A Dynamic Model for Therapy Selection in ICU Patients with VAP

Theodore Charitos[1], **Stefan Visscher**[2],
Linda C. van der Gaag[1], **Peter Lucas**[3], and **Karin Schurink**[2]

[1] Dept. of Inform. and Comp. Sciences, Utrecht University,The Netherlands
[2] Dept. of Internal Medic. and Infect. Diseases, University Medical Center Utrecht, The Netherlands
[3] Inst. for Comp. and Inform. Sciences, Radboud University, Nijmegen, The Netherlands

Abstract

Treating ventilator-associated pneumonia in mechanically ventilated patients in intensive care units is seen as a clinical challenge. In this paper, we develop a dynamic-decision model that explicitly captures the development of the disease over time. To represent the dependencies between the variables involved in a compact way we use a dynamic Bayesian network and combine it with the framework of partially observable Markov decision processes to choose optimal antimicrobial therapy for respiratory tract infections. We discuss implementation issues and modelling advantages of our model and demonstrate its use for a number of real patients.

1 Introduction

Many patients admitted to an intensive care unit (ICU) need respiratory support by a mechanical ventilator, which promotes the development of ventilator-associated pneumonia (VAP) in these patients. Effective and fast treatment of VAP is seen as an issue of major significance. The difficulty in diagnosing VAP is in the lack of an accurate, non-invasive (that is, patient-friendly) gold standard; VAP is therefore diagnosed by taking a number of different clinical features into account [9; 14].

A prominent role in the development of VAP is played by two stochastic processes: *colonisation* of the laryngotracheobronchial tree by pathogens and the onset and development of *pneumonia*. A dynamic Bayesian network, called dVAP was developed that explicitly captures the temporal relationships between the variables involved [5]. This network takes into account the patient's characteristics from earlier days when performing diagnosis. The numerical part of the network was constructed from estimations by infectious-disease experts and from the literature. In a later stage these probabilities were updated through machine learning using collected patient data, which resulted in a better diagnostic performance of the model.

The treatment of VAP is seen as a significant problem by ICU doctors. Firstly, many of the patients suffering from VAP are severely ill. Secondly, the presence of multi-resistant bacteria in clinical wards, in particular the ICU, makes prescription of antibiotics with a spectrum as narrow as possible essential; the description of broad-spectrum antibiotics promotes the development of antimicrobial resistance, and should therefore be avoided when possible. In this paper, we address optimal therapy selection using the dVAP model. For this purpose, we focus on the framework of partially observable Markov decision processes (POMDPs) [1; 7; 12; 15] for sequential decision making.

Although the standard POMDP framework in essence allows us to capture the main elements of choosing a therapy of VAP, it cannot be used directly, mainly because: (1) the number of parameters required can be huge, and (2) exact methods for solving the problem are computationally very demanding and only small problems can be solved exactly. In view of these considerations, we extend the dVAP network and construct a dynamic-decision model that incorporates the uncertainty included in the treatment procedure. We then use the Perseus algorithm for its evaluation [16]. Perseus is a point-based approximate value-iteration algorithm for POMDPs that achieves competitive performance both in terms of solution and speed comparing to alternative (and more complex) algorithms in the literature [3]. Perseus can moreover be easily implemented in practice [13]. Perseus, however, is designed for problems without any structure among the variables representing the state of the process. We enhance the applicability of Perseus for our structured domain to take advantage of the factorisations and independencies among the variables included in the dVAP model.

We tested the resulting model on a group of patients drawn from the files of the ICU of the University Medical Center Utrecht in the Netherlands. The solutions obtained indicate that our dynamic-decision model provides a useful framework for solving and analysing complex decision problems. Our results in fact advocate further application of Perseus in structured domains of other medical therapy problems.

The remainder of this paper is organised as follows. In Section 2, we describe the dVAP network for the diagnosis of VAP. In Section 3 we describe the basics of the POMDP framework and of the Perseus algorithm; in Section 4 we discuss modelling and computational issues related to applying Perseus to decision making for patients with VAP. Section 5 presents and discusses the results from an evaluation study. Finally, the paper ends with our conclusions in Section 6.

2 Diagnosing VAP

We begin by discussing the pathophysiology of VAP and then describe the dVAP model that captures the development of VAP.

2.1 Pathophysiology of VAP

Ventilator-associated pneumonia is, when looking at a daily level, a low-prevalence disease occurring in mechanically-ventilated patients in critical care and involves infection of the lower respiratory tract [2]. In contrast to infections of more frequently involved organs (such as the urinary tract), for which mortality is low, ranging from 1 to 4%, the mortality rate for VAP ranges from 24 to 50% and can reach 76% for some high-risk pathogens. Variables that change due to the development of VAP, among others, are an increased *body temperature*, an abnormal amount of coloured *sputum*, *signs* on the chest X-ray, the duration of *mechanical ventilation*, and an abnormal number of *leukocytes*.

The relationship between the *colonisation* by pathogens and the development of *pneumonia* is captured as follows. Periodically, a sample of the patient's sputum is cultured at the laboratory. When the culture shows a number of colonies of a particular bacterium that is above a particular threshold, the patient is said to be colonised by this bacterium. The seven groups of microorganisms that occur most frequently in critically ill patients and cause colonisation, are modelled in the *diagnostic part* of the network. Information about which bacterium or bacteria are currently present in a patient and the current signs and symptoms constitute the basis for choosing optimal antimicrobial treatment.

2.2 A dynamic model for diagnosis

A *dynamic Bayesian network* (DBN) is a graphical model that encodes a joint probability distribution on a set of stochastic variables, explicitly capturing the temporal relationships between them. More formally, let $\mathcal{V}_n = (V_n^1, \ldots, V_n^m), m \geq 1$, denote the set of variables at time n. Then, a dynamic Bayesian network is a tuple (B_1, B_2), where B_1 is a Bayesian network that represents the prior distribution for the variables in the first time slice $\mathcal{V}_1$, and B_2 defines the transitional relationships between the variables for two consecutive time slices, so that for every $n \geq 2$

$$p(\mathcal{V}_n \mid \mathcal{V}_{n-1}) = \prod_{i=1}^{m} p(V_n^i \mid \pi(V_n^i))$$

where $\pi(V_n^i)$ denotes the set of parents of V_n^i, for $i = 1, \ldots, m$.

DBNs are usually assumed to be time invariant, which means that the topology and the parameters of the network per time slice and across time slices do not change. Moreover, the Markov property for transitional dependence is assumed, which means that $\pi(V_n^i)$ can include variables either from the same time slice n or from the previous slice $n-1$, but not from earlier time slices [10]. Then, by unrolling B_2 for N time slices, a joint probability distribution $p(\mathcal{V}_1, \ldots, \mathcal{V}_N)$ is defined for which the following decomposition property holds:

$$p(\mathcal{V}_1, \ldots, \mathcal{V}_N) = \prod_{n=1}^{N} \prod_{i=1}^{m} p(V_n^i \mid \pi(V_n^i))$$

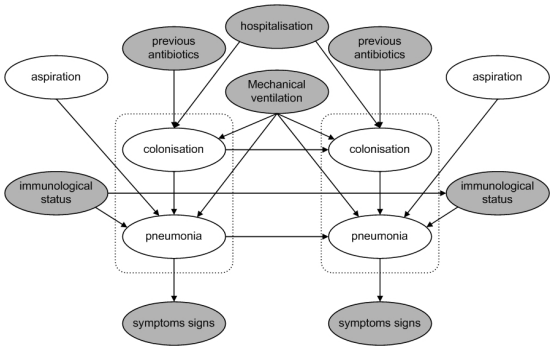

Figure 1: The dVAP network for the diagnosis of VAP; clear variables are hidden, shaded variables are observable. The dashed boxes indicate the hidden processes.

Monitoring in a DBN is the task of computing the probability distribution for a set of variables of interest $\mathcal{X}_n \subset \mathcal{V}_n$ at time n given the observations that are available up to and including time n.

2.3 Modelling and computational issues

An overview of the structure of the dynamic network constructed for the diagnosis of VAP [5] is depicted in Figure 1. The dVAP network includes two interacting dynamic hidden processes, modelled by the compound variables *colonisation* (7 variables) and *pneumonia* (8 variables). The process of colonisation is influenced by three input variables: *hospitalisation*, *mechanical ventilation* and *previous antibiotics*, and one hidden variable *aspiration* that in essence controls its dynamics. We note that the variables *hospitalisation* and *mechanical ventilation* are observed for a period that is longer than the transition interval of the model. The variables thus are modelled as affecting adjacent time slices. The variable *previous antibiotics* represents the effect of previous medication to the patient on the process of colonisation. The symptoms and signs of pneumonia are depicted in Figure 2. These variables are included in the *diagnostic part* of the network.

The practicability of the dVAP network depends to a large extent on the computational burden of inference with the network. For diagnosing patients with VAP, monitoring is performed at each time. For this purpose, the *interface algorithm* can be applied [10]. This algorithm is an extension of the *junction-tree algorithm* for inference with Bayesian networks in general [6] and efficiently exploits the forward interface of a dynamic network. Recall that the forward interface is the set of variables at time slice n that affect some variables at time slice $n+1$ directly. Note that the interface algorithm is linear in the total number of time slices and for large time scopes, the computation time can prove to be prohibitive in practice.

Recent results show that, in case consecutive similar observations are obtained, the probability distribution of the hidden process converges to a limit distribution within a given level of accuracy [4]. After some number of time slices, therefore, there is no need for further inference as long as similar observations are obtained. The phenomenon of consecutive similar observations was evident for several patients in the ICU files. For example, for these patients

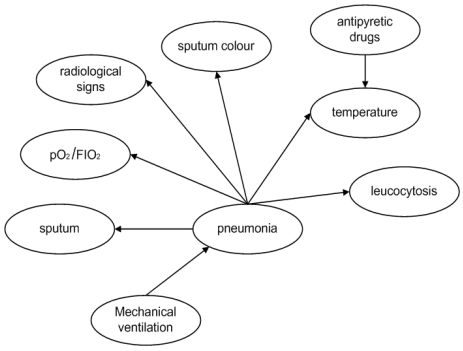

Figure 2: Symptoms and signs of pneumonia.

it was found that the same combination of values was observed for all or almost all of the observable variables for a number of consecutive days. On a set of ICU patients, test results indicated that representing time explicitly and taking into consideration the history of the patient increased diagnostic performance [5].

3 Therapy planning

In this section we describe our approach to solving the dynamic-decision model for patients with VAP. We begin with the theoretical background of POMDPs.

3.1 Basics of POMDPs

Partially Observable Markov Decision Processes (POMDPs) constitute a common framework for decision making about complex dynamic processes where the state of the process cannot be fully observed [1; 7; 15; 16]. A POMDP more specifically describes a stochastic process of which the states are hidden and for which decisions can only be based on observations seen and past actions performed.

Formally, a POMDP is a 6-tuple (S, Θ, A, P, O, R) where S is a finite set of states of the hidden process; Θ is a finite set of observations (findings, results of diagnostic tests); A is a finite set of actions; $P : S \times A \times S \rightarrow [0, 1]$ is a set of *Markovian transition models*, one for each action α, such that $p_\alpha(s' \mid s)$ represents the probability of going from state s to s' with action α; $O : S \times A \times \Theta \rightarrow [0, 1]$ is a set of *observation models*, one for each action α, such that $p_\alpha(o \mid s')$ represents the probability of making observation o after taking action α and transitioning to state s'; and R is a reward function $R : S \times A \times S \times \Theta \rightarrow \mathbb{R}$, such that $R(s, \alpha)$ represents the expected reward received in state s after taking action α.

Given a POMDP, the goal is to construct a *control policy* that maximizes an *objective (value) function*. The objective function combines rewards over multiple time slices, and typically is the expectation of the cumulative sum of rewards r_n at each time n over a *finite-horizon* of N slices, that is $E(\sum_{n=1}^{N} r_n)$, or over a *discounted infinite-horizon*, that is $E(\sum_{n=1}^{\infty} \gamma^n r_n)$, where $0 < \gamma < 1$ is a discount rate. In this paper we focus on the discounted infinite-horizon model as in previous applications of POMDPs in medicine [7].

A belief state b assigns a probability $b(s)$ to every possible state $s \in S$. There thus are an infinite number of possible belief states over S. An optimal policy for b has a *value function* that satisfies the Bellman optimality equation

$$V^*(b) = \max_{\alpha \in A} \left[r(b, \alpha) + \gamma \sum_{o \in \Theta} p(o \mid b, \alpha) V^*(\tau(b, \alpha, o)) \right] \quad (1)$$

where

- $r(b, \alpha) = \sum_{s \in S} b(s) R(s, \alpha)$;
- $p(o \mid b, \alpha) = \sum_{s' \in S} p(o \mid s', \alpha) \sum_{s \in S} p(s' \mid s, \alpha) b(s)$;
- $\tau(b, \alpha, o) \propto p(o \mid s, \alpha) \sum_{s' \in S} p(s \mid s', \alpha) b(s')$;

in which $r(b, \alpha)$ represents the expected reward for a belief state b and current action α, $p(o \mid b, \alpha)$ represents the probability of making observation o one time slice ahead under current action α for a belief state b, and $\tau(b, \alpha, o)$ is the update of the belief state given a previous belief state b and action α, and a current observation o. The optimal policy $\mu^* : b \rightarrow A$ now selects the value-maximizing action

$$\mu^*(b) = \arg\max_{\alpha \in A} \left[r(b, \alpha) + \gamma \sum_{o \in \Theta} p(o \mid b, \alpha) V^*(\tau(b, \alpha, o)) \right]$$

In order to compute the value function $V^*(b)$ in equation (1) we can use the *value iteration algorithm* [15], which guarantees that the sequence of value function approximations V_i defined as

$$V_i(b) = \max_{\alpha \in A} \left[r(b, a) + \gamma \sum_{o \in \Theta} p(o \mid b, \alpha) V_{i-1}(\tau(b, \alpha, o)) \right] \quad (2)$$

converges to the optimal solution. An important property of this approximation sequence is that the value functions $V_i(b)$ in equation (2) are piecewise linear and convex, which allows for computing the update in finite time for the complete belief space [1]. However, the computational cost of doing so is high for all but trivial problems, and thus several methods have been proposed in the literature that try to approximate the optimal value function V^* [8].

3.2 The Perseus algorithm

Perseus is an efficient point based approximate value iteration algorithm for POMDPs [16]. The main idea is to use a set of reachable belief states B that are sampled from the belief simplex to perform value function updates, ensuring that in each iteration the new value function is an upper bound to the previous value function, as estimated on the sampled set of belief states. The intuition behind this approach is that in most practical problems the belief simplex is sparse, in the sense that only a limited number of belief states can ever be reached by letting the hidden process interact with its environment. The algorithm performs value function updates, making sure that in each step the new value function estimate $V_{i+1}(b)$ is an upper bound for $V_i(b)$ for all $b \in B$. The major advantage of Perseus is that in each iteration i it uses only a (random) subset of states in B until the value $V_i(b)$ of every $b \in B$ has improved or remained the same. This property makes the algorithm efficient even in problem domains with large state spaces compared to other approximate methods [8]. We note, however, that the Perseus algorithm has been designed for POMDPs

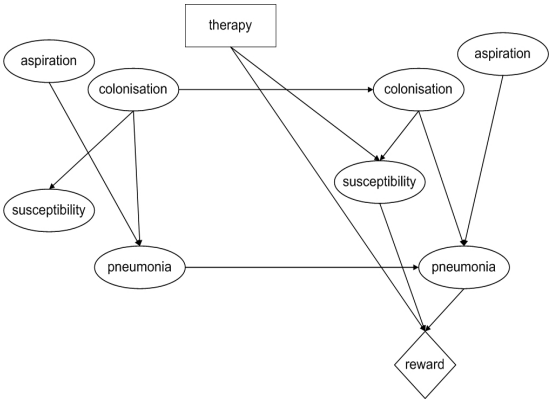

Figure 3: The dynamic-decision network for therapy selection of VAP. Action choice is represented as a rectangle and the reward function as a diamond. The observable variables are excluded for clarity.

with flat belief state, that is, for states without any type of internal structure. To incorporate Perseus into our approach to solving the dynamic-decision model for VAP, we have to enhance its applicability to more structured domains. We discuss such modifications in the next section.

4 Decision making for VAP

The aim of our dynamic-decision model is to aid clinicians in dealing with patients with suspected VAP. Optimal antimicrobial therapy for VAP is selected by balancing the expected efficacy of treatment, which is related to the number of pathogens causing the infection, against the spectrum of antimicrobial treatment. Each of the seven modelled groups of pathogens are susceptible to particular antibiotics. Some of these pathogens are easy to cover. For these pathogens, a narrow or even very narrow antimicrobial spectrum is sufficient. Some pathogens, however, are more difficult to eradicate. Here, we need broader spectrum antibiotics. The problem of prescribing unnecessarily broad-spectrum antibiotics is the occurrence of antibiotic resistance, which means that pathogens are no longer susceptible to a particular antibiotic. Antibiotic resistance is a well-known problem in health-care [2]. Our dynamic-decision model now incorporates the idea of prescribing antibiotic spectra as narrow as possible. The narrower the spectrum, the higher the preference. How well the pathogen is covered by an antibiotic times the preference of the broadness of its spectrum gives the final utility of prescribing this antibiotic [9]. The prescribed treatment thus is a trade-off between maximising coverage and narrowing broadness of spectrum.

To incorporate for decision making in the dVAP model, we add a decision-theoretic part that represents the effect of selected therapy on the probability distribution of VAP. Figure 3 depicts the resulting model. The dynamic-decision model includes the hidden compound variable *susceptibility* (8 variables) that represents the susceptibility of the suspected pathogens to particular antibiotics. A causal independence model, known as the noisy-AND gate [9], is used to model the conjunctive effect of antibiotics on the susceptibility of pathogens. The model thus includes 24 binary hidden variables with 2^{24} possible configurations. The *therapy* variable includes 26 different antibiotics or combinations of antibiotics and the value "none" indicating that the clinician does not prescribe any antibiotic to the patient. These antibiotics have been further classified into four different groups from very narrow to very broad, according to their spectrum. The reward function is thus based on these four spectrum groups and has been assessed by a domain expert [14]. Insight into the potential efficacy of treatment can be obtained by entering symptoms and signs of a patient. In total, the model contains 13 observable variables with 1382400 possible configurations.

At first sight, it seems impossible for Perseus or for any other algorithm to solve our model since both $|S|$ and $|\Theta|$ are extremely large. However, an important feature of our model is that its state and observation sets are not flat, but structured in a factored way. More specifically, the states and the observations of the model are not represented enumeratively but via hidden and observation variables respectively. We further note that although the hidden state of our model consists of 24 variables, only the variables pneumonia and susceptibility are important for decision making. Now, to make efficient use of the Perseus algorithm, we compute the joint probability distribution of just these two variables, which can be done in a similar manner to monitoring in the dVAP model. Our implementation of Perseus in addition takes into account that some variables in the forward interface are observable. Since, for example, *immunological status* is always observed and *colonisation* can be observed for some days, the belief state is modelled as a hybrid state with an observed and a hidden component [7]. Finally, we observe from Figure 2 that we have to consider just six observable variables that are probabilistically affected by pneumonia.

To decrease the computational burden of applying Perseus, we further do not take all observable variables into account when computing the summation in equation (2). That is, upon applying Perseus we sample belief states reflecting realistic data settings only. For example, VAP by definition may be initiated after a patient has been ventilated for more than two days. The state of the mechanical ventilation variable can thus be selected in every iteration of equation (2), from among just the states in which the duration of the ventilation is greater than two days.

As a result of the above modifications, the set Θ includes to 768 possible combinations, and thus is smaller in size than the original set by a factor 1800. The aforementioned considerations were used initially to create a set of reachable belief states B and then to apply Perseus with $\gamma = 0.95$. In our experiments on a 2.4 GHz Intel(R) Pentium computer, creating B took 1.5 seconds per belief state, while computing an optimal policy took approximately one minute using a total of 10000 sampled belief states.

5 Evaluation

We examined the performance of our dynamic-decision model on 5 patients diagnosed with VAP randomly selected from a prospectively collected database of ICU patients. Using the dVAP network we monitored these patients and computed their belief state per day for a total of 10 days.

day	p(VAP)	+ colon.path.	antibiotic
2	0.2295	- -	none meropenem (b)
3	0.0049	Enterobacteria2 -	cotrimoxazol (n) none
4	0.0052	- -	cotrimoxazol (n) none
5	0.0848	- -	cotrimoxazol (n) none
6	0.3401	- -	none cotrimoxazol (n)
7	0.0363	- -	none erythromycin (vn)
8	0.0017	P.aeruginosa Enterobacteria2	cotrimoxazol (n) none
9	0.0012	P.aeruginosa Enterobacteria2	cotrimoxazol (n) none
10	0.0046	- -	cotrimoxazol (n) none

(a) patient A

day	p(VAP)	+ colon.path.	antibiotic
2	0.028	- -	none cotrimoxazol (n)
3	0.0344	- -	none cotrimoxazol (n)
4	0.1905	Enterobacteria2 S.aureus	cotrimoxazol (n) erythromycin (vn)
5	0.5929	- -	cotrimoxazol (n) erythromycin (vn)
6	0.5445	- -	cotrimoxazol (n) erythromycin (vn)
7	0.9823	- -	cotrimoxazol (n) erythromycin (vn)
8	0.9791	Acinetobacter -	ceftazidim (i) aztreonam (i)
9	0.9459	Acinetobacter S.aureus, S.pneumoniae	cotrimoxazol (n) meropenem (b)
10	0.9918	- -	cotrimoxazol (n) meropenem (b)

(b) patient B

Table 1: The best two recommendations (and their spectrum in parenthesis) at each time slice for two patients. Abbreviations for antibiotic spectrum: vn=very narrow; n=narrow; i=intermediate; b=broad.

Contrary to an earlier evaluation of the diagnostic performance of the dVAP network [5], we took into account the sparse colonisation data that existed in the datasets of some patients. In contrast to the data for the observable variables that were readily available, the colonisation data were provided by the laboratory from sputum cultures and took on average 48 hours to become available. Also, these data concerned only a (small) subset of colonisation pathogens and were observed for a few days (maximum 3). To process the colonisation data, we assumed that whenever there was a positive culture for a specific pathogen on a specific day, then the values of the other non-observed pathogens were set to negative. We are aware that this assumption should be used with care. More specifically, the transition matrices estimated by the expert [5] suggested that, under particular conditions, if a pathogen is positive (negative) on one day then it cannot be negative (positive) on the next day. For one patient for example, we noticed upon processing the available data, that on day 8 we assumed the presence of S.aureus and S.pneumoniae to be negative while these two pathogens were actually observed to be positive on day 9. To resolve this issue, we made no assumption about these two pathogens on day 8 and left their value as unobserved.

We compared the recommended decisions from the model with an expert opinion as to the most appropriate antibiotics to cover the likely pathogens. The results were not entirely satisfactory in the opinion of the expert. For one patient, for whom no colonisation data were available, we found that the decisions recommended by the model were acceptable; for two patients the model recommended too broad a spectrum antibiotics, while for the other two patients the recommended antibiotics did not cover the observed pathogens. A possible explanation of this suboptimal performance of the model is that its decisions are strongly affected by the probability of VAP at each time and less by the colonising pathogens; that is, the prescription of antibiotics is heavily dominated by $p(\text{VAP})$, while less weight is given to the presence of colonisation data. For example, if on a specific day a patient has a very small $p(\text{VAP})$ but a positive colonising pathogen, then the model will abstain from prescribing antibiotics and will not use a narrow spectrum antibiotics as would be expected. Another reason is that the influence of the colonisation data on the recommended decision diminished with time according to the specification of the model. More precisely, since colonisation data are sparsely observed, a colonisation pathogen found to be positive on one day will have minor effect on the decision taken two days later because of the Markov assumption underlying the dVAP network.

In view of the above considerations, we enhanced our decision model to incorporate the influence of the colonisation data on the recommended decision in a more appropriate fashion. For each colonisation (group of) pathogen(s) found to be positive on a given day, we force the model to prescribe antibiotics to cover this pathogen. In this way our model considers a conglomeration of different decision plans that are influenced by the presence of positive pathogens in the patient's dataset. To cope with the sparsity of the colonisation data, we use the enhanced model for the following two days as well. As a result, the clinician is presented with a therapy plan that aims to cover positive observed pathogens for at least three days. For the remaining days for which no colonisation data were available, the original decision model was used. The evalua-

tion now showed now that the new recommendations better comply with the expert's recommendations.

We discuss the results for two patients in order to convey how our dynamic-decision model might be employed clinically, and to point out some of its limitations. For patient A, the dVAP network assigns a small probability to VAP for almost all the days. As a consequence, the decision not to prescribe any antibiotic is always recommended by the model. However, positive cultures of pathogens are observed for the days 3, 8 and 9. For these days (and for the next two days) the antibiotic cotrimoxazol (narrow spectrum) is recommended first. This recommendation reflects the ability of the model to prescribe an antibiotic even if the probability of VAP is very small. We note, however, that on day 2, the model suggests the antibiotic meropenem. This recommendation is far too broad for this patient, and raises the question whether alternative utility models might alleviate this problem. For patient B, the dVAP network assigns quite early (day 5) a relatively high probability to VAP which even further increases in the following days. In addition, positive cultures of pathogens are observed for the days 4, 8 and 9. Our dynamic-decision model takes into account both the high probability of VAP and the positive cultures to recommend appropriate antibiotics that belong to a narrow spectrum whenever possible. This is evident in the recommendations for days 4 to 7, while for days 8 to 10 the recommendation belongs to the intermediate or broad spectrum. The predictions made and the therapy suggested by the model for both patients are shown in Table 1.

6 Conclusions

We have described the development of a dynamic-decision model that is able to assist clinicians in the clinical management of ventilator-associated pneumonia. For the purpose of computing appropriate decisions from the model, we applied the framework of partially observable Markov decision processes for modelling the action-outcome uncertainty and partial observability. The application and potential of the POMDP framework to medical planning has been discussed in [12] and successfully explored in [7]; in the latter work, a hierarchical Bayesian network was used to represent the disease dynamics and to decrease the computational burden involved. Since exact computation in a POMDP is intractable, we discussed the application of the Perseus algorithm to our problem, in which the belief state of the hidden process is structured. The solutions obtained for a small set of patients from an initial evaluation of our model showed that POMDPs could provide a useful framework for solving complex decision problems. We feel that the promising results justify further refinement and extension of our current model as well as application of our framework to other complex structured decision problems [11].

Acknowledgements

This research was (partly) supported by the Netherlands Organization for Scientific Research (NWO). The first author would like to thank Matthijs Spaan and Nikos Vlassis for their help in implementing Perseus.

References

[1] J. Astrom (1965). Optimal control of Markov decision processes with incomplete state estimation. *Journal of Mathematical Analysis and Applications*, 10:174-205.

[2] M.J.M. Bonten (2004). Prevention of infection in the intensive care unit. *Current Opinion in Critical Care*, 10(5):364-368.

[3] R. Brafman (2006). Personal communication.

[4] T. Charitos, P. de Waal, and L.C. van der Gaag (2005). Speeding up inference in Markovian models. *Proceedings of the 18th International FLAIRS conference*, pp. 785-790.

[5] T. Charitos, L.C. van der Gaag, S. Visscher, K. Schurink, and P. Lucas (2005). A dynamic Bayesian network for diagnosing ventilator-associated pneumonia in ICU patients. Working notes of the 10th IDAMAP Workshop, pp. 32-37.

[6] R.G. Cowell, A.P. Dawid, S.L. Lauritzen, and D.J. Spiegelhalter (1999). *Probabilistic Networks and Expert Systems*. Springer.

[7] M. Hauskrecht and H.S.F. Fraser (2000). Planning treatment of ischemic heart disease with partially observable Markov decision processes. *Artificial Intelligence in Medicine*, 18(3):221-244.

[8] M. Hauskrecht (2000). Value-Function Approximations for Partially Observable Markov Decision Processes. *Journal of Artificial Intelligence Research*, 13: 33-94.

[9] P.J.F. Lucas, N.C de Bruijn, C.A.M Schurink, and A. Hoepelman (2000). A probabilistic and decision theoretic approach to the management of infectious disease at the ICU. *Artificial Intelligence in Medicine*, 19(3):251-279.

[10] K. Murphy (2002). *Dynamic Bayesian networks: Representation, Inference and Learning*. Ph.D. thesis, University of California Berkeley.

[11] P. Poupart (2005). *Exploiting Structure to Efficiently Solve Large Scale Partially Observable Markov Decision Processes*. Ph.D. thesis, University of Toronto.

[12] N.B. Peek (1999). Explicit temporal models for decision-theoretic planning of clinical management. *Artificial Intelligence in Medicine*, 15(2): 135-154.

[13] Perseus: URL: http://www.science.uva.nl/ mtjspaan/pomdp.

[14] C.A.M. Schurink (2003). *Ventilator Associated Pneumonia: a Diagnostic Challenge*. Ph.D. thesis, Utrecht University.

[15] E.J. Sondik (1978). The optimal control of partially observable Markov decision processes over the infinite horizon: Discounted costs. *Operations Research*, 26:282-304.

[16] M.T.J. Spaan and N. Vlassis (2005). Perseus: Randomized Point-based Value Iteration for POMDPs. *Journal of Artificial Intelligence Research*, 24:195-220.