Dynamic risk assessment in healthcare based on Bayesian approach

Li Mei[a,*] , Liu Zixian[a], Li Xiaopeng[a], Liu Yiliu[b]

[a] College of Management and Economics, Tianjin University, Tianjin, China

[b] Department of Mechanical and Industrial Engineering, Norwegian University of Science and Technology, Trondheim, Norway

*Corresponding author e-mail address: lm314835237@163.com

Abstract: Risks related with healthcare are always dynamic, and they are affected by situations of patients, human errors in treatment and even the states of medical devices. This paper proposes a dynamic medical risk assessment model, for capturing the impacts of factors on the occurrence of adverse events. In this model, a static fault tree is established to show risk scenarios. Dynamic Bayesian network and Bayesian inference are introduced to analyze the operations of medical devices, in consideration of their failures, repairs, and human errors over time. Hemodialysis infection is taken as the case to verify that the proposed method is helpful to demonstrate the changes of medical risks with time, and to identify the critical events contributing to the occurrence of the adverse event at different moments. These findings can act as the basis to assign and adjust safety measures.

Keywords: Medical risk; Dynamic risk assessment; Dynamic Bayesian network; Bayesian inference; Probability updating

1. Introduction

Medical risks refer to the probability of an adverse event during a particular treatment process, and they impact on patient safety and overall healthcare system, in

considering possible high medical costs, and wastage of medical resources. Medical risks can come from biological properties of patients, like age, gender and comorbidity, and they can be from those iatrogenic factors, such as device failures and human errors. Medical risks have received more attentions recently, after the publication of the report titled “To err is human: building a safer health system” by the Institute of Medicine (IOM) in 1999. The report has indicated that identification and control of iatrogenic factors are the most effective way of reducing medical risks [1]. When such iatrogenic risks can be revealed proactively, the critical events in a healthcare system are determined, safety measures can be implemented effectively.

1.1. Risk assessment in healthcare

Probabilistic risk analysis (PRA) has been introduced to assess medical risks. Among them, fault tree analysis (FTA) has been a common-accepted approach due to its ability for modeling risk scenario by illustrating relationships between an adverse event and the possible causes. Ekaette et al. have applied FTA in radiation treatment risk assessment [2], as well as Komal has evaluated risk of medication delivery and inpatient transfers with fuzzy FTA [3].

However, fault tree models have long been in the debate because of their weakness in representing the dependencies of events explicitly, analyzing multi-state variables, updating probabilities, and coping with uncertainties [4,5]. Bayesian network (BN) has gained more focuses recently, because it can release aforementioned limitations to some extent [6,7], with the strong capability in predictive and diagnostic analysis. Maglogiannis et al have introduced BN in healthcare to assess the risk of medical

information system [8]. Nevertheless, conventional BN can only represent relationships between variables at a particular time point or during a specific period of time, and it does not reflect temporal relationships between different times.

1.2. The dynamic nature of medical risk

It is noticeable that medical risks always evolve with time due to the dynamic natures of their impact factors in a healthcare process [9]. Here we consider medical workers and medical devices:

Behaviors of humans are full of uncertainties, and the probabilities of human errors are changing due to the ever-changing environment. It is reported, however, that probabilities of human errors are usually quantified as constant values while keeping same at different times [10,11]. However, it is possible to know the number of occurrences of a human error during a certain period of time, which can be used to conduct probability updating by Bayesian inference. Bayesian inference provides a mathematical framework for updating prior knowledge by integrating new data and information.

On the other hand, the performance of a medical device, affected by its age and usage intensity, can degrade over time. Compared with human errors, medical devices are of high reliability and failures seldom occur [12]. Therefore, Bayesian inference is not suitable for probability updating of device failures. When the device has a failure, a repair is needed to restore the failure. To model such behaviors, dynamic Bayesian network (DBN) can be introduced. This method is able to demonstrate the changes over time and relationships between the current state of a device and its past or future states

[13]. To our knowledge, there have been no specific studies of dynamic Bayesian network and Bayesian inference in risk assessment for device failures and human errors in healthcare.

Based on a fault tree to model medical risk scenario, the purpose of this study is to embed Bayesian approaches into the fault tree for dynamic medical risk assessment. Hemodialysis infection is involved as a case study due to its high incidence, and it is the main cause leading to the death of hemodialysis patients [14,15]. The proposed method includes two parts: (1) the qualitative and static analysis where risk identification of hemodialysis infection is carried out with the help of conventional FTA, (2) the quantitative and dynamic analysis in which probability updating is conducted using the dynamic Bayesian network and Bayesian inference to demonstrate risk profile over time. Additionally, critical events contributing to the occurrence of hemodialysis infection are identified using the ratio of variation (RoV). The rest of the paper is structured as follows: section 2 gives a brief introduction to dynamic Bayesian network and Bayesian inference; section 3 demonstrates the dynamic risk assessment model; section 4 is the case study followed by conclusions in section 5.

2. Dynamic Bayesian network

A Bayesian network consists of qualitative and quantitative parts. The qualitative part is a directed acyclic graph, including a set of nodes representing the system variables and a set of directed arcs representing the dependencies or the cause-effect relationships among the variables. The quantitative part is conditional probability tables (CPTs), which stand for conditional dependencies between nodes and their parents [16].

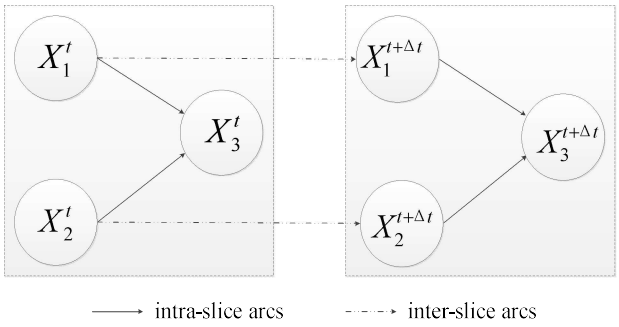


Fig. 1 A simple dynamic Bayesian network

As an extension of ordinary BN, DBN facilitates explicit modeling of temporal evolution for a set of random variables over a discretized timeline [13]. Assume that a timeline T is divided into several time-slices and the time interval is Δt. Each time-slice stands for a unit time which can be one day, one month or one year. So, a DBN can be defined as a pair $(B^1, B^{\rightarrow})$ over the timeline, where B^1 is a static BN at t=0, and $B^{\rightarrow}$ is a two-slice temporal BN which defines:

$$P(X^{t+\Delta t}|X^t) = \prod_{i=1}^{n} P\left(X_i^{t+\Delta t}|Pa(X_i^{t+\Delta t})\right). \quad (1)$$

Then, the joint probability distribution of the variables at time t can be represented as:

$$P(U^{t+\Delta t}) = P\left(X_1^{t+\Delta t}, X_2^{t+\Delta t}, \cdots, X_n^{t+\Delta t}\right) = \prod_{i=1}^{n} P\left(X_i^{t+\Delta t}|X_i^t, Pa(X_i^t), Pa(X_i^{t+\Delta t})\right), \quad (2)$$

where X_i^t and $X_i^{t+\Delta t}$ are different states of X_i in two consecutive time-slices, $Pa(X_i^t)$ and $Pa(X_i^{t+\Delta t})$ are parent sets of X_i at these two time-slices, respectively.

A simple DBN is shown in Fig. 1, where a similar BN is duplicated twice with variables of different values. Moreover, the two types of arcs show the relationships between variables, including intra-slice arcs which represent the relationships in a time-slice (denoted by solid arrows) and inter-slice arcs, representing the relationships between successive time-slices (denoted by dotted arrows).

DBNs can update the values of probability with Bayesian inference, which is a

key element to dynamic risk assessment [17]. Considering θ as a parameter, and $\pi_0(\theta)$ as the probability distribution function (prior distribution) of θ. New data about θ is used to form the likelihood function $f(x|\theta)$. Then, Bayesian inference can be employed to yield the posterior distribution $\pi_1(\theta|x)$ as shown in the following equation [18]:

$$\pi_1(\theta|x) \propto f(x|\theta)\pi_0(\theta). \quad (3)$$

3. Dynamic risk assessment model for hemodialysis infection

In this section, the qualitative part of our proposed model is described first, in which an adverse event and its possible causes are identified using conventional FTA. Then, the quantitative part introduces dynamic Bayesian network and Bayesian inference for probability updating of device failures and human errors, respectively. After that, the occurrence probability of the adverse event can be evaluated and the critical events will be identified. We take the hemodialysis infection as a case here to describe how the analysis should be conducted for healthcare activities.

Hemodialysis removes metabolic wastes and supplements bases through diffusion between blood and dialysate. With the continuous development and extensive applications of the technology, hemodialysis has been proved to be one of the most effective treatments for patients with end-stage renal disease (ESRD), which can increase the life span of patients. Even so, hemodialysis has a high level of risk due to a variety of influencing factors in the process. The process of hemodialysis is illustrated in Fig. 2 [19,20]. Specifically, nurses cannulate the vascular access, and connect them to electronically controlled dialysis machines. Sophisticated equipment purifies the

water for blending dialysate. Dialyzers are reprocessed and sterilized before they are delivered to the nurse for setting up the dialysis equipment. Nurses should take care of patients responsibly during the hemodialysis, but hazardous events are inclined to occur in so many interactions between dialysis staff, machines, and the environment [21]. Recently, the number of ESRD patients who receive hemodialysis is rapidly increasing because of the aging population and prevalence of diabetes and hypertension [22,23]. Due to sharing of dialysis machines and close contact among patients in dialysis units, the rate of infections acquisition of hemodialysis patients is significantly higher than general population [15,24]. Consequently, risk assessment for hemodialysis infection is to be urgently settled before serving the patients.

3.1. Fault tree analysis for hemodialysis infection

Fault tree analysis is a diagram which connect potential adverse event (that is a top event) with its possible causes (intermediate events and basic events) using Boolean logic where "AND" and "OR" are two basic types. Here we describe the "hemodialysis infection" as the top event (TOP) and causes of the top event mainly come from three stages of hemodialysis as shown in Fig. 2. Thus, "infection in preparation (G_1)", "infection during hemodialysis (G_2)" and "infection after hemodialysis (G_3)" are connected with "hemodialysis infection (TOP)" using an "OR" gate. According to expert suggestions and related literature, 23 possible basic events in the hemodialysis process are identified. The whole fault tree is constructed accordingly and shown in Fig. 3 [25,26,27]. The descriptions of symbols in the fault tree are listed in Table 1.

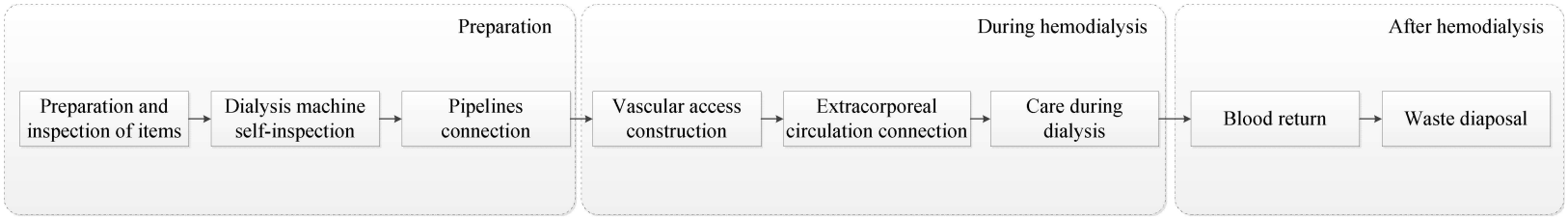


Fig. 2 The process of hemodialysis

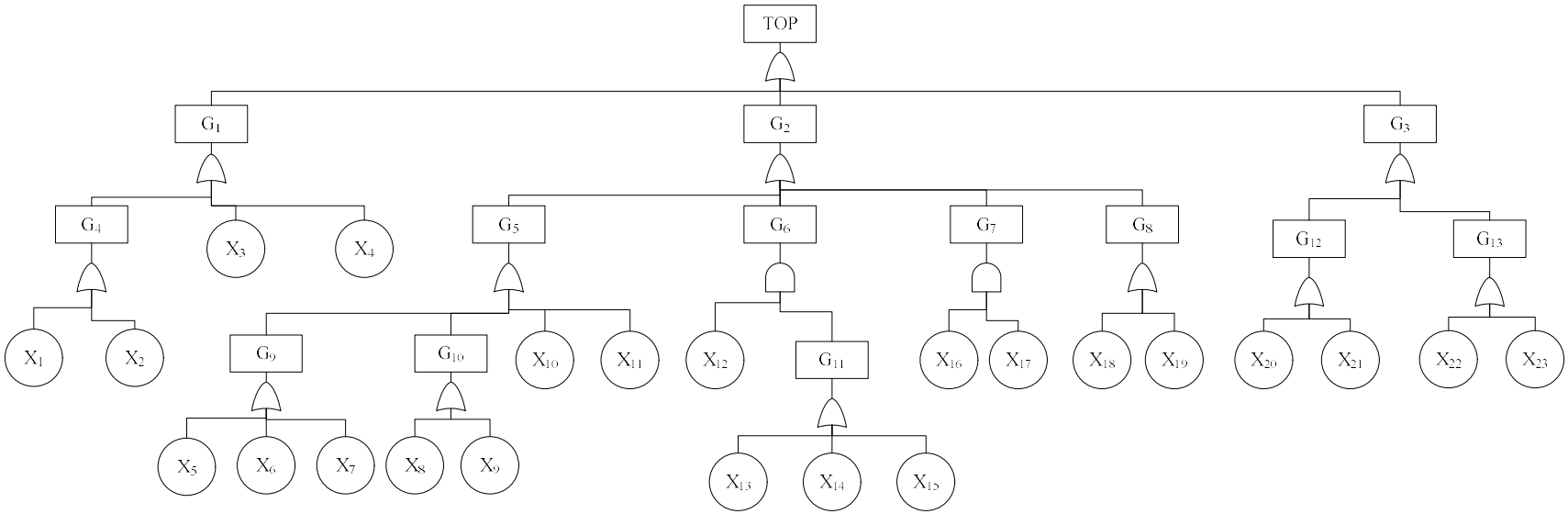


Fig. 3 Fault tree for hemodialysis infection

Table 1 Descriptions of events in fault tree for hemodialysis

Symbol	Description	Symbol	Description
TOP	Hemodialysis Infection	G_7	Arteriovenous pressure sensor failure
G_1	Infection in preparation	G_8	Improper care during hemodialysis
G_2	Infection during hemodialysis	G_9	Pretreatment system failure
G_3	Infection after hemodialysis	G_{10}	Reverse osmosis machine failure
G_4	Improper puncture	G_{11}	Blood leakage alarm failure
G_5	Water treatment system failure	G_{12}	Improper blood return
G_6	Dialyzer failure	G_{13}	Improper waste disposal
X_1	Puncture needle contamination	X_{13}	Sediment in blood leakage
X_2	Error in puncture	X_{14}	Detector contamination
X_3	Catheter damage	X_{15}	Exhaust failure
X_4	Catheter contamination	X_{16}	Sensor congestion
X_5	Multi-media filter failure	X_{17}	Sensor alarm failure
X_6	Carbon filter failure	X_{18}	Error in puncture point care
X_7	Demineralizer failure	X_{19}	Error in catheter care
X_8	Reverse osmosis membrane contamination	X_{20}	Error in fistula assessment
X_9	Reverse osmosis membrane damage	X_{21}	Error in blood return
X_{10}	Insufficient water disinfection	X_{22}	Error in closed drain
X_{11}	Insufficient water monitoring	X_{23}	Insufficient disinfection
X_{12}	Dialysis membrane damage		

3.2. Device failure probability updating with DBN

DBN is employed to show the state transition of a device at different times, and to calculate its corresponding failure probability. For a device i, X_i^t represents the failure probability at time *t*. The dynamic Bayesian network for i is depicted in Fig. 4.

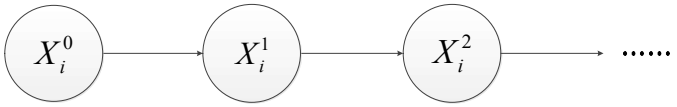


Fig. 4 Dynamic Bayesian network for medical device i

FTA assumes two states for each event, meaning that a device is either in its perfect functioning state (X_i=0) or failed state (X_i=1). In fact, it is a gradual process from X_i=0 to X_i=1 where many intermediate states exist between these two extreme states. For example, dialyzer can still normally operate for a period, even if it is a little wear out. The current paper assumes that for device failure events, each basic event has three states, namely X_i=1, X_i=0.5, and X_i=0, whereas every intermediate event still has two states in order to simplify calculation. The state X_i=0.5 refers to a degraded state between state X_i=0 and X_i=1.

At the beginning of a timeline (t=0), each device is in its state X_i=0. With the passage of time, it can either go to the degraded state (X_i=0.5) or the failed state (X_i=1), while all failure rates are assumed to be exponentially distributed. When a failure occurs (X_i=1), the device will no longer function continually, and repair is required. The state after the repair depends on the degree of repair that can be a perfect repair or an imperfect repair. The device can go to state "0" with perfect repair, or it can go to state "0.5" as well as "0" when performing imperfect repair. The repair rates also follow exponential distribution. The state transition of a device in dynamic Bayesian network

is shown in Fig. 5. The above symbols of the arcs are indicated as failure rates λ_i $(i = 1,2,3)$ and repair rates μ_j $(j = 1,2)$ between states, respectively.

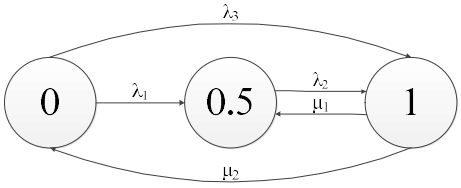


Fig. 5 State transition diagram for medical device

Assume that the current time is t, and the time interval is Δt, then the transition relationships of medical devices between two consecutive time-slices without repairs, with imperfect repairs, and with perfect repairs are given in Table 2-4, respectively [28]. Notably, the selection of Δt deserves consideration. For a given time period, risk variations may not be apparent under short Δt while a long Δt is prone to overlook some crucial information. Thus, a reasonable Δt should be determined carefully considering factors such as assessment cycle, observation period etc.

Table 2 Transition relationships between consecutive time-slices without repairs

t	$t+\Delta t$		
	X_i=0	X_i=0.5	X_i=1
X_i=0	$e^{-(\lambda_1+\lambda_3)\Delta t}$	$\frac{\lambda_1}{\lambda_1+\lambda_3}(1-e^{-(\lambda_1+\lambda_3)\Delta t})$	$\frac{\lambda_3}{\lambda_1+\lambda_3}(1-e^{-(\lambda_1+\lambda_3)\Delta t})$
X_i=0.5	0	$e^{-\lambda_2\Delta t}$	$1-e^{-\lambda_2\Delta t}$
X_i=1	0	0	1

Table 3 Transition relationships between consecutive time-slices with imperfect repairs

t	$t+\Delta t$		
	X_i=0	X_i=0.5	X_i=1
X_i=0	$e^{-(\lambda_1+\lambda_3)\Delta t}$	$\frac{\lambda_1}{\lambda_1+\lambda_3}(1-e^{-(\lambda_1+\lambda_3)\Delta t})$	$\frac{\lambda_3}{\lambda_1+\lambda_3}(1-e^{-(\lambda_1+\lambda_3)\Delta t})$
X_i=0.5	0	$e^{-\lambda_2\Delta t}$	$1-e^{-\lambda_2\Delta t}$
X_i=1	$\frac{\mu_2}{\mu_1+\mu_2}(1-e^{-(\mu_1+\mu_2)\Delta t})$	$\frac{\mu_1}{\mu_1+\mu_2}(1-e^{-(\mu_1+\mu_2)\Delta t})$	$e^{-(\mu_1+\mu_2)\Delta t}$

Table 4 Transition relationships between consecutive time-slices with perfect repairs

t	$t+\Delta t$		
	X_i=0	X_i=0.5	X_i=1
X_i=0	$e^{-(\lambda_1+\lambda_3)\Delta t}$	$\frac{\lambda_1}{\lambda_1+\lambda_3}(1-e^{-(\lambda_1+\lambda_3)\Delta t})$	$\frac{\lambda_3}{\lambda_1+\lambda_3}(1-e^{-(\lambda_1+\lambda_3)\Delta t})$
X_i=0.5	0	$e^{-\lambda_2\Delta t}$	$1-e^{-\lambda_2\Delta t}$
X_i=1	$1-e^{-(\mu_1+\mu_2)\Delta t}$	0	$e^{-(\mu_1+\mu_2)\Delta t}$

3.3. Human error probability updating with Bayesian inference

Bayesian inference is adopted to handle the dynamics of human errors. All human error events, both basic and intermediate, are assumed to involve two states since a human error just occurs or not. The difficulty of Bayesian inference is forming the likelihood function $f(x|\theta)$ using new data. Accident precursor data (APD) has been widely used to construct likelihood function which in turn updates the prior knowledge to yield the posterior distribution. APD is defined as the number of occurrences of the events which are not characterized as adverse events but indicate the increasing occurrence probability of an adverse event [29,30]. In healthcare, the number of occurrences of a human error can be regarded as APD. Considering P_i as the prior probability of a human error X_i, which needs to be updated, f is the number of occurrences of X_i in n patients during a time interval, then f follows a binomial distribution as follows [31]:

$$P(X_i = f) = f(APD|P_i) = P_i^f(1-P_i)^{n-f}. \quad (4)$$

To simplify calculation, it is common to choose a prior distribution and its corresponding likelihood function from the well-known conjugate families that in turn result in a posterior distribution from the same family [32]. Actually, the conjugate

distribution of binomial distribution is Beta distribution, thus this paper assumes that the P_i of X_i follows the $\mathrm{Beta}(\alpha, \beta)$ distribution [18, 31]:

$$f(P_i) \propto P_i^{\alpha-1}(1-P_i)^{\beta-1}, \tag{5}$$

where α and β are parameters of Beta distribution. In practice, α and β can be regarded as the numbers of occurrences and non-occurrences of X_i, respectively. Apparently, $\alpha+\beta$ is the number of served patients in the hospital in a time interval. Then, the posterior distribution of P_i can be obtained based on equation (3):

$$\begin{aligned} f(P_i|APD) &\propto f(APD|P_i)f(P_i) \propto P_i^{f}(1-P_i)^{n-f}P_i^{\alpha-1}(1-P_i)^{\beta-1} \\ &\propto P_i^{\alpha+f-1}(1-P_i)^{\beta+n-f-1}. \end{aligned} \tag{6}$$

Obviously, the posterior distribution of P_i is also a Beta distribution and the updated parameters are $\alpha+f$ and $\beta+n-f$, respectively.

In this paper, the value of P_i is considered as the mean value of Beta distribution. So, the prior probability of X_i can be presented as:

$$P_i = \frac{\alpha}{\alpha+\beta}, \tag{7}$$

and the posterior probability of X_i is adapted as:

$$P(X_i|APD) = \frac{\alpha+f}{(\alpha+f)+(\beta+n-f)} = \frac{\alpha+f}{\alpha+\beta+n}. \tag{8}$$

3.4. Dynamic risk profile for hemodialysis infection

When involving three-state device failures, conventional minimal cut sets methods in FTA to compute the probability of the top event become ineffective. Noisy OR-gate and AND-gate models can be introduced to handle corresponding “OR” gate and “AND” gate in FTA, and help to assign conditional probabilities for those intermediated events [33]. Let Y be an intermediate event, and $X_1, X_2, \cdots, X_n$ $(i =$

$1,2,\cdots,n)$ be the lower events of Y. Each X_i with three states is associated with a weight w_i which means X_i can cause Y with the probability of w_i. For noisy OR-gate, the conditional probabilities can be:

$$\mathrm{P}(Y|X_1,X_2,\cdots,X_n) = 1 - \prod\nolimits_{1\le i\le n}(1-w_i). \tag{9}$$

For noisy AND-gate, the conditional probabilities can be:

$$\mathrm{P}(Y|X_1,X_2,\cdots,X_n) = \prod\nolimits_{1\le i\le n} w_i. \tag{10}$$

Then, the probability of Y is obtained using the following equation:

$$\mathrm{P}_{noisy} = \prod\nolimits_{i=1}^{n} P(X_i)P(Y|X_i). \tag{11}$$

The weights for device failures are listed in Table 5. These values are estimated according to the expert suggestions, and the weights for state "0.5" are assigned as half of the values for state "1". For example, G_{10} in Fig. 3 is regarded as a noisy OR-gate and the conditional probabilities for G_{10} is computed using equation (9) (see Table 6).

Table 5 Weights for device failures

Symbol	Weight	Symbol	Weight
X_1	0.8	X_9	0.5
X_3	0.6	X_{12}	0.8
X_4	0.8	X_{13}	0.6
X_5	0.8	X_{14}	0.6
X_6	0.8	X_{15}	0.6
X_7	0.8	X_{16}	0.7
X_8	0.7	X_{17}	0.5

Table 6 Conditional probabilities for G_{10}

X_8 $(w_8 = 0.7)$	X_9 $(w_9 = 0.5)$	G_{10}
0	0	0
0.5	0	0.35

0	0.5	0.25
1	0	0.7
0	1	0.5
1	0.5	0.775
0.5	1	0.675
1	1	0.85

For other intermediate events with only two-state lower events, their probabilities can be computed as:

$$P_{AND} = \prod_{i=1}^{n} P(X_i), \tag{12}$$

$$P_{OR} = 1 - \prod_{i=1}^{n} \left(1 - P(X_i)\right). \tag{13}$$

Based on equation (11)-(13), the probability of the top event considering three-state events can be obtained by computing the probabilities of each logic gate from the bottom to the top.

Device failures and human errors can be prevented, or at least their occurrence probabilities can be reduced by adopting suitable safety measures. Considering the resources limitations, it is reasonable for healthcare to identify events with high criticalities, and assign corresponding measures onto these events. This paper measures the criticality of an event based on the ratio of variation (RoV), which takes both prior and posterior probabilities of an event into account and thus evaluates the criticality from a holistic perspective [34,35]. For a given basic event X_i, the posterior probability $\pi(X_i)$ can be computed as:

$$\pi(X_i) = \mathrm{P}(X_i|TOP) = \frac{\mathrm{P}(TOP|X_i)}{\mathrm{P}(TOP)}, \tag{14}$$

and RoV can be represented as:

$$\mathrm{RoV} = \frac{\pi(X_i) - \theta(X_i)}{\theta(X_i)}, \tag{15}$$

where $\theta(X_i)$ denotes the prior probability of X_i.

4. Case study

In the case study, data is collected from a large general hospital in Tianjin, China and used to estimate the prior parameters. Failure rates λ and repair rates μ of device failures following exponential distribution are listed in Table 7. Parameters α and β of human errors following Beta distribution are listed in Table 8. During data collection, we are allowed to go to the unit once a week for almost three months, so Δt in this study is selected as one week and all the results are within 10 weeks. More detailed data analysis and computational procedures can be found in the Appendix.

Table 7 Device failures in fault tree for hemodialysis infection

Symbol	Failure rate (λ, 10^{-5}/h)	Repair rate (μ, /h)	Prior probability (5^{th} week)	Posterior probability (10^{-3}, 5^{th} week)	Ratio of Variation (RoV, 5^{th} week)
X_1	2.324	2	0.0010	5.17	4.17
X_3	2.561	0.138	0.0010	8.01	6.28
X_4	4.487	0.97	0.0020	18.77	8.38
X_5	3.835	0.67	0.0017	15.96	8.39
X_6	3.007	0.55	0.0013	12.21	8.39
X_7	3.056	0.64	0.0013	12.21	9.39
X_8	3.107	1	0.0013	10.84	7.33
X_9	2.930	1	0.0013	8.10	5.23
X_{12}	3.741	1	0.0016	1.91	0.20
X_{13}	3.000	0.82	0.0013	1.35	0.04
X_{14}	3.160	0.83	0.0014	1.46	0.04
X_{15}	2.090	0.45	0.0009	0.94	0.04
X_{16}	2.163	0.12	0.0009	0.91	0.01
X_{17}	2.467	0.37	0.0011	1.11	0.007

Table 8 Human errors in fault tree for hemodialysis infection

Symbol	Parameters of Beta distribution		Prior probability	Posterior probability	Ratio of Variation
	α	β	(5^{th} week)	(10^{-2}, 5^{th} week)	(RoV, 5^{th} week)
X_2	2.987	1457.013	0.0032	1.14	2.61
X_{10}	2.245	1457.755	0.0021	2.37	10.52
X_{11}	2.172	1457.828	0.0020	2.31	10.52
X_{18}	5.411	1454.589	0.0047	5.41	10.52
X_{19}	2.201	1457.799	0.0027	3.06	10.52
X_{20}	3.626	1456.374	0.0029	3.37	10.52
X_{21}	1.123	1458.877	0.0014	1.55	10.52
X_{22}	1.633	1458.367	0.0023	2.65	10.52
X_{23}	3.132	1456.868	0.0020	2.28	10.52

4.1. Dynamic risk profile for hemodialysis infection

In order to compute the probability of the top event “hemodialysis infection”, the occurrence probabilities of 23 basic events in Fig. 3 should be determined first. For each device failure in Table 7, to simplify calculation, failure rates λ_i $(i = 1,2,3)$ and repair rates μ_j $(j = 1,2)$ between states are assumed to be:

$$\lambda_1 = \lambda_2, \tag{16}$$

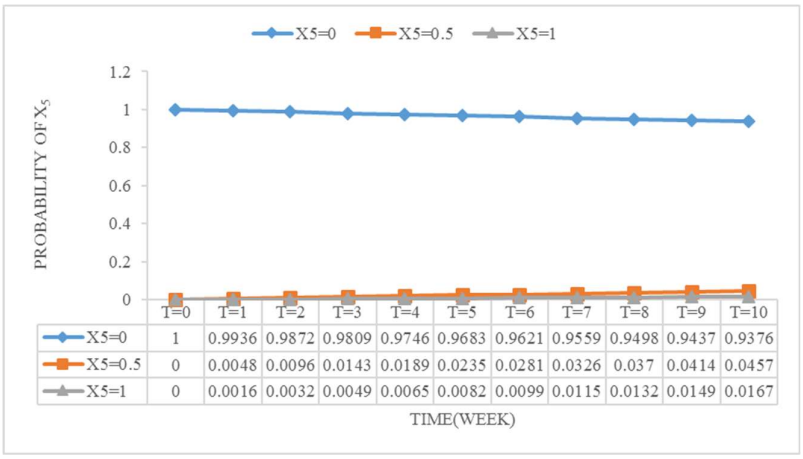
$$\lambda_1 + \lambda_3 = \lambda, \tag{17}$$

$$\lambda_1/\lambda_3 = 3, \tag{18}$$

$$\mu_1 + \mu_2 = \mu, \tag{19}$$

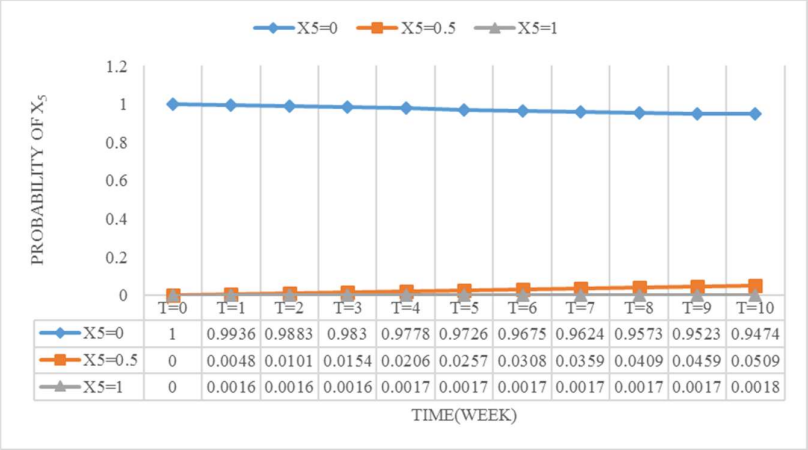
$$\mu_1/\mu_2 = 1/2. \tag{20}$$

Based on the above equations, λ_i $(i = 1,2,3)$ and μ_j $(j = 1,2)$ of each device can be computed and further, failure probabilities at different times can be obtained according to Table 2-4. X_5 (Multi-media filter failure) is taken as an example. The failure probabilities at different time points without repairs, with imperfect repairs, and with perfect repairs, are shown in Fig. 6. As shown in Fig. 6, the probability of X_5 in state “0” is decreasing while that of X_5 in state “0.5” and “1” continues to increase, under all repair degrees. Repair modes make a difference in the slopes of different states. The results are consistent with the degrading process in reality, verifying the effectiveness and practicality of the proposed method. In addition, the failure probabilities are increasing slightly, demonstrating the high reliability of medical devices [12].



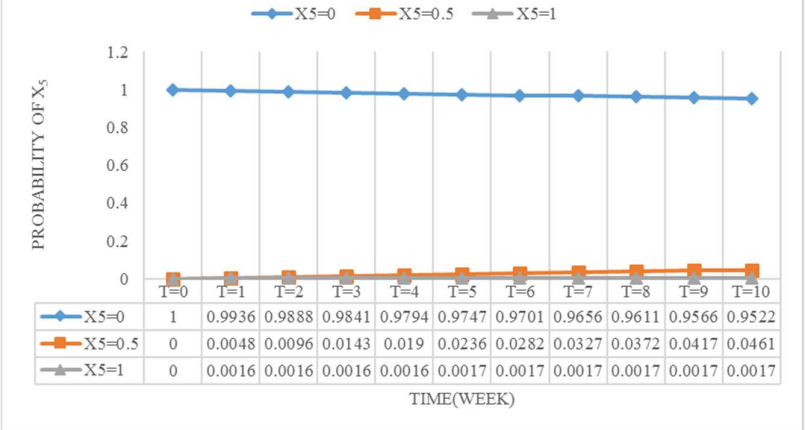
	T=0	T=1	T=2	T=3	T=4	T=5	T=6	T=7	T=8	T=9	T=10
X5=0	1	0.9936	0.9872	0.9809	0.9746	0.9683	0.9621	0.9559	0.9498	0.9437	0.9376
X5=0.5	0	0.0048	0.0096	0.0143	0.0189	0.0235	0.0281	0.0326	0.037	0.0414	0.0457
X5=1	0	0.0016	0.0032	0.0049	0.0065	0.0082	0.0099	0.0115	0.0132	0.0149	0.0167

(a)Probability of X_5 without repairs within 10 weeks



	T=0	T=1	T=2	T=3	T=4	T=5	T=6	T=7	T=8	T=9	T=10
X5=0	1	0.9936	0.9883	0.983	0.9778	0.9726	0.9675	0.9624	0.9573	0.9523	0.9474
X5=0.5	0	0.0048	0.0101	0.0154	0.0206	0.0257	0.0308	0.0359	0.0409	0.0459	0.0509
X5=1	0	0.0016	0.0016	0.0016	0.0017	0.0017	0.0017	0.0017	0.0017	0.0017	0.0018

(b)Probability of X_5 with imperfect repairs within 10 weeks



	T=0	T=1	T=2	T=3	T=4	T=5	T=6	T=7	T=8	T=9	T=10
X5=0	1	0.9936	0.9888	0.9841	0.9794	0.9747	0.9701	0.9656	0.9611	0.9566	0.9522
X5=0.5	0	0.0048	0.0096	0.0143	0.019	0.0236	0.0282	0.0327	0.0372	0.0417	0.0461
X5=1	0	0.0016	0.0016	0.0016	0.0016	0.0017	0.0017	0.0017	0.0017	0.0017	0.0017

(c)Probability of X_5 with perfect repairs within 10 weeks

Fig. 6 Occurrence probability of X_5 within 10 weeks

For each human error (Table 8), Table 9 lists the accident precursor data during 10 weeks. According to equation (8), probability updating can be conducted using these data. X_2 (Error in puncture) is taken as an example to illustrate the proposed method and the result is shown in Fig. 7. As shown in Fig. 7, different from the probability of device failures, that of human errors fluctuates rather than increases continuously. Specifically, human error probability goes up when an error occurs in the current week and goes down when there is no error. It can be attributed to the uncertainty and complexity of human behaviors, which are influenced by various aspects, such as psychological factors, circumstances, etc. [36].

Table 9 Accident precursor data of human errors within 10 weeks

TIME	T=1	T=2	T=3	T=4	T=5	T=6	T=7	T=8	T=9	T=10
n	20	26	25	25	22	27	24	24	20	23
	f	*f*	*f*	*f*	*f*	*f*	*f*	*f*	*f*	*f*
X_2	0	1	1	0	0	1	0	1	0	0
X_{10}	1	0	0	0	0	0	0	2	0	0
X_{11}	0	0	0	1	0	0	1	0	0	1
X_{18}	1	1	0	0	0	0	1	0	0	0
X_{19}	0	1	1	0	0	0	0	0	1	0
X_{20}	0	0	0	0	1	1	0	0	1	0
X_{21}	0	0	0	1	0	0	0	0	0	0
X_{22}	0	0	0	0	2	1	0	1	0	0
X_{23}	0	0	0	0	0	1	0	0	0	1

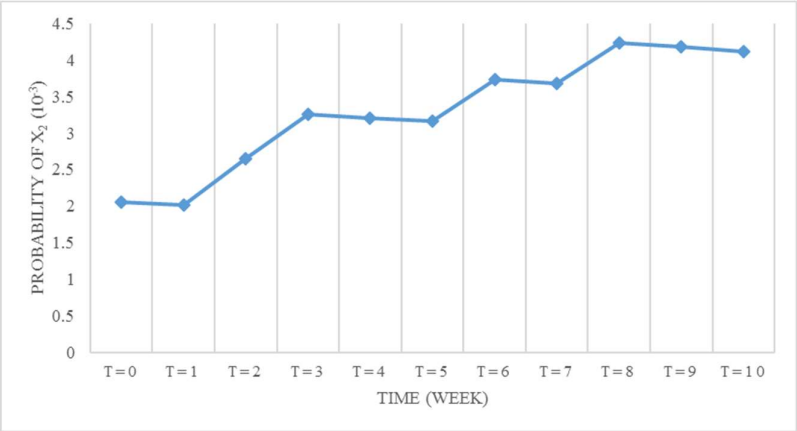


Fig. 7 Occurrence probability of X_2 within 10 weeks

So far, probabilities of basic events per week are obtained, and the probability of hemodialysis infection can be calculated. Fig. 8 shows the occurrence probability of hemodialysis infection without repairs, with imperfect repairs, and with perfect repairs within 10 weeks. In Fig. 8, the occurrence probability of hemodialysis infection increases over time due to the joint impact of device failures and human errors. Especially, the increasing rates of the probability are decreasing when considering no repairs, imperfect and perfect repairs.

Criticality analysis is conducted using the concept of RoV and the results are shown in Table 7 and 8 (the 5th week under imperfect repairs). As illustrated from Table 7, X_4 (catheter contamination), X_5 (multi-media filter failure), X_6 (carbon filter failure), X_7 (demineralizer failure) have the highest RoV, since most hemodialysis patients rely on the catheters that are associated with high infection rates [37]. Besides, hemodialysis patients usually are exposed to a large amount of water that contains a lot of chemicals and bacteria [38]. In addition, almost all the human errors show high RoV of 10.52 (Table 8). In a dialysis unit, limited number of nurses, whose compliances are easily

influenced by such a high-stress circumstance, take care many patients [39]. When the workloads of nurses cannot be alleviated in an effective way, staff training, personal protective equipment and other means are necessary for healthcare to reduce human errors [40].

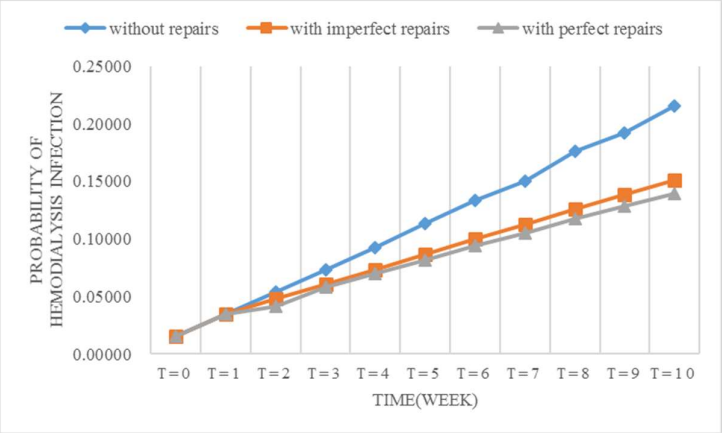


Fig. 8 Occurrence probability of hemodialysis infection within 10 weeks

4.2. Validation of the model

Validation is to demonstrate that the model is a reasonable representation of the actual healthcare system. Due to the operability and intelligibility, a validation method introduced by Jones B et al [41], is adopted. The method requires that the results of the sensitivity analysis should satisfy the following three axioms:

Axiom 1. A slight increase/decrease in the prior probabilities of each basic event should certainly result in a relatively increase/decrease of the posterior probabilities of the top event.

Axiom 2. Given the variation of probability distributions of each basic event, its influence magnitude to the top event values should keep consistency.

Axiom 3. The total influence magnitudes of the combination of the probability

variations from x attributes on the values should be always greater than the one from the set $x - y(y \in x)$ attributes.

Take the basic events $X_1 - X_4$ with imperfect repairs at the 5th week as an example. When X_1 is set to 100%, the probability of the top event "hemodialysis infection" increases to 44.99% from 8.68%. When both X_1 and X_2 are set to 100%, the probability increases to 49.55%. Besides, when X_3 is also set to 100%, the probability increases to 79.71%. When all these four basic events occur, the probability of the top event increases to 95.88%. The sensitivity analysis satisfies all the axioms above, therefore providing a partial validation to the model. In order to conduct a full validation, parameters monitoring and data collecting for a long period are required, but it is impractical in practices in terms of the confidentiality in healthcare, as well as involved time and cost.

5. Conclusions and perspectives

Dynamic risk assessment model for hemodialysis infection is constructed and validated in this paper. Considering unique features of different risk factors, device failures and human errors are handled separately, with the help of dynamic Bayesian network and Bayesian inference, respectively. FTA serves as a modular framework to integrate these factors and evaluate the medical risk as a whole.

The proposed method is applied in a case study of hemodialysis infection from a hospital. The results show that by introducing Bayesian approaches, the probability variations of device failures, human errors and hemodialysis infection during 10 weeks can be obtained. The failure probabilities of devices are increasing slightly, verifying

the high reliability of medical devices. In contrast, the occurrence probabilities of human errors exhibit irregularity as they fluctuate within the 10 weeks, due to their uncertainty and complexity. In addition, the occurrence probability of hemodialysis infection increases over time due to the joint impact of device failures and human errors. Especially, the increasing rates of the probability are decreasing when considering no repairs, imperfect and perfect repairs. Also, catheter contamination, multi-media filter failure, carbon filter failure and demineralizer failure, are identified as the most critical device failures at the 5^{th} week under imperfect repairs. Besides, most of the human errors have high criticalities. Critical events at different times can be identified in the similar way. This offers healthcare the basis for assigning and adapting safety measures. Although this paper uses a specific example of hemodialysis infection, the developed model can also be applied to other healthcare domains, since device failure and human error are common risk factors in healthcare [9].

There are also some limitations in using Bayesian based model. Firstly, the prior probabilities and assumptions, e.g. failure and repair times of device failures follow exponential distributions, influence the accuracy of the model parameters and posterior distributions. Therefore, reliable and abundant input data is necessary to build more accurate models. Secondly, healthcare is much more complex due to the diversity of medical processes, the heterogeneity of patients, the wide range of providers, among others [9]. When those factors and their interactions need to be considered, Bayesian methods become ineffective. In addition, when involving more states, the size of the model explodes, which will consume computational time and cost. To release the

aforementioned limitations, simulation methods, such as Monte Carlo simulation and Petri net, can be helpful tools to supplement the Bayesian approaches [42,43].

Acknowledgements

This research is supported by the National Natural Science Foundation of China (No. 91746205 and No. 71871158).

A. Appendix

This appendix demonstrates how the input data in the case study is collected and obtained. Besides, some examples are given to demonstrate the specific computational procedures of the proposed model.

A.1. Data collection

In the case study, the occurrence probability of hemodialysis infection is computed using the proposed model. According to the proposed method, the required input parameters include: failure rates $\lambda_i\ (i = 1,2,3)$ and repair rates $\mu_j\ (j = 1,2)$ between the states of the medical devices following exponential distribution, as well as α and β of human errors following $\mathrm{Beta}(\alpha, \beta)$ distribution. In order to obtain these data, we investigated a dialysis unit in a hospital from Tianjin, China.

For device failures (Table 7), the unit has the data of average failure rates λ and repair times provided by the devices manufacturers. Then, the repair rates μ are set as the inverse of the repair times. However, the manufacturers didn't provide the more detailed data, such as the failure rates $\lambda_i\ (i = 1,2,3)$ and repair rates $\mu_j\ (j = 1,2)$ between states, since that is relevant to the quality of their products. Therefore, to

simplify calculation, we estimate these parameters by making assumptions between their values (Equation (16)-(20)).

For human errors (Table 8), α and β can be assigned as the number of occurrence and non-occurrence of the human error. However, we couldn't find required data to estimate the values of parameters, since the unit does not have an established medical error reporting mechanism. In addition, infections usually do not occur immediately, creating difficulties in identifying causes and estimating infection rate. Only serious infections had been recorded, but the detailed data is not available due to confidentiality. There are about 1460 cases in the unit last year and the rate of serious infections is around 1.68% (24.53 serious infections per year). Therefore, we assume the number of infections is 24.53 and assign it as the sum of α for all human errors. Then, α of each human error is obtained using random number generation and β is the difference between α and 1460. This method had negative impact on the accuracy of the model at the beginning, which would be compensated by repeating Bayesian inference-based updating.

A.2. Calculation of the occurrence probabilities of X_5

X_5 is taken as an example to demonstrate the computation procedures of the occurrence probabilities of device failures. The failure rate λ and repair rate μ of X_5 are 3.835×10^{-5} and 0.67 per hour, respectively. Based on equation (16)-(20), λ_1-λ_3, and μ_1-μ_2 of X_5 can be computed as: 2.876×10^{-5}, 2.876×10^{-5}, 9.59×10^{-6}, 0.2233 and 0.4467, respectively. At week 0 (T=0), X_5 is assumed to at state "0", meaning the probabilities of X_5 are as follows: $\mathrm{P}(X_5^{T=0}=1)=\ \mathrm{P}(X_5^{T=0}=0.5)=0,\ \mathrm{P}(X_5^{T=0}=$

$0) = 1$, under all repair degrees. At T=1, according to Table 2, the transition probabilities between T=0 and T=1 without repairs are computed as:

Supplementary table 1 transition probabilities of X_5 between T=0 and T=1 (without repairs)

T=0	T=1		
	X_5=0	X_5=0.5	X_5=1
X_5=0	0.9936	0.0048	0.0016
X_5=0.5	0	0.9952	0.0048
X_5=1	0	0	1

According to the total probability formula, the probabilities of X_5 at T=1 without repairs can be obtained as follows:

$$\begin{aligned}\mathrm{P}(X_5^{T=1} = 1) &= \mathrm{P}(X_5^{T=1} = 1 | X_5^{T=0} = 0)\mathrm{P}(X_5^{T=0} = 0) \\ &+ \mathrm{P}(X_5^{T=1} = 1 | X_5^{T=0} = 0.5)\mathrm{P}(X_5^{T=0} = 0.5) \\ &+ \mathrm{P}(X_5^{T=1} = 1 | X_5^{T=0} = 1)\mathrm{P}(X_5^{T=0} = 1) = 0.0016,\end{aligned}$$

$$\begin{aligned}&\mathrm{P}(X_5^{T=1} = 0.5) \\ &= \mathrm{P}(X_5^{T=1} = 0.5 | X_5^{T=0} = 0)\mathrm{P}(X_5^{T=0} = 0) \\ &+ \mathrm{P}(X_5^{T=1} = 0.5 | X_5^{T=0} = 0.5)\mathrm{P}(X_5^{T=0} = 0.5) \\ &+ \mathrm{P}(X_5^{T=1} = 0.5 | X_5^{T=0} = 1)\mathrm{P}(X_5^{T=0} = 1) = 0.0048\end{aligned}$$

$$\mathrm{P}(X_5^{T=1} = 0) = 1 - \mathrm{P}(X_5^{T=1} = 0.5) - \mathrm{P}(X_5^{T=1} = 1) = 0.9936$$

Similarly, the probabilities of X_5 at different times and under different repair degrees can be obtained as shown in Fig. 6.

A.3. Calculation of the occurrence probabilities of X_2

X_2 is taken as an example to demonstrate the computation procedures of the occurrence probabilities of human errors. At T=0, the parameters of X_2 are $\alpha_2^{T=0} = 2.987$ and $\beta_2^{T=0} = 1457.013$. Then, the probability of X_2 is computed based on

equation (7) as follows:

$$\mathrm{P}(X_2^{T=0}) = \frac{2.987}{2.987 + 2457.013} = 0.00205.$$

At T=1, X_2 didn't occur and 20 patients received hemodialysis this week, thus the probability can be updated based on equation (8) as follows:

$$P(X_2^{T=1}) = \frac{2.987 + 0}{1457.013 + 20} = 2.02 \times 10^{-3}.$$

Similarly, the probabilities of X_2 at different times can be obtained as shown in Fig. 7.

A.4. Calculation of the occurrence probabilities of hemodialysis infection

The probability of hemodialysis infection at 5th week with imperfect repairs is taken as an example to demonstrate the computation procedures. The probabilities of basic events are computed using aforementioned procedures and shown as follows:

Supplementary table 2 probabilities of device failures (5th week, with imperfect repairs)

Symbol	$\mathrm{P}(X_i = 0)$	$\mathrm{P}(X_i = 0.5)$	$\mathrm{P}(X_i = 1)$
X_1	0.9833	0.0157	0.001
X_3	0.9816	0.0172	0.0011
X_4	0.968	0.03	0.002
X_5	0.9726	0.0257	0.0017
X_6	0.9784	0.0203	0.0013
X_7	0.9781	0.0206	0.0013
X_8	0.9777	0.0209	0.0013
X_9	0.979	0.0197	0.0013
X_{12}	0.9731	0.0253	0.0016
X_{13}	0.9786	0.0201	0.0013
X_{14}	0.9773	0.0213	0.0014
X_{15}	0.985	0.0141	0.0009
X_{16}	0.9845	0.0145	0.0009
X_{17}	0.9823	0.0167	0.0011

Supplementary table 3 probabilities of human errors (5th week, with imperfect repairs)

Symbol	P(X_i)	Symbol	P(X_i)
X_2	0.00316	X_{20}	0.00293
X_{10}	0.00206	X_{21}	0.00135
X_{11}	0.00201	X_{22}	0.0023
X_{18}	0.0047	X_{23}	0.00198
X_{19}	0.00266		

Examples of computation procedures for different gates in FTA are as follows:

Noisy OR-gate: $\mathrm{P}(G_{10}) = \prod_{i=8}^{9} P(X_i)P(Y|X_i) = 0.0137.$

Noisy AND-gate: $\mathrm{P}(G_7) = \prod_{i=16}^{17} P(X_i)P(Y|X_i) = 0.00276.$

OR gate: $\mathrm{P}(G_{12}) = 1 - \prod_{i=20}^{21}\big(1 - P(X_i)\big) = 0.00427.$

Similarly, the probabilities of hemodialysis infection at different times can be obtained, as shown in Fig. 8, by calculating the probabilities of each logic gate from the bottom to the top.

References

[1] Kohn L T, Corrigan J M, Donaldson M S, et al. To Err is Human: Building a Safer Health System: National Academies Press, 2000.

[2] Ekaette E, Lee R C, Cooke D L, et al. Probabilistic fault tree analysis of a radiation treatment system. Risk Analysis: An Official Publication of the Society for Risk Analysis, 2010, 27(6):1395-1410.

[3] Komal. Fuzzy fault tree analysis for patient safety risk modeling in healthcare under uncertainty. Applied Soft Computing, 2015, 37:942-951.

[4] Khakzad N, Khana F. Dynamic risk analysis using bow-tie approach. Reliability Engineering & System Safety, 2012, 104(104):36-44.

[5] Villa V, Paltrinieri N, Khan F, et al. Towards dynamic risk analysis: A review of the risk assessment approach and its limitations in the chemical process industry. Safety Science, 2016, 89:77-93.

[6] Wittberg P. Overview on Bayesian networks applications for dependability, risk analysis and maintenance areas. Engineering Applications of Artificial Intelligence, 2012, 25(4):671-682.

[7] Ale B, Gulijk C V, Hanea A, et al. Towards BBN based risk modelling of process plants. Safety Science, 2014, 69(1):48-56.

[8] Maglogiannis I, Zafiropoulos E, Platis A, et al. Risk analysis of a patient monitoring system using Bayesian Network modeling. Journal of Biomedical Informatics, 2006, 39(6):637-647.

[9] Kazemi R, Mosleh A, Dierks M. A Hybrid Methodology for Modeling Risk of Adverse Events in Complex Health‐Care Settings. Risk analysis, 2017, 37(3): 421-440.

[10] Trucco P, Cavallin M. A quantitative approach to clinical risk assessment: The CREA method. Safety Science, 2006, 44(6):491-513.

[11] Ong M S, Coiera E. Safety through redundancy: a case study of in-hospital patient transfers. Quality & Safety in Health Care, 2010, 19(5): e32.

[12] Weininger S, Kapur K C, Pecht M. Exploring Medical Device Reliability and Its Relationship to Safety and Effectiveness. IEEE Transactions on Components & Packaging Technologies, 2010, 33(1):240-245.

[13] Khakzad, Nima. Application of dynamic Bayesian network to risk analysis of

domino effects in chemical infrastructures. Reliability Engineering & System Safety, 2015, 138:263-272.

[14] Feroze U, Kalantarzadeh K, Sterling K A, et al. Examining Associations of Circulating Endotoxin with Nutritional Status, Inflammation and Mortality in Hemodialysis Patients. J Ren Nutr, 2012, 22(3):317-326.

[15] Su Y, Norris J L, Zang C, et al. Incidence of hepatitis C virus infection in patients on hemodialysis: A systematic review and meta‐analysis. Hemodialysis International, 2013, 17(4): 532-541.

[16] Nielsen T D, Jensen F V. Bayesian networks and decision graphs. Springer Science & Business Media, 2009.

[17] Khakzad N, Khan F, Amyotte P. Safety analysis in process facilities: Comparison of fault tree and Bayesian network approaches. Reliability Engineering & System Safety, 2011, 96(8):925-932.

[18] Kelly D, Smith C. Bayesian Inference for Probabilistic Risk Assessment. Springer, 2011.

[19] Alquist M, Bosch J P. Treatment mapping--a systematic methodology to assess quality, efficiency and variability in the hemodialysis delivery process. Blood Purification, 2008, 26(5):417-422.

[20] Thomas A, Silver S A, Rathe A, et al. Feasibility of a hemodialysis safety checklist for nurses and patients: a quality improvement study. Clinical Kidney Journal, 2016, 9(3):335-342.

[21] Kliger A S. Maintaining Safety in the Dialysis Facility. Clinical Journal of the

American Society of Nephrology Cjasn, 2015, 24(1):19-21.

[22] United States Renal Data System. USRDS 2017 Annual Data Report: Atlas of Chronic Kidney Disease and End-Stage Renal Disease in the United States. Vol. 2. Ch. 1: Incidence, Prevalence, Patient Characteristics, and Treatment Modalities. Bethesda, MD: National Institutes of Health, National Institute of Diabetes and Digestive and Kidney Diseases, 2018.

[23] Liyanage T, Ninomiya T, Jha V, et al. Worldwide access to treatment for end-stage kidney disease: a systematic review. Lancet, 2015, 385(9981):1975-1982.

[24] Toida T , Sato Y , Nakagawa H , et al. Glycaemic control is a predictor of infection-related hospitalization on haemodialysis patients: Miyazaki Dialysis Cohort study (MID study). Nephrology, 2016, 21(3):236-240.

[25] Wong L P . Systems Thinking and Leadership How Nephrologists Can Transform Dialysis Safety to Prevent Infections. Clin J Am Soc Nephrol, 2018, 13(4):655-662.

[26] Samani S , Saffari M , Charkhchian M , et al. Incidence and risk factors of bloodstream catheter-related infections in hemodialysis patients. Comparative Clinical Pathology, 2015, 24(2):275-279.

[27] Sahli F, Feidjel R, Laalaoui R. Hemodialysis catheter-related infection: rates, risk factors and pathogens. Journal of infection and public health, 2017, 10(4): 403-408.

[28] Kohda T, Cui W. Risk-based reconfiguration of safety monitoring system using dynamic Bayesian network. Reliability Engineering & System Safety, 2007,

92(12):1716-1723.

[29] Vose, D. Risk analysis, a quantitative approach. 2000, John Wiley & Sons Ltd.

[30] Kalantarnia M, Khan F, Hawboldt K. Dynamic risk assessment using failure assessment and Bayesian theory. Journal of Loss Prevention in the Process Industries, 2009, 22(5):600-606.

[31] Groth K M, Smith C L, Swiler L P. A Bayesian method for using simulator data to enhance human error probabilities assigned by existing HRA methods. Reliability Engineering & System Safety, 2014, 128:32-40.

[32] Meel A, Seider W D. Plant-specific dynamic failure assessment using Bayesian theory. Chemical Engineering Science, 2006, 61(21):7036-7056.

[33] Neapolitan R E. Learning Bayesian networks. 2003, London, Prentice Hall.

[34] Zarei E, Azadeh A, Khakzad N, et al. Dynamic safety assessment of natural gas stations using Bayesian network. Journal of hazardous materials, 2017, 321: 830-840.

[35] Zarei E, Azadeh A, Aliabadi M M, et al. Dynamic safety risk modeling of process systems using bayesian network. Process Safety Progress, 2017, 36(4): 399-407.

[36] Onofrio R, Trucco P, Torchio A. Towards a taxonomy of influencing factors for human reliability analysis (HRA) applications in surgery. Procedia Manufacturing, 2015, 3: 144-151.

[37] Xue H, Ix J H, Wang W, et al. Hemodialysis access usage patterns in the incident dialysis year and associated catheter-related complications. American journal of kidney diseases, 2013, 61(1): 123-130.

[38] Coulliette A D, Arduino M J. Hemodialysis and water quality. Seminars in Dialysis, 2013, 26(4):427-438.

[39] Oliveira A C, Cardoso C S, Mascarenhas D. Contact precautions in intensive care units: facilitating and inhibiting factors for professionals' adherence[J]. Revista Da Escola De Enfermagem Da Usp, 2010, 44(1):161-165.

[40] Karkar A, Bouhaha B M, Dammang M L. Infection control in hemodialysis units: a quick access to essential elements. Saudi J Kidney Dis Transpl, 2014, 25(3):496-519.

[41] Jones B, Jenkinson I, Yang Z, et al. The use of Bayesian network modelling for maintenance planning in a manufacturing industry. Reliability Engineering & System Safety, 2010, 95(3):267-277.

[42] Rubinstein R Y, Kroese D P. Simulation and the Monte Carlo method. John Wiley & Sons, 2016.

[43] David R, Alla H. Discrete, continuous, and hybrid Petri nets. Berlin: Springer, 2005.