Bayesian Networks in Healthcare: the chasm between research enthusiasm and clinical adoption

Evangelia Kyrimi^{1,*}, Scott McLachlan^{1,2}, Kudakwashe Dube^{2,3}, Norman Fenton¹

¹ Risk and Information Management, Queen Mary University of London, United Kingdom

² Health informatics and Knowledge Engineering Research (HiKER) Group

³ School of Fundamental Sciences, Massey University, New Zealand

Abstract:

Problem: Bayesian Networks (BN) can address real-world decision-making problems, and there is enormous and rapidly increasing interest in their use in healthcare. Yet, despite thousands of BNs in healthcare papers published yearly, evidence of their adoption in practice is extremely limited and there is no consensus on why.

Method: A preliminary review was conducted to identify research gaps and justify the conduct of a broader scoping or systematic review of BNs in healthcare

Results: We highlight that: (1) there have been no significant attempts to systematically review the domain; (2) there are weaknesses in the way BN development processes are presented in the literature; and (3) in contrast to enthusiasm (including from clinicians) there has been negligible adoption of published BNs into clinical practice.

Conclusion: A systematic review of BNs in healthcare is needed to a) understand the chasm between research enthusiasm and clinical adoption; and b) improve clinical adoption.

Keywords: Bayesian networks, medical reasoning, clinical decision support, healthcare

1. Introduction

Clinical decision making is a complex evolving process where evidence is gathered and diagnostic and treatment decisions are made (Tiffen, Corbridge and Slimmer, 2014). While clinicians are generally considered to be good decision-makers, it can be challenging to combine all available evidence and accurately reason under conditions of uncertainty (Bornstein and Emler, 2001). Many clinical decision support systems (CDSS) have been developed, and graphical probabilistic models such as Bayesian Networks (BNs) are one popular approach (Druzdzel and Flynn, 2002; Beattie and Nelson, 2006; Patel *et al.*, 2009; Adams and Leveson, 2012; Shortliffe and Cimino, 2013; Lucas, 2001; Lucas, van der Gaag and Abu-Hanna, 2004).

BNs are based on Bayes' theorem and model causal or influential relationships between variables. Bayes' theorem represents a formal method for understanding the natural thought process people undergo while evolving their understanding of the likelihood of one phenomenon under observation as their knowledge of related influential phenomena increases. For instance, when clinicians perform differential diagnosis where multiple diseases present with similar symptoms: as new information from diagnostic tests and clinical examination of the patient become available, clinicians update their belief on what they believe is the illness causing the patient's poor health. BNs use a graphical approach for *compact representation of multivariate probability distributions* and *efficient reasoning under uncertainty*. Their popularity in healthcare results from their interpretable graphical structure and their ability to: (i) model complex problems with causal dependencies where a significant degree of uncertainty is involved; (ii) combine different sources of information such as data and experts' judgement; and (iii) model interventions and reason both diagnostically and prognostically. Based

* Corresponding author:

Email address: e.kyrimi@qmul.ac.uk (Evangelia Kyrimi)

Postal address: 10, Godward Square, Queen Mary University of London, Mile End Rd, Bethnal Green, London E1 4FZ

on limited marketing materials, there are indications that some major healthcare technology companies are using BNs in their non-public research and application development (Armstrong, 2018), including what is claimed as the largest BN in the world (MMC Ventures, 2017).

While working on several projects related to clinical decision support using BNs we realised that: (i) even though a vast number of healthcare related BNs are published every year, no existing paper describes the scope of BNs developed for use in healthcare; (ii) apart from a small number of micro-reviews about BNs for specific medical conditions, e.g. (Bielza and Larranaga, 2014) and one small-scale epidemiology-focused review on directed acyclic graphs that is yet to be peer reviewed (Tennant *et al.*, 2019), the domain lacks a systematic or scoping review; and, (iii) that while some researchers have investigated potential reasons why a model might not be useful in practice, e.g. (Wyatt and Altman, 1995; Blackmore, 2005; Reilly and Evans, 2006; Moons *et al.*, 2009; Adams and Leveson, 2012), no review investigates which of the published medical BNs identifies or acknowledges the reasons why a model may or may not have been adopted in healthcare.

In contrast to widely discussed and accepted benefits of BNs, it appears that negligible clinical adoption of *published* BNs into healthcare practice has occurred. We believe that it is only when we understand the necessary rules for adopting a BN in clinical practice, and whether or not the BNs described in the literature adhere to these rules, that we may bridge the gap between development and adoption of BNs in healthcare. *It is our intention in this work to draw attention to this gap through a small-scale investigation that briefly explores the challenges inhibiting BN development and adoption for healthcare application.*

The remainder of this paper is organised as follows: Section 2 presents the method used to achieve the objectives of this paper; Section 3 discusses the results; Section 4 proposes the way forward; Section 5 presents our concluding remarks.

2. Method

A search using Queen Mary University's library search system that includes access to PubMed, MedLine, ScienceDirect, Scopus, DOAJ and Elsevier was performed using a selection of keywords arranged in the following search algorithm:

```
"((Bayes OR Bayesian) AND network) OR (probabilistic AND graphical AND model)) AND (medical OR clinical)"
```

To identify the existing knowledge gaps in this research area, a small-scale review was proposed to analyse a limited number of papers from which we could identify research gaps and develop an overarching view of the domain sufficient to justify effort for conduct of a broader scoping or systematic review of BNs in healthcare. Initially, 50 papers were selected based on their relevance to the search term as reported by the search engine. Relevance was computed based on both the number of times the search terms appear in the paper, and where within the paper the search terms appeared (title, abstract, keywords, keywords plus).

Table 1 identifies the basic review framework which was agreed by all of the authors of this paper. This framework was applied to the preliminary survey presented in this paper. It consists of a set of elements relating to generalisability of the development methodology and adoption of the BN in healthcare

Table 1: Preliminary Review Framework

1. <i>Generalisability of BN development process and its outcomes:</i>	☐ Overall development process ☐ BN structure elicitation from knowledge ☐ BN parameters elicitation from knowledge ☐ BN structure learning from data ☐ BN parameters learning from data
2. <i>BN adoption in healthcare:</i>	☐ Clinical benefit ☐ Clinical credibility ☐ Accuracy ☐ Generalisability ☐ Usability ☐ Potential impact

	Actual impact In use

3. Results and Discussion

Literature on BNs in Healthcare: Figure 1 demonstrates the year on year increase of publications on BNs in healthcare both in total and when narrowed to identify *the most highly relevant ones*.

Figure 1: Number of publications on medical Bayesian networks per year

Application of Preliminary Review Framework: Our objectives were to investigate: (1) generalisability of the BN development process; and, (2) BN adoption in healthcare. For inclusion, all authors agreed that papers had to possess all of the following criteria:

1. Describe a genuine BN model (or BN modelling process) rather than Bayesian statistics or naive Bayes;
2. Be targeted clearly at a medical problem; and,
3. Be intended to support clinicians or patients in decision making or prediction.

Our initial search identified a collection of 1894 papers for screening. Screening involved excluding papers: (i) that were duplicates; (ii) published outside the period 1995-2018; (iii) not published in English; (iv) not published in journal or conference proceedings; and (v) whose primary content was not healthcare related. After screening, 834 papers remained for eligibility assessment. From these, the 50 identified by the search engine as most relevant were selected. The 26 listed in Table 2 were those that met all three of the above inclusion criteria. The literature selection process is presented diagrammatically in Figure 2. Each paper was reviewed by two authors (EK and SM) and in cases of disagreement, consensus was required. In what follows we refer to papers listed in Table 2 by their number in column 1.

Figure 2: Preliminary review literature selection

Table 2 papers in preliminary review that met all inclusion characteristics

Paper Number	Authors (year)	Title	Source
1	Akhtar, M. J. and Utne, I. B. (2014)	Human fatigue's effect on the risk of maritime groundings - A Bayesian Network modeling approach	Safety Science 62
2	Burnside, E. S. <i>et al.</i> (2006)	Bayesian Network to Predict Breast Cancer Risk of Mammographic Microcalcifications and Reduce Number of Benign Biopsy Results: Initial Experience	Radiology 240(3)
3	Celi, L. <i>et al.</i> (2008)	An artificial intelligence tool to predict fluid requirement in the intensive care unit: a proof-of-concept study	Critical Care, 12(6)
4	Chang, Y. S. <i>et al.</i> (2015)	Mobile cloud-based depression diagnosis using an ontology and a Bayesian network	Future Generation Computer Systems, 43-44
5	Cruz-Ramírez, N. <i>et al.</i> (2007)	Diagnosis of breast cancer using Bayesian networks: A case study	Computers in Biology and Medicine, 37(11)
6	Dewan, A. and Sharma, M. (2015)	Prediction of heart disease using a hybrid technique in data mining classification	2015 International Conference on Computing for Sustainable Global Development
7	Díez, F. J. <i>et al.</i> (1997)	DIAVAL, a Bayesian expert system for echocardiography	Artificial Intelligence in Medicine, 10(1)
8	Van der Heijden, M., Velikova, M. and Lucas, P. J. F. (2014)	Learning Bayesian networks for clinical time series analysis	Journal of Biomedical Informatics, 48
9	Kline, J. A. <i>et al.</i> (2005)	Derivation and validation of a bayesian network to predict pretest probability of venous thromboembolism	Annals of Emergency Medicine, 45(3)
10	Magrini, A., Luciani, D. and Stefanini, F. M. (2018)	A probabilistic network for the diagnosis of acute cardiopulmonary diseases	Biometrical Journal, 60(1)
11	Moreira, M. W. L. <i>et al.</i> (2016)	A Preeclampsia Diagnosis Approach using Bayesian Networks	IEEE 2016 International Conference on Communications (ICC),
12	Nordmann, J. P. and Berdeaux, G. (2007)	Use of a Bayesian Network to Predict the Nighttime Intraocular Pressure Peak from Daytime Measurements	<i>Clinical Therapeutics</i> , 29(8)
13	Onisko, A., Druzdzel, M. J. and Austin, R. M. (2009)	Application of Dynamic Bayesian Networks to Cervical Cancer Screening	Intelligent Information Systems, 9
14	Onisko, A., Druzdzel, M. J. and Wasyluk, H. (1999)	A Bayesian Network Model for Diagnosis of Liver Disorders	7th Symposium on Intelligent Information Systems
15	Sandri, M. <i>et al.</i> (2014)	Dynamic Bayesian Networks to predict sequences of organ failures in patients admitted to ICU	Journal of Biomedical Informatics, 48
16	Seixas, F. L. <i>et al.</i> (2014)	A Bayesian network decision model for supporting the diagnosis of dementia, Alzheimer's disease and mild cognitive impairment	Computers in Biology and Medicine, 51
17	Sesen, M. B. <i>et al.</i> (2013)	Bayesian networks for clinical decision support in lung cancer care	PLoS ONE, 8(12)

18	Sesen, M. B. <i>et al.</i> (2014)	Lung Cancer Assistant: a hybrid clinical decision support application for lung cancer care	Journal of the Royal Society, 11
19	Velikova, M. <i>et al.</i> (2014)	Exploiting causal functional relationships in Bayesian network modelling for personalised healthcare	International Journal of Approximate Reasoning, 55
20	Wang, X.-H. <i>et al.</i> (1999)	Computer-assisted diagnosis of breast cancer using a data-driven Bayesian belief network	International Journal of Medical Informatics, 54(2)
21	Wang, Z. <i>et al.</i> (2015)	3-D Stent Detection in Intravascular OCT Using a Bayesian Network and Graph Search	IEEE Trans Med Imaging, 34(7)
22	Wasylyuk, H., Onisko, A. and Druzdzal, M. J. (2001)	Support of diagnosis of liver disorders based on a causal Bayesian network model	Medical Science Monitor, 7
23	Yet, B. <i>et al.</i> (2013)	Decision support system for Warfarin therapy management using Bayesian networks	Decision Support Systems, 55(2)
24	Yet, B., Perkins, Z., <i>et al.</i> (2014)	Not just data: a method for improving prediction with knowledge	Journal of biomedical informatics, 48
25	Zhang, Q. and Shi, X. (2017)	A mixture copula bayesian network model for multimodal genomic data	Cancer Informatics, 16(Section 4)
26	Zhang, Y. <i>et al.</i> (2015)	Identification of reciprocal causality between non-alcoholic fatty liver disease and metabolic syndrome by a simplified Bayesian network in a Chinese population	BMJ Open, 5(9)

BN development process: Only 5 out of the 26 reviewed papers provided a detailed repeatable *overall development process* [24, 16, 1, 6, 10]. Similarly, from the 13 papers in which the BN structure was elicited from knowledge, only 4 provided a generic structure such as a template [17, 18, 19, 10]. Moreover, no paper provided a detailed explanation of how the knowledge was elicited from experts. Often, references were given to activities such as interviews that had been conducted to elicit the expert knowledge without providing details of the interview process, questions asked, or responses received from the participants. While it is not uncommon to elicit the BN structure from knowledge, eliciting BN parameters from knowledge is generally difficult. In 5 of the 26 reviewed papers, we identified BN models whose parameters were elicited partially or completely from knowledge [2, 23, 19, 11, 10]. From these 5 papers, only one provided a repeatable *parameter elicitation process* [24]. Similar results were also identified where BN structure and/or parameters were learned from data. From the 15 papers where the BN structure was learned from data, only 4 provided detail of the *structure learning process* [17, 18, 8, 16]. BN parameters were learned from data in 24 of the 26 papers. yet only 3 offered detailed description of the *parameter learning process* and the way missing data were treated [23, 8, 16].

BN adoption in healthcare: The frequency of elements related to BN adoption in healthcare is presented in Figure 3. Most of the reviewed papers clearly described the *clinical benefit* of the medical BN and provided BNs with acceptable *accuracy*. However, *generalisability* was identified only once [24]. Generalisability is important for achieving BN adoption in healthcare as a model could not be used widely in practice if it has not been shown to work on alternative populations. This strongly indicates a research gap limiting the application of BNs to populations that differ from the one used in the original training of the model. Two of the rarest elements to identify were the model's *usability* and *impact* (either potential or actual). This agrees with our initial remark that there is a significant gap between developing an accurate model and having a useful and usable model that can be adopted and have an impact in healthcare. In this preliminary review no BN adhered to all the rules necessary for having a 'useful' model. **Finally, and crucially, no paper in the preliminary review collection provided any indication at all that the BN models were in current clinical use.**

Figure 3: Frequency of the elements related to the BN adoption in healthcare

4.4. Summary and the way forward

This small-scale preliminary review found that literature on medical BNs can be broadly grouped into: (1) method-driven papers, where the focus is on a particular aspect of the modelling methodology; and, (2) case-driven papers, where the focus is on the medical application.

The key findings are that:

1. The literature lacks a well described overall BN development process;
2. Most works only briefly present the way the BN structure and parameters were elicited or learned from knowledge or data, respectively. This makes their methodology difficult, if not impossible, to replicate.

3. Important elements, such as the BN's generalisability, usability and impact, which are necessary for having a useful model, have received little or no attention. This creates a fundamental gap between developing an accurate BN and adopting it in healthcare.
4. Despite immense interest and research effort in BNs to support clinical decision making, it appears that negligible clinical adoption of *published* BNs into healthcare practice has occurred.

The fact that BNs do not appear to have achieved a significant (publicly accessible) contribution to healthcare suggests that researchers in BNs for healthcare must start to address serious challenges. While our analysis used a sound review framework, it is only a preliminary review where a small number of papers have been selected based on their relevance to the search term to establish the rationale for a more comprehensive investigation. We believe systematic or scoping reviews that summarise and organise the research contributions in the domain may help. The lack of such reviews has likely led to duplication of effort, while allowing important research gaps to stay hidden. The high volume of published works and lack of significant attempts to systematically review the domain are significant motivating factors for undertaking a systematic or scoping review to summarise the literature and highlight research gaps and future directions. A robust scoping review conforming to the PRISMA statement is underway and promises further contributions.

Funding

SM, NF and EK acknowledge support from the Engineering and Physical Sciences Research Council (EPSRC) under project EP/P009964/1: PAMBAYESIAN: Patient Managed decision-support using Bayes Networks. KD acknowledges financial support from Massey University for his study sabbatical with the PamBayesian team.

Competing Interests

No author identified a competing interest relevant to this research.

Contributors

EK prepared the first draft. EK proposed the preliminary framework, which was refined by SM, KD and NF. EK and SM conducted the preliminary review. KD and NF supervised the research. All authors contributed, commented and approved the final draft.

References

- Adams, S. T. and Leveson, S. H. (2012) 'Clinical prediction rules', *BMJ*, 344, p. d8312. doi: 10.1136/bmj.d8312.
- Armstrong, S. (2018) 'The apps attempting to transfer NHS 111 online', *BMJ (Online)*, 360(January 2018), pp. 1–5. doi: 10.1136/bmj.k156.
- Beattie, P. and Nelson, R. (2006) 'Clinical prediction rules: what are they and what do they tell us?', *The Australian journal of physiotherapy*. Elsevier, 52(3), pp. 157–163. doi: 10.1016/S0004-9514(06)70024-1.
- Bielza, C. and Larranaga, P. (2014) 'Bayesian networks in neuroscience: a survey', *Frontiers in Computational Neuroscience*, 8, pp. 1–23. doi: 10.3389/fncom.2014.00131.
- Blackmore, C. C. (2005) 'Clinical prediction rules in trauma imaging: who, how, and why?', *Radiology*, 235, pp. 371–374. doi: 10.1148/radiol.2352040307.
- Bornstein, B. H. and Emler, A. C. (2001) 'Rationality in medical decision making: A review of the literature on doctors' decision-making biases', *Journal of Evaluation in Clinical Practice*, 7(2), pp. 97–107. doi: 10.1046/j.1365-2753.2001.00284.x.
- Druzdel, M. J. and Flynn, R. R. (2002) *Decision Support Systems, Encyclopedia of Library and Information Science, Second Edition*. doi: 10.1081/E-ELIS3-120043887.
- Lucas, P. (2001) 'Bayesian networks in medicine: a Model-based Approach to Medical Decision Making', in *EUNITE workshop on Intelligent Systems in Patient Care*.
- Lucas, P., van der Gaag, L. C. and Abu-Hanna, A. (2004) 'Bayesian networks in biomedicine and health-care', *Artificial intelligence in medicine*, 30, pp. 201–214. doi: 10.1016/j.artmed.2003.11.001.
- MMC Ventures (2017) *Beyond The Hype: Artificial Intelligence Podcast Interview - Babylon Health*. Available at: <https://www.healthcare.digital/single-post/2018/03/29/Healing-Healthcare-with-Artificial-Intelligence- 'Beyond-The-Hype'-with-Dr-Ali-Parsa-Founder-and-CEO-of-Babylon-Health#>!

- Moons, K. G. M. *et al.* (2009) 'Prognosis and prognostic research: what, why, and how?', *BMJ*, (338), p. 375.
- Patel, V. L. *et al.* (2009) 'The coming of age of artificial intelligence in medicine', *Artificial Intelligence in Medicine*, 46(1), pp. 5–17. doi: 10.1016/j.artmed.2008.07.017.
- Reilly, B. M. and Evans, A. T. (2006) 'Translating clinical research into clinical practice: Impact of using prediction rules to make decisions', *Annals of Internal Medicine*, 144, pp. 201–209. Available at: <http://www.scopus.com/inward/record.url?eid=2-s2.0-33644855006&partnerID=40&md5=903fb2ad9598fcbdc1984db9a9af8784>.
- Shortliffe, E. H. and Cimino, J. J. (2013) *Biomedical Informatics: Computer Applications in Health Care and Biomedicine, Fourth Edition*, Springer. doi: 10.1007/978-1-4471-4474-8.
- Tennant, P. W. G. *et al.* (2019) 'Use of directed acyclic graphs (DAGs) in applied health research: review and recommendations', *medRxiv*. doi: <http://dx.doi.org/10.1101/2019.12.20.19015511>.
- Tiffen, J., Corbridge, S. J. and Slimmer, L. (2014) 'Enhancing Clinical Decision Making: Development of a Contiguous Definition and Conceptual Framework', *Journal of Professional Nursing*, 30(5), pp. 399–405. doi: 10.1016/j.profnurs.2014.01.006.
- Wyatt, J. C. and Altman, D. G. (1995) 'Commentary: Prognostic models: clinically useful or quickly forgotten?', *BMJ*, 311, pp. 1539–1541. doi: 10.1136/bmj.311.7019.1539.