# Using a Bayesian-network Model for the Analysis of Clinical Time-series Data

Stefan Visscher[1], Peter Lucas[2], Karin Schurink[1], and Marc Bonten[1]

[1] Department of Internal Medicine and Infectious Diseases,
University Medical Centre Utrecht, The Netherlands
Email: {S.Visscher, K.Schurink, M.J.M.Bonten}@azu.nl
[2] Institute for Computing and Information Sciences,
Radboud University Nijmegen, The Netherlands
Email: peterl@cs.ru.nl

**Abstract.** Time is an essential element in the clinical management of patients as disease processes develop in time. Not surprisingly, the amount of temporal clinical information that is being collected to understand what is happening in the patient is therefore continuously increasing. A typical example of a disease process where time is considered important is the development of ventilator-associated pneumonia (VAP) during the stay in the ICU. A Bayesian network was developed previously to support clinicians in the diagnosis and treatment of VAP in the ICU. In the research described in this paper, we have investigated whether this Bayesian network can also be used to analyse the temporal data collected in the ICU for patterns indicating development of VAP. In addition, it was studied whether the Bayesian network was able to suggest appropriate antimicrobial treatment. A temporal database with over 17700 patient days was used for this purpose.

## 1 Introduction

Diagnosing disorders in critical care is a challenging task, as when a disorder is missed and thus treatment is delayed this may actually be the cause of death of a patient, as many of the patients in critical care are severely ill. Diagnosing a disorder may also be difficult if there are few signs and symptoms that are typical for the disorder, in particular if a disorder does not occur very frequently. This is the situation with ventilator-associated pneumonia, or VAP for short. VAP is a form of pneumonia that occurs in patients whom are mechanically ventilated in critical care units and involves infection of the lower respiratory tract. This has implications for the management of patients with VAP, as it is quite unlikely that doctors are able to identify this low-prevalence disorder reliably, and they may therefore need some form of decision support. As diagnosing a disorder in medicine involves reasoning with uncertainty, we used a Bayesian network as our primary tool for building a decision-support system. The development of VAP is a time-based process, involving colonisation of the patient by microorganisms. As diagnosing and monitoring a patient in critical care is a dynamic process

**Fig. 1.** Bayesian network for Ventilator-Associated Pneumonia.

involvinf uncertainty, we will analyse clinical time series for patients who were admitted to the ICU using a Bayesian network.

## 2   Bayesian networks in Biomedicine

Health care institutions collect large amounts of information about patients' state and interventions taken by physicians and nursing staff, such as therapies. The practical use of these 'rough' process data is still limited, since they are not collected in controlled clinical trials [1]. However, the interest has grown to analyse these data collections for management purposes. For understanding the complex decision-making processes in health-care it is important to reason about them. Since these processes often include uncertainty, Bayesian-network models are constructed to support decision-making in real-life practice as a Bayesian network is an excellent tool for reasoning with uncertainty [4].

Formally, a Bayesian network $\mathcal{B} = (G, \mathrm{Pr})$ is a directed acyclic graph $G = (V(G), A(G))$ with set of vertices $V(G) = \{V_1, \ldots, V_n\}$, representing stochastic variables, and a set of arcs $A(G) \subseteq V(G) \times V(G)$, representing statistical dependences and independences among the variables. On the set of stochastic

**Table 1.** Data description including mean values.

| Diagnosis | VAP $n = 157$ | no VAP $n = 17553$ |
|---|---|---|
| temperature | 38.7 $^\circ C$ | 37.9 $^\circ C$ |
| mech. ventilation | 271h = 11.3d | 268h $\approx$ 11.2d |
| leukocytes | 13.8 | 13.0 |
| pO$_2$/FiO$_2$ | 187 | 244 |
| sputum | 22.7 | 13.0 |
| sputum pur. | 74% | 35% |
| culture | positive | negative |
| antipyr. drugs | 82% | 61% |
| x-chest | positive | negative |

variables, a joint probability distribution $\Pr(V_1, \ldots, V_n)$ is defined that is factorised respecting the (in)dependences represented in the graph:

$$\Pr(V_1, \ldots, V_n) \prod_{i=1}^{n} \Pr(V_i \mid \pi(V_i))$$

where $\pi(V_i)$ stands for the variables corresponding to the parents of vertex $V_i$.

## 3 Patient Data

We had a temporal database of 17710 records at our disposal, each record representing data of a patient in the ICU during a period of 24 hours. The database concerned 2233 distinct patients, who were all together 2424 times admitted to the ICU. Thus, multiple admissions for one patient occurred. For 157 of the 17710 episodes, ventilator-associated pneumonia (VAP) was diagnosed by two infectious-disease specialists (See Table 1 for a description of the patients' data).

There is not a real gold standard for diagnosing ventilator-associated pneumonia. Important entities that are related to the development of VAP include *body temperature*, *drugs* to suppress fever, amount of *sputum*, *signs* on the chest X-ray, the ratio between the amount of oxygen in the arterial blood and the fractional inspired oxygen concentration, i.e. *pO$_2$/FiO$_2$*, the duration of *mechanical ventilation*, too few or too many *leukocytes* (white blood cells) and a *sputum culture* [5]. These symptoms involve the diagnostic part of the Bayesian network for VAP.

ICU patients become colonised by bacteria present in the ICU environment. When colonisation of the lower respiratory tract occurs within 2–4 days after intubation, this is usually caused by antibiotic-sensitive bacteria. When a VAP is developed after 4 days of intubation, often antibiotic-resistant bacteria are involved, which are more difficult to destroy. Therefore, it is important to start treatment as soon as possible when an infection is suspected [2].
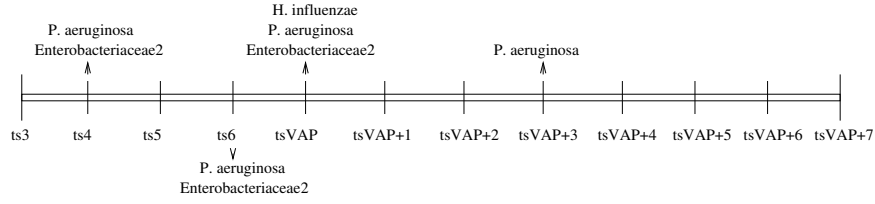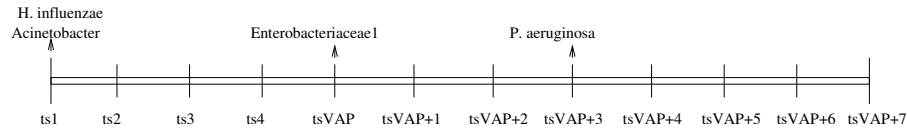
**Fig. 2.** Timeline for patient 2.

It has been stated that the duration of hospital stay and the severity of illness of a patient are associated with the prevalence of Gram-negative bacteria. We modelled the bacteria species each as one vertex in the Bayesian network (See Fig. 1). Also, for each modelled microorganism, the pathogenity, i.e. the influence of the microorganism on the development of VAP, was included in the model. The presence of certain bacteria has influence on the antimicrobial therapy. Each microorganism is susceptible to some particular antibiotics. Susceptibility, in this case, is stated as the sensitivity to or degree to which a microorganism is injured by treatment with a specific antibiotic. The susceptibility of each microorganism was taken into account while constructing the model. Our infectious-disease experts assigned utilities to each combination of microorganism(s) and antimicrobial drug(s) using a decision-theoretic model and estimated all conditional probabilities [3]. These variables belong to the therapeutic part of the Bayesian network for VAP.

## 4 The Role of Time in the Colonisation Process

The development of ventilator-associated pneumonia is a time-based process. By including the variable *mechanical ventilation* in the Bayesian network, we modelled the influence of time. Also, the node 'hospitalisation' represents a certain notion of time. The first 48 hours a patient is mechanically ventilated, the probability of colonisation is low. After a period of two days, this probability is higher and as a result, the probability of developing a VAP is also higher. Our database represents the situation for several patients from admission to dismission. It would not be very usefull to look at the daily data individually, since a physician does not work in this way either.



**Fig. 3.** Timeline for patient 3.

We consider the period from admission to dismission of the patient as a *time series* $\langle X_t \rangle$, $t = 0, \ldots, n_p$, where $t = n_p$ is the time of dismission of patient $p$. For each record, we collected the output of the Bayesian network, i.e., the best possible antimicrobial treatment. This antimicrobial treatment is based on the utility, i.e. the preference - according to the experts, and the overall-susceptibility of the microorganisms, i.e. the coverage. The optimal coverage is computed by varying the therapy choise, and selecting the therapies with maximum expected utility. Prescribing antibiotics is a trade-off between maximising coverage and minimising broadness of spectrum. When a patient is known to be colonised with a particular microorganism, the antimicrobial spectrum should be as narrow as possible.

We randomly selected 8 patients with VAP. During the period of seven days from timepoint of diagnosis, the patient is treated with antibiotics.

For each of the 8 patients, a timeline was drawn which shows colonisation information and the period from the timepoint a VAP was concerned. As mentioned before, it is not clinically relevant to look at the data of one particilar patient day. To judge a patient's condition, we have to look at the data of more than one day, thus we considered the time window $[t, t + 7]$, also taking into account information about (previous) colonisation, since this has direct influence on the antimicrobial therapy. We took the *minimal* set of advised antibiotics as the Bayesian network's output. This advice was compared to actually prescribed antibiotics, and agreement of disagreement between the two was judged two infectious disease specialists.

For patients 2 and 3, the colonisation process is shown by a timeline in Fig. 2 and 3. The present microorganisms are depicted. The therapy advice of the Bayesian network and the antibiotic treatment that was prescribed are presented for each patient, together with the judgement of the infectious-disease experts, in Table 2.

**Table 2.** Bayesian network versus infectious-disease specialists.

| Patient | Physician | Bayesian Network | Judgement |
|---|---|---|---|
| 1 | cefpirom | meropenem | spectrum much too broad |
| 2 | cefpirom | clindamycin+aztreonam | not the same, but correct |
| 3 | augmentin, ceftriaxon | meropenem | spectrum too broad |
| 4 | cefpirom | meropenem | not the same, but correct |
| 5 | augmentin | meropenem | spectrum much too broad |
| 6 | pipcil, ciproxin | clindamycin+ciprofloxacin | not the same, but correct |
| 7 | tazocin | ceftriaxone | not the same, but correct |
| 8 | ceftriaxon | cotrimoxazol | not the same, but correct |

## 5  Conclusions and Future Work

Since the development of ventilator-associated pneumonia is a time-based processes involving uncertainty, we have analysed temporal patient data using a Bayesian network. In this paper, we have only analysed time-series data of 8 patient with VAP. One of the most important processes in VAP is the colonisation by microorganisms. The results show that for 3 of the 8 patients, the antibiotic spectrum was too broad. We noticed that the patients for whom a correct therapy advice was given by the Bayesian network, were colonised by P. aeruginosa.

## References

1. Riccardo Bellazzi, Christiana Larizza, Paolo Magni, and Roberto Bellazi. Temporal data mining for the quality assessment of hemodialysis services. *Artificial Intelligence in Medicine*, in press.
2. Michael T. Brenman, Farah Bahrani-Mougeot, Philip C. Fox, Thomas P. Kennedy, Sam Hopkins, Rcihard C. Boucher, and Peter B. Lockhart. The role of oral microbial colonization in ventilator-associated pneumonia. *Oral Medicine*, 98(6), 2004.
3. Peter J.F. Lucas, Nicolette C. de Bruijn, Karin Schurink, and Andy Hoepelman. A probabilistic and decision-theoretic approach to the management of infectious disease at the icu. *Artificial Intelligence in Medicine*, 19(3):251–279, 2000.
4. Peter J.F. Lucas, Linda C. van der Gaag, and Ameen Abu-Hanna. Bayesian networks in biomedicine and health-care. *Artificial Intelligence in Medicine*, 30(3):201–214, 2004.
5. Bonten MJM, Kollef MH, and Hall JB. Risk factors for ventilator-associated pneumonia: from epidemiology to patient management. *Clinical Infectious Diseases*, 38(8), 2004.