A Review of Dynamic Bayesian Network Techniques with Applications in Healthcare Risk Modelling

Mohsen Mesgarpour[1], Thierry Chaussalet[2], and Salma Chahed[3]

1 **HSCMG, Faculty of Science and Technology, University of Westminster**
115 New Cavendish Street, London, W1W 6UW, UK
mohsen.mesgarpour@my.westminster.ac.uk

2 **HSCMG, Faculty of Science and Technology, University of Westminster**
115 New Cavendish Street, London, W1W 6UW, UK
chausst@westminster.ac.uk

3 **HSCMG, Faculty of Science and Technology, University of Westminster**
115 New Cavendish Street, London, W1W 6UW, UK
S.Chahed@westminster.ac.uk

Abstract

Coping with an ageing population is a major concern for healthcare organisations around the world. The average cost of hospital care is higher than social care for older and terminally ill patients. Moreover, the average cost of social care increases with the age of the patient. Therefore, it is important to make efficient and fair capacity planning which also incorporates patient centred outcomes. Predictive models can provide predictions which their accuracy can be understood and quantified. Predictive modelling can help patients and carers to get the appropriate support services, and allow clinical decision-makers to improve care quality and reduce the cost of inappropriate hospital and Accident and Emergency admissions. The aim of this study is to provide a review of modelling techniques and frameworks for predictive risk modelling of patients in hospital, based on routinely collected data such as the Hospital Episode Statistics database. A number of sub-problems can be considered such as Length-of-Stay and End-of-Life predictive modelling. The methodologies in the literature are mainly focused on addressing the problems using regression methods and Markov models, and the majority lack generalisability. In some cases, the robustness, accuracy and re-usability of predictive risk models have been shown to be improved using Machine Learning methods. Dynamic Bayesian Network techniques can represent complex correlations models and include small probabilities into the solution. The main focus of this study is to provide a review of major time-varying Dynamic Bayesian Network techniques with applications in healthcare predictive risk modelling.

1998 ACM Subject Classification G.3 Probability and Statistics

Keywords and phrases Healthcare Modelling, Dynamic Bayesian Network, Predictive Risk Modelling, Time-Varying, Hospital Administrative Data

Digital Object Identifier 10.4230/OASIcs.SCOR.2014.89

1 Introduction

Costs of care are increasing at a rate that is unaffordable in the current economy. This is mainly due to the impact of ageing population, population growth, deprivation, increased expectations and cost of treatment and technology [45, 19]. The current system is unsustainable and unfair, and the current financial options to support people in meeting care costs are limited.

4th Student Conference on Operational Research (SCOR'14).
Editors: Pedro Crespo Del Granado, Martim Joyce-Moniz, and Stefan Ravizza; pp. 89–100
OpenAccess Series in Informatics
OASICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

In the fourth consecutive year, in 2013/14 local authorities have reduced social care budget [46]. Looking further ahead, it is projected that people aged over 85 to almost double by 2030, with an additional 600,000 of ageing population will be needing significant care [44]. While a quarter of people aged over 65 will need to spend very little on care over their life, half can expect cost of up to £20,000, and one in 10 can expect cost of £100,000 [44]. The fairness and quality of care can only be guaranteed if the cost saving changes are sustainable in the way care is provided. Because, there is some evidence that health and social care act as substitutes, pressure on one system can affect the other [27].

The main content is divided into seven sections. In the second section, the background is described briefly. The Predictive Modelling (PM) with applications to risk adjustment and predictive risk modelling is explained in the third section. The fourth section highlights the techniques in healthcare PM with emphasis on Dynamic Bayesian Network (DBN) modelling. Then, the main approaches in modelling time-varying DBN are highlighted in the fifth section. In the sixth section, the gaps and limitations of the current solutions are outlined. In the seventh section, the challenges are listed. Finally, in section eight, the future research areas are stated.

2 Background

Identification of high risk events has been a major concern for many healthcare providers and purchasers. According to Lewis et al. [43], there are three major sources of risks to the healthcare systems:

- Ageing population;
- Increase in number of people with long-term conditions;
- Rise in rate of emergency (or unplanned) admission to hospitals.

Firstly, a major concern in healthcare organisations throughout the world is about coping with an ageing population [10]. Survey results [12, 42] show that many people would prefer to die with an appropriate support at home rather than hospital, and yet the number of death in hospital can reach to 65% in the future if no appropriate policy would be in place.

Moreover, the average cost of hospital care is higher than social care for older and terminally ill patients. Also, the costs of care in the final phases before death are very high in hospital. According to the Nuffield Trust report [27], the average cost of social care increases with the age of the patient. However, the cost of social care stays cheaper than the hospital care for those whose age is below 85. Based on gender, the intersection point of hospital and social care costs for male patients are close to the age 90 and for the female patients are approximately at the age 80.

Sometimes a hospital admission can be avoided by residential setting substitution or social care. According to the analysis by Bardsley et al. [5] on a wide population in England, use of a social care may prevent the need for the hospital care. The End-of-Life researches help patients to get appropriate support services towards the end of life by better management of resources and patients. The ambition of National Health Service (NHS) is to increase the number of people who die in their usual place of residence to 60% . This baseline in 2007 was 38%, and with the End-of-Life practices that are in place it reached to 42% in 2012 [51, 52].

Secondly, ageing population and change of lifestyles are related to increase in number of people with long-term conditions and morbidity conditions (any condition that represent departure from a state of well-being). For instance, there was a significant rise in chronic kidney disease, diabetes and cancer between years 2006 and 2011 [18]. Also, it has been predicted that the number of people with co-morbidity condition (dual diagnoses) to rise from 1.9 million in 2009 to 2.9 million in 2018 [47].

Thirdly, an estimated £11 billion per year is the cost of unplanned admissions to the NHS in England. There are two main causes of emergency admissions: early discharge and unpredictable accidents and emergency events. Based on a retrospective study by Clarke et al. [9], about half of the 30-day emergency readmissions were potentially preventable between 2004 and 2010. Discharging patients is a primary way of providing free beds in the healthcare systems, and administrators are in charge of evaluating the risk of emergency readmission in advance. Patient flow modelling such as Length-of-Stay PM enables managers to better understand the operational and clinical functions [2]. Length-of-Stay PM includes capturing the flow of patient from the admission to discharge. The flow is through a number of conceptual (virtual) phases that each patient goes through them. The PM of Length-of-Stay uses the time spent in phases in addition to the clinical data and the demographic data to identify events.

Therefore, it is important to make efficient and fair capacity planning. Predictive risk models that include the patient centred outcomes, as well as the aggregated levels, can help patients and carers to get the appropriate support services in clinical decision making. Also, PM of risks can improve care quality and reduce the cost of inappropriate hospital and Accident and Emergency (A&E) admissions.

The aim of this study is to provide a review of DBN techniques in healthcare predictive risk modelling, such as Length-of-Stay and End-of-Life PM.

3 Predictive Modelling

PM is often associated with Machine Learning (ML), pattern recognition and data mining. The practice of PM defines the process of development of models that their prediction accuracy can be understood and quantified [38]. Geisser [25] defined PM as "the process by which a model is created or chosen to try to best predict the probability of an outcome."

In the field of decision analysis, risk modelling [22] can answer questions such as:

- Is the risk high or low, and is the level acceptable?
- Which are the most critical causal factors?
- What are the differences between the risks of the alternative solutions?
- What can be achieved in terms of risk-reducing effect in comparison with the other options?

Generally, the identification of emergent risk can be categorised into modelling of three main aspects: stratification, clinical profiles and resource utilisation profiles. Moreover, in modelling of the events, the time dimension is usually modelled as time-to-event or as a risk score.

For instance, Gotz et al. [29] at the Watson Research Centre introduced an approach based on patient similarity, which extract a cohort of patients from an Electronic Health Record that is similar to specific patient, and provides interactive visualisation for complex decision making. It has been shown that this clustering, dynamic visualisation and expert refinement had been effective for near-term prognosis [21] and risk evaluation [11] of physiological data.

There are two major branches of risk modelling in healthcare: PM and Risk Adjustment. PM is frequently used for finding high-risk member Case Finding, such as finding patients with high risk of readmission and also to predict member cost and utilisation.

Risk Adjustment is a normalisation technique for comparison purposes, such as classifying patients by potential risk level for the purpose of the insurance provider reimbursement [43, 33]. These two methods are briefly summarised in the following subsections.

3.1 Predictive Risk Modelling

In healthcare systems various types of scoring systems (e.g. Glasgow Coma Scale for patients with brain injury) are used to support clinical and administrative decisions; however, statistical and stochastic models are needed to estimate the risks according to the changes in the care and environmental variables [60].

Physicians are interested in the evaluation and forecast of the adverse events that may provoke mortality or longer hospital stay for the patient, and assign a quantity to the patient's risk profile [14]. In terms of risk impact, healthcare risk analysis can be categorised into two categories: Operational Risks and Clinical Risks [36].

Data mining in healthcare can predict risks in healthcare, improve health status of high-risk patients and make overall savings. There are various predictive risk models in the literature, but each can forecast a small range of healthcare and social care outcomes. They differ in terms of the predicted time range, variables, data sources and the modelling approaches [43].

For instance, Billing et al. [6] developed a predictive risk model for re-admission using a multivariate logistic regression based on Length-of-Stay, diagnoses and demographic data derived from the Hospital Episode Statistics (HES) database. The algorithm produces risk scores for each admitted patient and a prediction of people in risk of readmission within 30 days. Also, 20 different risk bands were defined and associated to cost estimates for the business case analysis.

3.2 Risk Adjustment

Although the medical advances have contributed to improvement of life expectancy, it has little to do with the life expectancy and has much more to do with the quality of life. Risk Adjustment methods are either used for direct selection by health insurer for selection of good (profitable) risks from an insurer pool, or indirectly by designing insurance products. The models are often based on a linear utility function framework [48], and the objective is to minimise the outcome (risk) [15, 16]. The data mining techniques in Risk Adjustment modelling help to find the characteristics that have predictability power, which do not result into unfair risk assessment.

4 Techniques

Five major modelling methodologies presented in the previous studies on PM in healthcare are: simulation, formula-based methods, statistical, probabilistic and queuing. The focus of this study is on Dynamic Bayesian Network (DBN) techniques in PM, therefore initially regression modelling approach is briefly summarised. Afterwards, a brief introduction to the ML methods is provided. Then, the graphical networks with the main focus on DBN modelling are reviewed.

4.1 Regression

Regression modelling methods, such as logistic regression and mixed models, have been applied extensively in previous literature in social science and healthcare modelling [24, 39, 2].

An application of regression modelling is in the pathway modelling of the End-of-Life and frailty (i.e. the factors that arise from heterogeneity amongst patients) function modelling. A multinomial logit model was developed by Adeyemi et al. [1] for modelling patient's pathway. In the model, the patient frailties were regarded as mixed effect, and the random effects distributions were modelled based on patient pathways. The model could identify the high probability pathways for survival and cost objective functions.

An evaluation of multiple logistic regression-based PMs was performed by Aylin et al. [4] for inpatient death using the HES database. Different number of factors, such as age, sex, admission parameters and diagnostics, were included, and ultimately the performance was compared by using Receiver-Operating Characteristic (ROC). The models were designed for very specific mortality risk index in healthcare, but it was demonstrated that the use of only an administrative database could effectively predict the risk.

4.2 Machine Learning

Since late 1980s, ML methods have been used in extending the statistical analysis for making inferences from data. There are a lot to be done in the area of automated methods for learning and forecasting in healthcare. In this research, the major area of contribution is in developing techniques for doing transfer learning. The transfer learning refers to the methods that harness and adapt models to a specific new predictive task at hand. Transfer learning methodologies can help to use forecasting and PM to provide a systematic methodology of analysis for similar cases with smaller number of visible parameters. This may also be extended to perform active learning for use in complex real-world settings [34, 30].

Based on the knowledge of interest, Bayesian Networks (BNs), Neural Networks, Decision Tree and Kernel methods, such as Support Vector Machines, and Gaussian Processes are often used in healthcare data mining. Other approaches in ML can be found in the work of Bishop and Nasrabadi [7]. In the following, BNs which is a subset of graphical networks is discussed.

4.2.1 Bayesian Networks

There are two main approaches in incorporating stochastic models to the statistical modelling: discriminative (conditional distribution model) and generative (joint probability model). The discriminative modelling does not make any assumption about the prior distribution and only includes the conditional probability. Therefore, discriminative modelling is also known as the frequentist approach, and a linear classifier and logistic regression are examples of it. On the other hand, Bayes modelling methods are known as generative and they include the prior (marginal) distribution of the evidence data to the discriminative model [49].

Graphical networks are commonly used in probabilistic and statistical modelling. BNs [54] are standard approaches for modelling structured domains by allowing explicit representation of dependencies. BNs use probabilistic inference methods to present parameters with high dimensional distributions. These parameters are presented as nodes in graphical networks, where links between nodes represent direct correlation. To be able to update and infer large networks' probabilities in real-time, inference approximation methods, such as Expectation Maximisation and Particle Filter, have been introduced to estimate the joint and marginal probabilities of the nodes.

In BN modelling, template-based representations are used to produce a single compact model that can represent properties of system dynamics and to produce distribution over the different trajectories (DBN modelling) or to produce a distribution over different worlds (e.g. Genetics networks). To be able to reason about non-static situations, DBNs [17] are used to represent nodes with system-states. The system-states are either considered as stationary time-slices (homogeneous or invariant) like Markov Models or as the state observation model like Hidden Markov Models (HMMs). In state observation models, the states are variant and evolve on their own separately from the observations [37].

There are always two major challenges in using BN modelling: design and training. The accuracy and the efficiency of BN models depend on four main design choices: the framework of the causal tree, the framework of the system-states, inference approximation algorithm and finally the assignment and update method of prior probabilities.

It is not feasible computationally to design a graphical network that is too large, dynamic, inhomogeneous, noisy and incomplete. Therefore, the modelling assumptions are often relaxed, and specialised approximation and heuristic methods are widely used. In the next section, five main Bayesian Network (BN) techniques in PM of dynamic systems are outlined.

5 Time-Varying Dynamic Bayesian Networks Approaches

In a PM problem, such as Length-of-Stay or End-of-Life, the features are mainly inhomogeneous, because the processes in the models are either non-stationary (e.g. length of illness or treatment stages) or the events are sparse (e.g. morbidity conditions or patient states). The unobservable and inhomogeneous properties of models cause the momentum of the system dynamic to change across temporal access.

Generally, it is not statistically tractable to consider all of the variances for every time-point, because of the complexity in the inference and the lack of enough training evidence. Also, it is not possible to segment the time, since the model characteristics are unknown for each segment. There are five main approaches to model a time-varying Dynamic Bayesian Network (DBN). The approaches are highlighted in the following and the summary of the studies is presented in Table 1.

Firstly, a basic indirect approach is to transform time in order to make the process homogeneous. A naive approach is to use a time interpolation technique. Instead of using direct time transformation, a method like Kalman filter can be used, which is a Linear Dynamical System technique and is based on an autoregressive function to estimate a value at a timepoint [62, 13]. Xu et al. [63] proposed a state space model based on Kalman filter to estimate mean and variance for equally and unequally spaced longitudinal count data with serial correlation. The model applied to Epileptic Seizure and Primary Care Visits Data, and with high number of observations, the model produced comparable results to those by numerical approach.

Moreover, another indirect approach is to re-weight the likelihoods at each time-point using a particle based approach, such as feed-forward and sparse Kalman filtering. Since DBN is a generative model, it often works better for sparse models, because of its assumption about the underlying probabilities. However, it needs to be applied with an extreme care, since inappropriate sampling technique can rapidly slide the weights to zero or make the model assumptions and prior probabilities incorrect [37].

Furthermore, another approach is network rewiring which is also known as time-evolving graphical models [57]. It is a feasible option for large-scale time-varying networks. Time-evolving graphical models have been recently used in designing large networks in biological and social studies [3, 64, 31], with the objective to find unobserved network topologies or to rewire network under different conditions (e.g. edge-stability, and transitivity). For instance, recently a new modelling approach known as temporal Exponential Random Graph Model has been proposed for modelling networks evolving over discrete time-steps with Monte Carlo Markov Chain based or convex optimisation algorithms for posterior inference [3, 31, 32].

Another approach is conditional BN modelling, which is also known as multilevel or hierarchical BN model and is popular in the literature. In BN modelling a time-varying framework on top of a Markov Chain (MC) technique can be used to model multilevel time properties. A Cox phase-type model is used for modelling durations on top of a Hidden Semi-Markov Model by [20] for human activity recognition. In this research, an extension added to the Coxian Hidden Semi-Markov Model, which incorporates both duration and hierarchical modelling. Also, a DBN framework is proposed by Lappenschaar et al. [41] for modelling non-stationary events in multi-morbidity modelling. This PM of the interactions between heart failure and diabetes mellitus could closely resemble the PM techniques which use multilevel Linear Mixed Model. Moreover, a framework is designed by Lappenschaar et al. [40] for formulating an Linear Mixed Model into a BN using a Logistic Regression function.

Finally, the Linear Dynamical Systems are useful temporal models, which represent one or more real-valued variables that evolve linearly over time, with some Gaussian noise [37]. There are two categories of the Linear Dynamical Systems methods in modelling of time-varying DBN: Switching Linear Dynamic System and Time Varying Autoregression.

Switching Linear Dynamic Systems have been studied extensively for piecewise modelling of linear systems [61]. Based on the Switching Linear Dynamic Systems modelling, time-varying observations [28] and time-varying duration [8, 53] can be formulated using a latent MC. But, the MC method which is a piecewise stationary, does not have a very general application in learning and inference, and a time-varying linear regression can be used instead. For instance, a time-varying DBN has been introduced by Song et al. [59], which aggregates observations of adjacent time points by a kernel re-weighting function.

Time Varying Autoregression models are another type of the Linear Dynamic Systems [61] models, which focus on non-stationary models with fixed structure. Time Varying Autoregression models have been applied on a wide range of research applications, such as PM of equity market [35], inferring time-varying data from gene expression [56, 55] and modelling non-Gaussian autoregression [26].

6 Gaps and Limitations

Presented solutions in the literature mainly lack robustness and re-usability in different environments and for different care plans. These approaches mainly use descriptive, regression or Markov methods. The main shortcomings of these models can be summarised in the following:

- Not being fit for the modelling of complex correlations;
- Not accounting for small probabilities in an appropriate way;
- Not updating the beliefs (prior probabilities) based on the environmental variables and change of policies;
- Not using an automated parallel workflow system to compare different models and settings [50, 22].

Moreover, after the breakdown of financial markets at 2008, Rodriguez [58] wrote, "PM, the process by which a model is created or chosen to try to best predict the probability of an outcome has lost credibility as a forecasting tool". This is due to either the modeller's expertise or knowledge, or the lack of resources. The main common reasons that make a PM model to fail are as below:

- Inadequate pre-processing of the data;
- Inadequate model validation;
- Unjustified extrapolation;
- Over-fitting the model to available data [38].

Table 1 The summary of the studies in modelling time-varying processes.

Approach	**Method(s)**	**Study**	**Domain**	**Findings/Outcomes**
Time transformation	Auto-regression and Kalman filter	Xu et al. [63]	Healthcare events	The likelihood estimation approach performs better than the numerical integration approach
Time Evolving Graphical Network	Prevailing networks	Robinson et al. [57]	Generic non-stationary data	Demonstrating the feasibility
	Scalable inference for time-evolving networks	Ahmed et al. [3]	Biological systems to social science	Having asymptotically value-consistent under fixed model dimension
Conditional BN Modelling	A multilevel BN	Lappenschaar et al. [40]	Multimorbidity condition prediction	Providing more insight into interaction of multiple diseases
	A multilevel BN	Lappenschaar et al. [41]	The course of a medical condition	An informative clinical decision making tool
	Semi-Markov model, Coxian and HMM	Duong et al. [20]	Recognition of human activities of daily living	Having high and comparable accuracy
Switching Linear Dynamic Systems	Using hidden variables for network changes	Wang et al. [61]	Camera tracking	Being successful for both simulated non-stationary data and video sequences
Time Varying Autoregression	Kalman filter	Johnson et al. [35]	Equity market prediction	Performing as well as the Capital Asset Pricing Model benchmark, despite of using non-traditional pricing measures

The PM of Operational Risks in healthcare modelling, such as Length-of-Stay and End-of-Life, varies across systems and often are not robust. There is a few examples in the literature that uses ML techniques including BNs to address Length-of-Stay and End-of-Life problems.

7 Challenges

There are a number of challenges in applying ML techniques in healthcare modelling. Firstly, there are uncertainties in patient flows and resource demands as well as complex interdependency relations in the healthcare system.

Secondly, there are parameters, associations and timeframes that are missing from the primary and the secondary database due to security and confidentiality concerns. Thirdly, understanding the parameters, dependencies and independences are very crucial. Fourthly, there are various challenges that lie ahead of the robustness and generalisation of the solution, such as supporting adequate training data and selection of performance measures [23, 4].

Finally, it is very common in PM to build, evaluate and compare different models with varying features, algorithms, parameters and cohorts. Therefore, use of an optimised and automated parallel workflow system, like the framework developed by Ng et al. [50] is desirable to facilitate large-scale modelling.

8 Future Research Directions

Available solutions and proposed methodologies in the literature are mainly focused on addressing healthcare problems, such as PM of Length-of-Stay and End-of-Life, using regression methods and Markov models on very small number of case studies. The robustness, accuracy and re-usability of the models can be improved using ML methods.

ML methods, such as BNs, can represent complex correlation models and include small probabilities into the solution. BN methodologies allow composing risk utility functions and the causal and dependency relationship between variables, events, risks and outcomes. This enables BN models to appropriately estimate small probabilities that are associated to healthcare system as well as the statistical evidence [22]. There has been a little research done on using BNs in healthcare modelling, and the time-varying DBN techniques have great potentials in PM of non-stationary and hierarchical events and risks in healthcare problems.

References

1 Shola Adeyemi and Thierry J Chaussalet. Models for extracting information on patient pathways. In *Intelligent Patient Management*, pages 171–182. Springer, 2009.

2 Shola Adeyemi, Eren Demir, and Thierry Chaussalet. Towards an evidence-based decision making healthcare system management: Modelling patient pathways to improve clinical outcomes. *Decision Support Systems*, 55, 1:117–125, 2013.

3 Amr Ahmed and Eric P Xing. Recovering time-varying networks of dependencies in social and biological studies. *Proceedings of the National Academy of Sciences*, 106(29):11878–11883, 2009.

4 Paul Aylin, Alex Bottle, and Azeem Majeed. Use of administrative data or clinical databases as predictors of risk of death in hospital: comparison of models. *BMJ: British Medical Journal*, 334(7602):1044, 2007.

5 Martin Bardsley, Theo Georghiou, Ludovic Chassin, Geraint Lewis, Adam Steventon, and Jennifer Dixon. Overlap of hospital use and social care in older people in england. *Journal of health services research and policy*, 17(3):133–139, 2012.

6 John Billings, Ian Blunt, Adam Steventon, Theo Georghiou, Geraint Lewis, and Martin Bardsley. Development of a predictive model to identify inpatients at risk of re-admission within 30 days of discharge (parr-30). *BMJ Open: e001667*, 2(4), 2012.

7 Christopher M Bishop and Nasser M Nasrabadi. *Pattern recognition and machine learning*, volume 1. springer New York, 2006.

8 A. Blake, B. North, and M. Isard. Learning multi-class dynamics. In *Proceedings of the 1998 Conference on Advances in Neural Information Processing Systems II*, pages 389–395, Cambridge, MA, USA, 1999. MIT Press.

9 Ian Blunt, Martin Bardsley, Amy Grove, and Aileen Clarke. Classifying emergency 30-day readmissions in england using routine hospital data 2004-2010: what is the scope for reduction? *Journal of Epidemiology and Community Health*, 66(Suppl 1):A45–A45, 2012.

10 Michael Caley and Khesh Sidhu. Estimating the future healthcare costs of an aging population in the uk: expansion of morbidity and the need for preventative care. *Journal of Public Health*, 33(1):117–122, 2011.

11 S Chattopadhyay, P Ray, HS Chen, MB Lee, and HC Chiang. Suicidal risk evaluation using a similarity-based classifier. In *Advanced Data Mining and Applications*, pages 51–61. Springer, 2008.

12 Xavier Chitnis, T Georghiou, A Steventon, and M Bardsley. The impact of the marie curie nursing service on place of death andhospital use at the end of

life. https://www.mariecurie.org.uk/Documents/HEALTHCARE-PROFESSIONALS/Our%20impact/Marie%20Curie_Full%20Report_Final_Web.pdf, 2012.

13 Richard J Cook and Jerald F Lawless. Statistical issues in modeling chronic disease in cohort studies. *Statistics in Biosciences*, 6:127–161, 2013.

14 Chiara Cornalba. Clinical and operational risk: A bayesian approach. *Methodology and Computing in Applied Probability*, 11(1):47–63, 2009.

15 Anthony J. Culyer and Joseph P. Newhouse. *Handbook of Health Economics: Vol. 1A*. Elsevier, 2000.

16 Anthony J. Culyer, Mark V. Pauly, Joseph P. Newhouse, Thomas G. McGuire, and Pedro P. Barros. *Handbook of Health Economics: Vol. 2*. Elsevier, 2012.

17 Thomas Dean and Keiji Kanazawa. A model for reasoning about persistence and causation. *Computational intelligence*, 5(2):142–150, 1989.

18 DoH. Long term conditions compendium of information (third edition). https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/216528/dh_134486.pdf, Apr. 2012.

19 DoH. Business case: for the health and care modernisation transition programme. https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/224140/TP_business_case_docx.pdf, 2013.

20 Thi Duong, Dinh Phung, Hung Bui, and Svetha Venkatesh. Efficient duration and hierarchical modeling for human activity recognition. *Artificial Intelligence*, 173(7):830–856, 2009.

21 Shahram Ebadollahi, Jimeng Sun, David Gotz, Jianying Hu, Daby Sow, and Chalapathy Neti. Predicting patient's trajectory of physiological data using temporal trends in similar patients: A system for near-term prognostics. In *AMIA Annual Symposium Proceedings*, volume 2010, page 192. American Medical Informatics Association, 2010.

22 Norman E Fenton and Martin D Neil. *Risk Assessment and Decision Analysis with Bayesian Networks*. CRC Press, 2012.

23 Jonathan Agner Forsberg, Rikard Wedin, Bjarne H Bauer, Henrik CF andHansen, Minna Laitinen, Clement S Trovik, Johnny Keller, Patrick J Boland, and John H Healey. External validation of the bayesian estimated tools for survival (bets) models in patients with surgically treated skeletal metastases. *BMC cancer*, 12(1):493, 2012.

24 G David Garson. *Hierarchical linear modeling: Guide and applications*. Sage, 2012.

25 Seymour Geisser. *Predictive interference: an introduction*, volume 55. CRC Press, 1993.

26 Deniz Gencaga, Ercan E. Kuruoglu, and Aysin Ertuzun. Modeling non-gaussian time-varying vector autoregressive processes by particle filtering. *Multidimensional Systems and Signal Processing*, 21(1):73–85, 2010.

27 Theo. Georghiou, Sian Davies, Alisha Davies, and Martin Bradley. Understanding patterns of health and social care at the end of life. http://www.nuffieldtrust.org.uk/sites/files/nuffield/121016_understanding_patterns_of_health_and_social_care_full_report_final.pdf, Oct. 2012.

28 Zoubin Ghahramani and Geoffrey E Hinton. Variational learning for switching state-space models. *Neural computation*, 12(4):831–864, 2000.

29 David Gotz, Jimeng Sun, Nan Cao, and Shahram Ebadollahi. Visual cluster analysis in support of clinical decision intelligence. In *AMIA Annual Symposium Proceedings*, volume 2011, pages 481–490. American Medical Informatics Association, 2011.

30 Susan Graham, Deborah Estrin, Eric Horvitz, Isaac Kohane, Elizabeth Mynatt, and Ida Sim. Information technology research challenges for healthcare: From discovery to delivery. *ACM SIGHIT Record*, 1(1):4–9, 2011.

31 Fan Guo, Steve Hanneke, Wenjie Fu, and Eric P Xing. Recovering temporally rewiring networks: A model-based approach. In *Proceedings of the 24th international conference on Machine learning*, pages 321–328. ACM, 2007.

32 Steve Hanneke, Wenjie Fu, Eric P Xing, et al. Discrete temporal models of social networks. *Electronic Journal of Statistics*, 4:585–605, 2010.

33 Dawn E Holmes and Lakhmi C Jain. *Data Mining: Foundations and Intelligent Paradigms: Volume 3: Medical, Health, Social, Biological and Other Applications*, volume 3. Springer, 2012.

34 Eric Horvitz. From data to predictions and decisions: Enabling evidence-based healthcare. `http://research.microsoft.com/pubs/141911/Evidence_based_healthcare_essay.pdf`, 2010.

35 Lorne D Johnson and Georgios Sakoulis. Maximizing equity market sector predictability in a bayesian time-varying parameter model. *Computational Statistics and Data Analysis*, 52(6):3083–3106, 2008.

36 Linda T Kohn, Janet M Corrigan, Molla S Donaldson, et al. *To err is human: building a safer health system*, volume 627. National Academies Press, 2000.

37 Daphne Koller and Nir Friedman. *Probabilistic Graphical Models: Principles and Techniques*. MIT Press, 2009.

38 Max Kuhn and Kjell Johnson. *Applied Predictive Modeling*. Springer, 2013.

39 Elena Kulinskaya, Diana Kornbrot, and Haiyan Gao. Length of stay as a performance indicator: robust statistical methodology. *IMA Journal of Management Mathematics*, 16(4):369–381, 2005.

40 Martijn Lappenschaar, Arjen Hommersom, Peter JF Lucas, Joep Lagro, and Stefan Visscher. Multilevel bayesian networks for the analysis of hierarchical health care data. *Artificial intelligence in medicine*, 57, 3:171–183, 2013.

41 Martijn Lappenschaar, Arjen Hommersom, Peter JF Lucas, Joep Lagro, Stefan Visscher, Joke C Korevaar, and François G Schellevis. Multilevel temporal bayesian networks can model longitudinal change in multimorbidity. *Journal of clinical epidemiology*, 66(12):1405–1416, 2013.

42 Charles Leadbeater and Jake Garber. Dying for change. `http://www.demos.co.uk/files/Dying_for_change_-_web_-_final_1_.pdf`, 2010.

43 Geraint Lewis, Natasha Curry, and Martin Bardsley. Choosing a predictive risk model: a guide for commissioners in england. `https://www.primis.nottingham.ac.uk/attachments/article/643/choosing_predictive_risk_model_guide_for_commissioners_nov11.pdf`, 2011.

44 Funding of Care and Support. Fairer care funding - the report of the commission on funding of care and support. `http://www.ilis.co.uk/uploaded_files/dilnott_report_the_future_of_funding_social_care_july_2011.pdf`, 2011.

45 NHS England. Nhs england publishes ccg funding allocations for next two years following adoption of new formula. `http://www.england.nhs.uk/2013/12/18/ccg-fund-allocs/`, 2013.

46 The King's Fund. Briefing: The care bill - second reading in the house of commons. `http://www.kingsfund.org.uk/sites/files/kf/field/field_publication_file/briefing-care-bill-house-of-commons-second-reading-kingsfund-dec13.pdf`, 2013.

47 The King's Fund. The health and social care system in 2025 - a view of the future. `http://www.kingsfund.org.uk/sites/files/kf/field/field_publication_file/The%20Health%20and%20Social%20Care%20System%20in%202025%20-%20supplementary%20note%20from%20The%20King's%20Fund.pdf`, 2013.

48 Joseph P Newhouse. Reimbursing health plans and health providers: efficiency in production versus selection. *Journal of economic literature*, 34(3):1236–1263, 1996.

49 Andrew Y. Ng and Michael I. Jordan. On discriminative vs. generative classifiers: A comparison of logistic regression and naive bayes. *Advances in neural information processing systems*, 14:841, 2002.

50 Kenney Ng, Amol Ghoting, Steven R Steinhubl, Walter F Stewart, Bradley Malin, and Jimeng Sun. Paramo:a parallel predictive modeling platform for healthcare analytic research using electronic health records. *Journal of Biomedical Informatics*, 2013.

51 NHS. Whole system partnership - end of life care engagement event. Presentation, Jul. 2012.

52 NHS. Improving end of life care through early recognition of need. `http://www.thewholesystem.co.uk/docs/eolc_predictive_modelling_report.pdf`, 2013.

53 Vladimir Pavlovic, James M Rehg, and John MacCormick. Learning switching linear models of human motion. In *NIPS*, pages 981–987. Citeseer, 2000.

54 Judea Pearl. *Probabilistic reasoning in intelligent systems: networks of plausible inference*. Morgan Kaufmann, 1988.

55 Bruno-Edouard Perrin, Liva Ralaivola, Aurelien Mazurie, Samuele Bottani, Jacques Mallet, and Florence d'Alche Buc. Gene networks inference using dynamic bayesian networks. *Bioinformatics*, 19(suppl 2):ii138–ii148, 2003.

56 Arvind Rao, Alfred O Hero III, James Douglas Engel, et al. Inferring time-varying network topologies from gene expression data. *EURASIP Journal on Bioinformatics and Systems Biology*, 2007:7–7, 2007.

57 Joshua W. Robinson and Alexander J. Hartemink. Learning non-stationary dynamic bayesian networks. *J. Mach. Learn. Res.*, 11:3647–3680, December 2010.

58 Manuel A. Rodriguez. The failure of predictive modeling and why we follow the herd. `http://www.cfclaw.com/files/r_201145133147.pdf`, 2011.

59 Le Song, Mladen Kolar, and Eric P Xing. Time-varying dynamic bayesian networks. In *NIPS*, pages 1732–1740, 2009.

60 Roger Stedman. Scoring systems and outcomes. In Fang Gao Smith, editor, *Core Topics in Critical Care Medicine*, pages 27–33. Cambridge University Press, 2010.

61 Zhaowen Wang, Ercan E Kuruoglu, Xiaokang Yang, Yi Xu, and Thomas S Huang. Time varying dynamic bayesian network for nonstationary events modeling and online inference. *Signal Processing, IEEE Transactions on*, 59(4):1553–1568, 2011.

62 Zhu Wang, Wayne A Woodward, and Henry L Gray. The application of the kalman filter to nonstationary time series through time deformation. *Journal of Time Series Analysis*, 30(5):559–574, 2009.

63 Stanley Xu, Richard H Jones, and Gary K Grunwald. Analysis of longitudinal count data with serial correlation. *Biometrical journal*, 49(3):416–428, 2007.

64 Shuheng Zhou, John Lafferty, and Larry Wasserman. Time varying undirected graphs. *arXiv preprint arXiv:0802.2758*, 2008.