient information as input, the algorithm will iterate through all combinations of interventions (treatments) to find out how likely the desired outcomes can be achieved. These recommendations are then sorted by their joint probability: how likely it is that all desired outcomes are achieved together. As this is a complex concept to grasp for the user, who lacks detailed probabilistic knowledge, just the individual probabilities are displayed, showing how likely each individual goal can be achieved. Bar charts were chosen for display as they allow for easy comparison of the probabilities. The list of *recommendations* displayed to the user is shown in Figure 3.

5.4. Outcome Explanation

When evidence and desired outcomes are stated or a recommendation is chosen, the user will get additional information presented in the “explanation” window to the right. The explanation window consists of multiple tabs for multiple use cases to provide additional information or explanations (Fig. 2, right).

Relevance View The relevance view is shown as the standard tab of the explanation view (Fig. 4). It provides a relevance-based explanation of the current outcomes: all evidence items are listed in a list view, ordered by their global relevance for the calculation. The local relevance is displayed too: for each desired outcome, a two-sided bar diagram visualizes through its width

Variable	Relevance	Survival 5yr: yes	Lymph node metastasis: no
Adjuvant therapy: radiotx			
Postoperative grade: grade_3			
L1CAM: positive			

Figure 4: The relevance view displays how relevant each evidence and treatment item is for the desired outcomes. For each item its global relevance is shown, as well as its impact on each individual desired outcome.

Variable \updownarrow	Prediction	Likelihood
CA125	lt_35	
Endometrium in cervical cytology	no	
Myometrial invasion	lt_50	

Figure 5: The “All predictions” view lists all variables of the network, together with each variable’s most likely prediction and it’s likelihood, or if it’s given by the user.

the relevance of the evidence, and through its direction and color if the impact was positive or negative.

Predictions List View This view lists all nodes of the network with their most likely state (Fig. 5). How probable that state is, is displayed through a color-coded tag either stating whether the value was given by the user or whether the probability is very likely, less likely, or not likely at all. This allows for fast information about the state of the network and therefore the consequences of the chosen decision. The omission of edges makes the view very compact and simple, especially for users who lack experience with Bayesian networks.

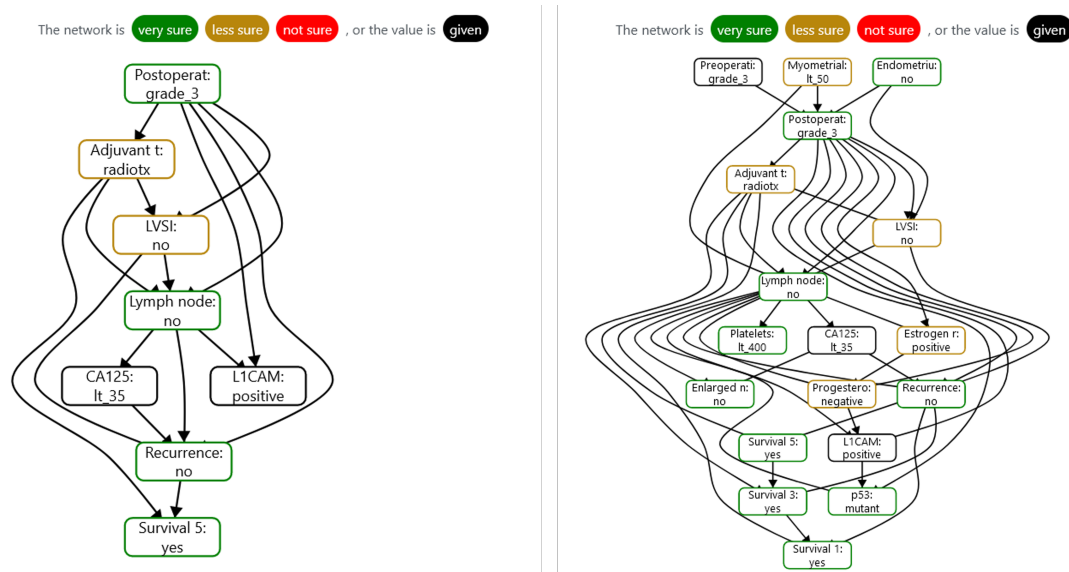


Figure 6: The two provided network views. Left: A compact view that shows the most relevant nodes for the current calculation. Right: The complete network.

Full Network View The full network view shows a graph view of the underlying Bayesian network (Fig. 6, right). Each node is shown with its most likely state. The probability of that state is shown through the border color, again color-coded in the categories “given”, “very likely”, “less likely” and “not likely”. The graph is displayed using the Sugiyama layout [14], as it provides a deterministic layout and structures the nodes causally from bottom to top.

Compact Network View The compact view simplifies the full network view by just showing the most relevant nodes (Fig. 6, left). This reduces visual clutter in complex networks and provides information on the reasoning process. Starting from the desired outcomes, possible relevant paths are computed using the Markov blanket and hidden sets based on the work by Timmer et al. [9]. Paths are just recursively followed when the change in probability distribution compared to an uninstantiated network is increasing in the direction of the evidence item. This is used as indication that the item is influencing the nodes on the path more than other influences and is therefore having a relevant influence on the desired outcome using this path. When a path ends with an evidence item its nodes are considered relevant. To avoid confusion with other network views in the final view, the edges of the network instead of the paths are shown.

5.5. Case Comparison

One important use case for the CDSS was the ability to compare two patient cases or recommendations. All views were adapted to provide a compact overview of two configurations. For example, the data input views were changed to a table format and the node border colors in the network views now highlight nodes that have different values in both configurations, instead

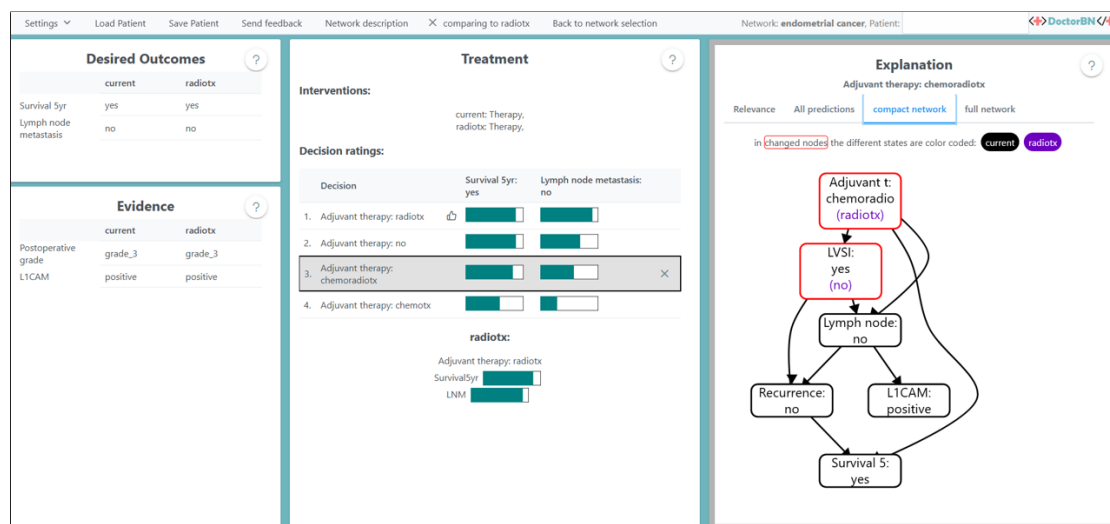


Figure 7: Comparison view comparing chemotherapy with radiotherapy.

of showing probabilities.

5.6. Feedback on System and Model Functionality

Bayesian networks are often partially or fully learned and are usually too complex to be fully seen as white-box models. This means that even after evaluating a Bayesian network, it could produce surprising or wrong output results. A feedback form allows users to describe such cases or other issues and optionally include the current configuration in the collected feedback.

6. Evaluation

An evaluation of the system was conducted through interviews with six clinicians with expertise on endometrial cancer. The participants accessed the DoctorBN software on their own computers. This allowed for a more realistic user experience. To obtain comparable feedback, all clinicians used the *ENDORISK* network. During the interviews, the participants were asked to complete a series of tasks and think aloud while performing them. The tasks were designed in such a way as to allow for the evaluation of whether or not a participant detected a functionality, was able to understand it, and knew how to explore it. Afterwards, the participants were asked to fill out a feedback form evaluating their experience using the website.

The evaluation results were generally positive (Fig. 8). All clinicians stated that they would like to use the software in clinical practice. They mostly agreed that the software provides sufficient explanation. However, they differed in how they perceived the usability of the website. Some problems in usability were addressed after the evaluation by reworking the concerned features. Especially the connection between the treatment view and the explanation view was simplified. Users were unable to find some system features, such as the comparison view,

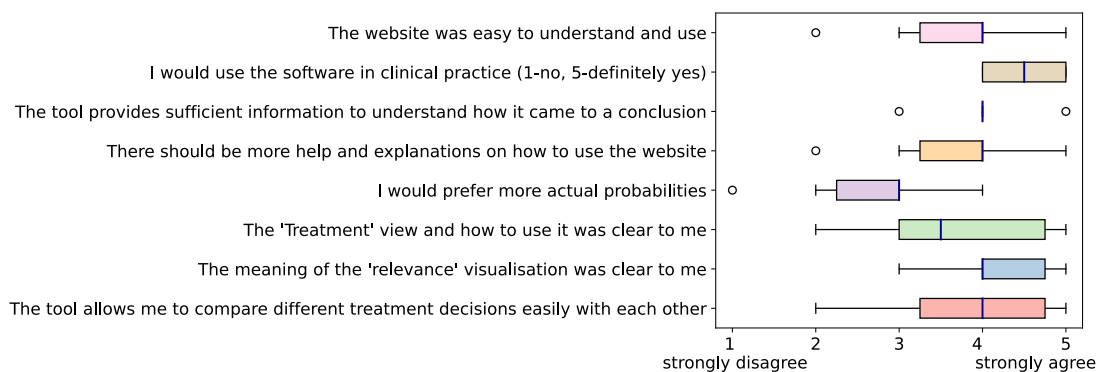


Figure 8: Results of the evaluation with six clinicians, summarized using boxplots.

which were accordingly made more accessible. Some results of the evaluation, however, require more complex revisions. This includes how to provide more specific network support for some networks without losing the ability to generally work with all networks. Additionally, the evaluation revealed some bias in the underlying network regarding adjuvant therapy, which has to be improved for clinical use.

7. Discussion and Conclusions

This work presents an approach for computer-based clinical decision support, illustrated by the Bayesian network on endometrial cancer. The approach developed is reusable with other Bayesian networks. Treatment recommendations are automatically generated from user input and ranked by their ability to reach the desired outcomes, such as the survival of the patient. The approach also provides different visualizations to let the user gain additional insight and understand how the network came to its conclusions. Those explanations are crucial to generate trust when working with complex models such as Bayesian networks. The evaluation revealed a high interest in the system, although there are still some usability problems that have to be addressed. The participating clinicians agreed to be willing to use the system in clinical practice. The approach is a contribution to the general goal of making complex computational models like Bayesian networks available to domain experts with little to no experience in using them. CDSS will improve patient treatment and safety by supporting healthcare and therapy tailored to the individual patient.

Acknowledgments

Our work was supported by the German Federal State of Saxony-Anhalt (FKZ: I 88).

References

- [1] D. Pritchard, F. Moeckel, M. Villa, L. Housman, C. McCarty, H. McLeod, Strategies for integrating personalized medicine into healthcare practice, *Per Med* (2017) 141–152. doi:10.2217/pme-2016-0064.
- [2] A. L. Madsen, M. Lang, U. B. Kjærulff, F. Jensen, The Hugin tool for learning bayesian networks, in: *European conference on symbolic and quantitative approaches to reasoning and uncertainty*, Springer, 2003, pp. 594–605.
- [3] I. Beinlich, E. H. Herskovits, Ergo: a graphical environment for constructing bayesian, arXiv preprint arXiv:1304.1095 (2013).
- [4] R. T. Sutton, D. Pincock, D. C. Baumgart, D. C. Sadowski, R. N. Fedorak, K. I. Kroeker, An overview of clinical decision support systems: benefits, risks, and strategies for success, *NPJ digital medicine* 3 (2020) 1–10.
- [5] C. Reijnen, E. Gogou, N. C. M. Visser, ..., H. V. N. Küsters-Vandeveld, P. J. F. Lucas, J. M. A. Pijnenborg, Preoperative risk stratification in endometrial cancer (endorisk) by a bayesian network model: A development and validation study, *PLoS medicine* 17 (2020) e1003111.
- [6] M. J. Druzdzel, Smile: Structural modeling, inference, and learning engine and genie: a development environment for graphical decision-theoretic models, in: *Aaai/Iaai*, 1999, pp. 902–903.
- [7] J. Müller, M. Stoehr, A. Oeser, J. Gaebel, M. Streit, A. Dietz, S. Oeltze-Jafra, A visual approach to explainable computerized clinical decision support, *Computers & Graphics* 91 (2020) 1–11.
- [8] G.-E. Yap, A.-H. Tan, H.-H. Pang, Explaining inferences in bayesian networks, *Applied Intelligence* 29 (2008) 263–278.
- [9] S. T. Timmer, J.-J. C. Meyer, H. Prakken, S. Renooij, B. Verheij, Explaining bayesian networks using argumentation, in: *European Conference on Symbolic and Quantitative Approaches to Reasoning and Uncertainty*, Springer, 2015, pp. 83–92.
- [10] A. Shih, A. Choi, A. Darwiche, Compiling bayesian network classifiers into decision graphs, in: *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 33, 2019, pp. 7966–7974.
- [11] I. Ben-Gal, Bayesian networks, *Encyclopedia of statistics in quality and reliability* 1 (2008).
- [12] F. V. Jensen, et al., *An introduction to Bayesian networks*, volume 210, UCL press London, 1996.
- [13] D. M. Eddy, *Probabilistic reasoning in clinical medicine: Problems and opportunities*, Cambridge University Press (1982) 249–267.
- [14] K. Sugiyama, S. Tagawa, M. Toda, Methods for visual understanding of hierarchical system structures, *IEEE Transactions on Systems, Man, and Cybernetics* 11 (1981) 109–125.