

# Improving Antibiotic Therapy of Ventilator Associated Pneumonia using a Probabilistic Approach\*

Nicolette de Bruijn and Peter Lucas  
Department of Computer Science, Utrecht University  
Padualaan 14, 3584 CH Utrecht, The Netherlands  
E-mail: {nbruijn,lucas}@cs.uu.nl

Karin Schurink and Andy Hoepelman  
Department of Internal Medicine, Utrecht University Hospital  
Heidelberglaan 100, 3584 CX Utrecht, The Netherlands

## Abstract

PTA is a decision-theoretic expert system that aims to assist clinicians in diagnosing and treating patients with pneumonia at the intensive care unit. The underlying probabilistic network model includes knowledge for diagnosing pneumonia on the basis of the likelihood of tracheobronchial-tree colonisation by pathogens, and symptoms and signs actually present in the patient. Optimal antibiotic therapy is selected by balancing the expected efficacy of treatment, which is related to the likelihood of pathogen-specific pneumonias, against costs and side effects of treatment. In this article, the structure of the system and results of a preliminary evaluation are described.

*Keywords:* medical decision support, decision-theoretic expert systems, probabilistic networks, Bayesian belief networks.

## 1 Introduction

The medical community is presently in a state of transition from a situation dominated by the paper medical record, containing most clinical patient data, to a future situation where all patient data will be available on-line by an electronic clinical information system. In many hospitals, laboratory data have already been stored and distributed by a hospital information system for some time. In technically more advanced hospitals, a number of departments have their clinical patient data already fully managed by a local clinical information system, usually supplemented with a traditional hospital-wide information system. This typically holds for departments in which the amount of data collected for a patient is huge, such as intensive care units (ICUs).

However, providing facilities to store and retrieve patient data to the clinician is not enough. Clinical information systems should also offer facilities to assist clinicians in dealing

---

\*Published in: P. Lokol, B. Zupan, J. Stare, M. Premik, R. Engelbrecht (eds.), *Medical Informatics Europe '99*, IOS Press, Amsterdam, pp. 690–695.

with hard clinical problems, such as facilities for decision support. In this paper, we describe the development of a decision-theoretic expert system [4], i.e. a system based on a combination of the theory of probabilistic networks (Bayesian belief networks) and decision theory [6, 7], called *PTA* (Pneumonia Therapy Advisor), that is capable of providing advice about the administration of appropriate antibiotic therapy to patients with pneumonia. The aim of the development of this system is the provision of decision support concerning the treatment of serious infectious diseases as an extension to the current clinical information system at our ICU.

The structure of this paper is as follows. In the next section, the problem of the management of patients with pneumonia is brought in clinical perspective. In Section 3, we describe the pneumonia model that is used as a knowledge base of the decision-theoretic expert system. In Section 4, we present results of an evaluation of the system, and, finally, in Section 5 it is discussed what has been achieved by our research.

## 2 Management of pneumonia at the ICU

Pneumonia is a frequently occurring clinical problem at the ICU. Many patients admitted to an ICU need respiratory support by a mechanical ventilator; in addition, many of these patients are affected by severe disease which may result in depression of their immune system. Both conditions promote the development of bacterial pneumonia in these patients.

Usually, a distinction is made between *community-acquired pneumonia* (CAP), which is pneumonia originating outside the hospital, *hospital-acquired pneumonia* (HAP), which develops during a stay at the hospital, and *ventilator-associated pneumonia* (VAP), which may arise in patients whom are mechanically ventilated [1]. Mechanically ventilated patients are intubated, which appears to be a major factor promoting the development of this type of infection. Since CAP is caused by pathogens outside the clinic, common antibiotic therapy is usually effective; there is often no need to send the patient to the hospital. This disease can be handled by primary care physicians. The situation with HAP and VAP is quite different. The occurrence of HAP or VAP in a patient during a stay at the hospital is seen as a significant problem due to the presence of multiresistant bacteria in clinical wards, in particular at the ICU. These bacteria are likely to be the cause of infection, but it is often impossible to eradicate these bacteria using standard antibiotic treatment regimes, i.e. antibiotic therapy that is usually prescribed when the chances for multiresistance are small.

Choosing an appropriate therapy not only involves the issue of susceptibility of pathogens to antibiotic drugs, but also of possible side effects and of future development of resistance to drugs. In the case of antibiotic therapy possible side effects are: renal failure, diminished hearing, epileptic seizures and allergic reactions varying from skin rash to anaphylactic shock. Clearly, the decisions about appropriate antibiotic therapy must be made on the grounds of a lot of uncertain medical knowledge. In particular, making decisions concerning the initial therapy, called *empirical therapy*, of these patients is difficult because it takes at least 48 hours before the results of sputum cultures become available. Sputum cultures yield information of the identity and antibiotic susceptibility of pathogens. Hence, empirical antibiotic therapy, the subject on which *PTA* focusses, is usually prescribed without actually knowing the identity of the causative pathogens.

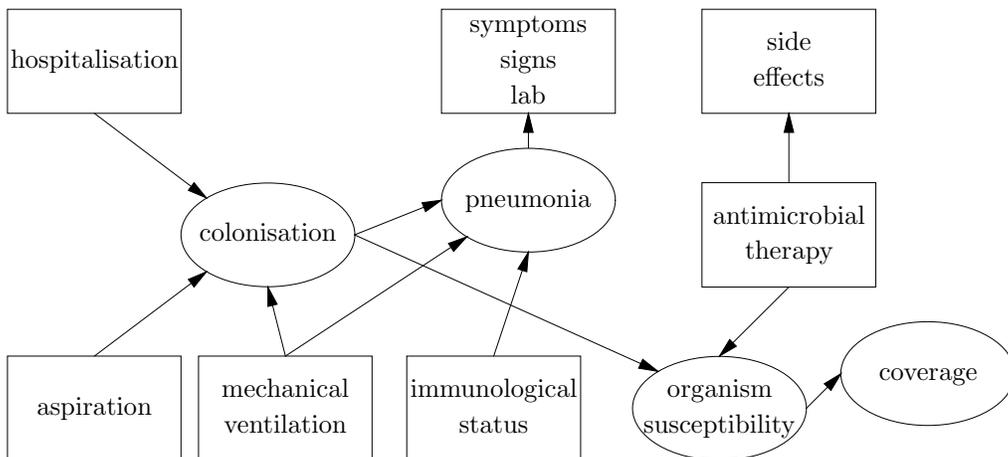


Figure 1: Global structure of probabilistic model.

### 3 A probabilistic model of pneumonia

#### 3.1 The probabilistic network formalism

A natural representation of the uncertainties involved in dealing with treatment of pneumonia is offered by the probabilistic-network formalism [5, 6]. A *probabilistic network* is an acyclic directed graph  $G = (V(G), A(G))$ , with a set of vertices  $V(G) = \{V_1, \dots, V_n\}$ , where each vertex  $V_i \in V(G)$  represents a discrete stochastic variable, and a set of arcs  $A(G) \subseteq V(G) \times V(G)$ , where each arc  $(V_i, V_j) \in A(G)$  represents stochastic dependence. Arcs are often informally seen as to mirror causal or correlational influences among variables. On the set of variables  $\{V_1, \dots, V_n\}$  is defined a joint probability distribution  $\Pr$  that can be factorised according to the topology of the graph as follows:

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

where  $\pi(V_i)$  represents the set of variables associated with the parent vertices of  $V_i$ . This means that the joint probability distribution  $\Pr(V_1, \dots, V_n)$  can be defined in terms of ‘local’ probability tables  $\Pr(V_i | \pi(V_i))$  by assuming the variable  $V_i$  to be conditionally independent of all predecessors of the associated vertex  $V_i$  given the parents  $\pi(V_i)$ . Hence, the given parents shield a variable from stochastic influences of the other predecessors; the entire acyclic directed graph reflects all (un)conditional independencies among the corresponding variables [6].

#### 3.2 VAP model

The probabilistic network model was developed with the help of a number of infectious disease experts and by consultation of the medical literature. Since the diagnosis and treatment of VAP appeared clinically most significant, the model presented here is restricted to VAP. However, extension of the model towards HAP and CAP would have little effect on the global structure of the present model. Figure 1 gives an overview of the structure of the model. Boxes indicate processes that can be (partially) observed, or decisions (selection of therapy) that can be made by the clinician at the time of empirical therapy; ellipses indicate processes that cannot be observed at the time of empirical therapy selection.

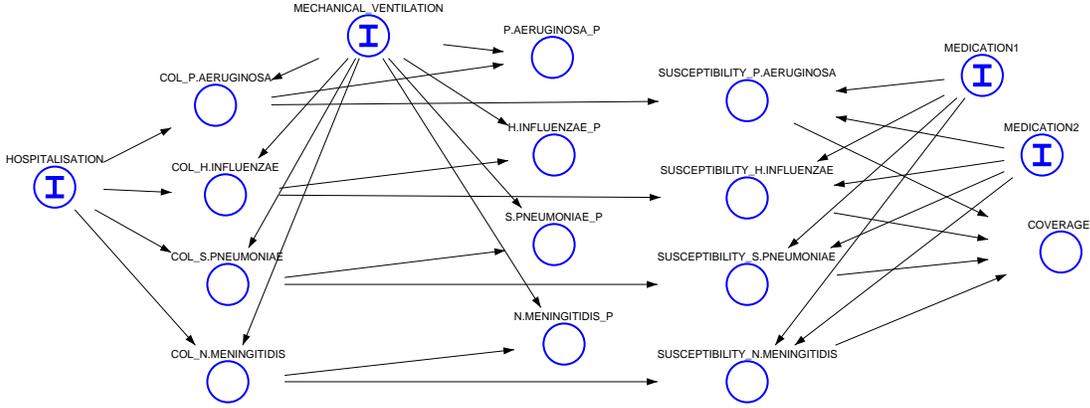


Figure 2: Part of a probabilistic VAP model.

Central to the model is the temporal process of *colonisation* of the laryngotracheobronchial tree by pathogens. The temporal nature of the process is expressed by the interaction between duration of stay at the ICU (*hospitalisation*) and the duration of *mechanical ventilation*: both duration of stay at the ICU and duration of mechanical ventilation are positively correlated to colonisation by pathogens. *Aspiration* of stomach content is another factor positively correlated to colonisation by certain pathogens (*Enterobacteriaceae*). When a patient is colonised by specific pathogens, there is a certain probability that *pneumonia* will develop. This is expressed in the graph by an arc from *colonisation* to *pneumonia*. Duration of mechanical ventilation and immunological status influence the probability that pneumonia develops as well; hence, an arc from these vertices to the *pneumonia* vertex is included in the model. When the patient is affected by pneumonia, certain symptoms, signs or laboratory abnormalities can be observed, as summarised in a corresponding vertex in the graph. Finally, the *susceptibility* of pathogens to particular antibiotic treatment is determined by the choice of medical treatment and the actually present pathogens causing infection. Some antibiotic drugs may give rise to certain side effects.

Some of the vertices shown in the graph in Figure 1 actually comprise a number of separate vertices. For example, colonisation by pathogens was modelled as a biological process, in which it was assumed that pathogens occur independently. Given this assumption, the processes of colonisation and development of pneumonia by each possible pathogen have been modelled by separate vertices, with corresponding stochastic variables, in the probabilistic model. Note that this representation allows for the presence in the patient of a pulmonary infection due to *multiple organisms*; if, in contrast, colonisation (pneumonia) by different pathogens had been represented as values of a single stochastic variable, they would have been assumed mutually exclusive, which obviously is incorrect. Colonisation by 11 of the most frequently occurring pathogens, such as *Pseudomonas aeruginosa* and *Neisseria meningitidis*, is represented in the current model. In Figure 2, a small part of the model, only involving the vertices *hospitalisation*, *mechanical ventilation*, *colonisation* (e.g. COL.P.AERUGINOSA, i.e. colonisation by *Pseudomonas aeruginosa*), and *pneumonia* (e.g. P.AERUGINOSA Pneumonia), *susceptibility* of pathogens to antimicrobial drugs, and *medication* is shown, restricted to 4 of the 11 included pathogens. Vertices with capital I are *instantiation vertices* (see below).

The probabilistic model has been implemented using the Ideal probabilistic expert-system

shell [8]. Figure 3 shows the complete model as seen by a user of the current version of the system.

Antibiotic treatment consists of the selection of one or two different antibiotics – possibly also none – modelled by two identical therapy vertices; each therapy vertex includes 24 different antibiotic drugs (including none), yielding  $24^2 = 576$  possible combinations, of which  $\binom{23}{2} = 253$  (excluding ‘none’) are different, yielding a total of  $253 + 24 = 277$  (now including ‘none’ and single drugs) different therapies.

### 3.3 Decision-making capabilities

The efficacy of treatment for a patient can be examined by entering observed symptoms and signs in the patient, duration of hospitalisation, mechanical ventilation, previously experienced side effects of antibiotic therapy, and laboratory data. Treatment selection is flexible in the sense that only data concerning vertices that are denoted as instantiation vertices must be supplied always; although entering more patient data may provide additional evidence concerning the cause of pneumonia, the model can be used even when most other findings are unknown. Using the probabilistic inference algorithm as described in [3], the probability distribution defined on the *coverage* variable can be updated, yielding an a posteriori probability of coverage of all possible pathogens,  $\Pr^*$ , reflecting knowledge concerning all entered evidence (*Evidence*), i.e.

$$\Pr^*(coverage) = \Pr(coverage|Evidence)$$

By indicating particular vertices without parents as instantiation vertices, meaning that a value for the corresponding variable must always be supplied, yielding additional stochastic independence information, the speed of the probabilistic inference could be dramatically improved. By defining a utility measure in terms of coverage, it is possible, in principle, to automatically determine the treatment that gives optimal coverage by computation of the maximum expected utility. By including the side effect variables, it is also possible to automatically determine the antibiotic treatment that balances maximum coverage against minimum side effects. However, since the number of possible drug combinations was 576, computation of optimal therapy is practically infeasible, because even probabilistic inference for one antibiotic drug combination takes seconds to minutes.

As a practical solution to this problem, we have restricted the 277 different therapies to the 32 different therapies considered adequate for most patients, and represented in a single vertex.

## 4 Evaluation

### 4.1 Soundness of the structure of the model

The structure of the probabilistic model was designed, taking colonisation as a starting point. The structure of the model reflects evolution of hospital-acquired infectious disease, starting with hospitalisation and mechanical ventilation, and finally resulting in pneumonia and therapeutic intervention. Hence, the present structure of the network has a strong logical foundation, and we, therefore, believe it to be basically correct. There may be particular arcs missing due to gaps in the medical knowledge concerning pneumonia. However, correlations



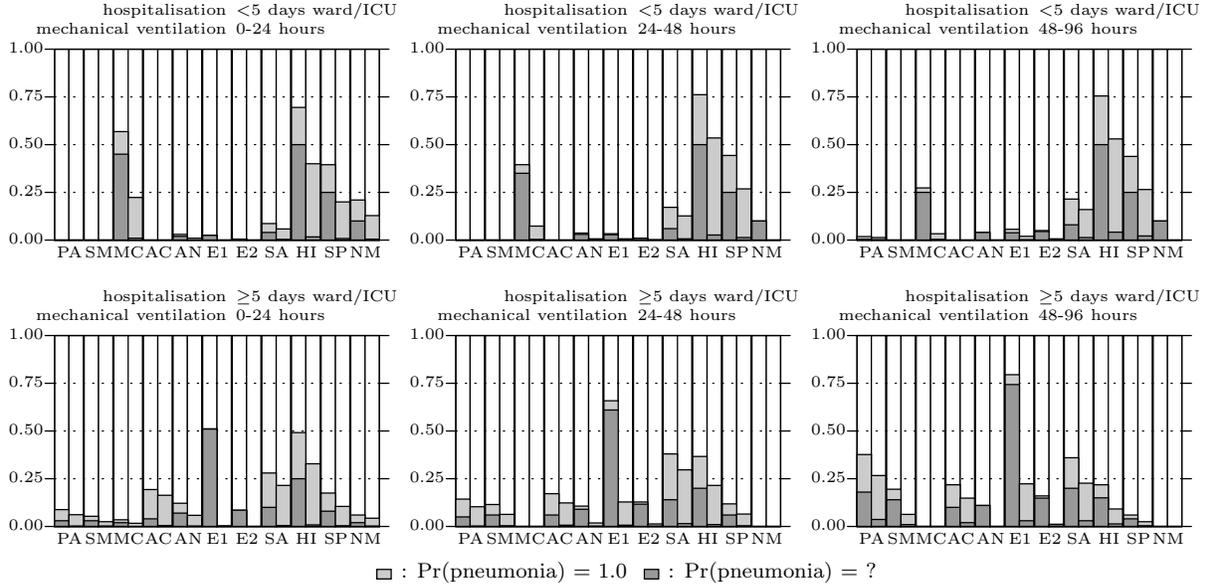


Figure 4: Obtained predictions after entering information concerning duration of hospitalisation and mechanical ventilation. Names of pathogens have been abbreviated. For each pathogen, the probability of colonisation and pneumonia have been depicted, in that order.

due to missing medical knowledge would likely to be weak, and hence have little effect on the probability distribution.

## 4.2 Accuracy of probabilistic information

In addition to the structure of a probabilistic network, the accuracy of the represented probabilistic information is an issue that also requires attention. This has been investigated in two ways. Firstly, a number of simulation experiments was carried out, and secondly the model's conclusions were evaluated for a number of different patients.

Some of the vertices in the model were instantiated (i.e. a specific value was chosen), and the resulting a posteriori probability distribution was compared with frequency information available from the ICU. The results of these experiments are shown in Figure 4.

It appears that the patterns of frequency of colonisation and pneumonia for particular pathogens changes with duration of stay at the hospital. For example, the relative frequency of colonisation and pneumonia by *Moraxella Catharrhalis* (MC in the figure), which is high when entering the hospital, decreases during the first 96 hours of mechanical ventilation (upper three graphs), to become nearly absent when the patient has stayed more than five days in hospital. This frequency decreases to nearly zero for patients whom are mechanically ventilated during the next 96 hours (lower graph). These observations did agree with expert opinion.

Furthermore, the conclusions obtained by the model were examined for a group of 12 patients, drawn from the files of the ICU. For these patients, the system selected the therapy that best covered suspected pathogens. Similarly, a specialist in infectious disease was requested to select the best antibiotic therapy, and also to assess recommendations made by

n	Hosp	Vent	Extra Info	Expert Advice	Diff	PTA	Assessment	C
1	<5days	24-48	colonisation HI	Augmentin	3/21	Aug+Gen	acceptable	>
2	<5days	48-96	colonisation HI+SA	Augmentin	8/19	Aug+Gen	acceptable	>
3	<5days	48-96	colonisation SA	Augmentin	8/21	Aug+Gen	acceptable	>
4	<5days	48-96	-	Augmentin	3/21	Aug+Gen	acceptable	>
5	<5days	48-96	-	Augmentin	4/22	Aug+Gen	2nd best	>
6	<5days	48-96	-	Augmentin	5/21	Aug+Gen	acceptable	>
7	<5days	48-96	-	Augmentin	4/23	Aug+Gen	acceptable	>
8	<5days	24-48	COPD	Cefta+Tobra	5/23	Aug+Gen	2nd best	<
9	$\geq 5$ dW/IC	0-24	aspiration	Aug+Gen	7/28	Tazocin	acceptable	>
10	$\geq 5$ dIC	96-144	-	Cefta+Tobra	8/24	Tazocin	2nd best	=
11	$\geq 5$ dIC	>144	phagocytys dys	Cefta+Tobra	7/19	Tazocin	2nd best	=
12	$\geq 5$ dIC	>144	colonisation E1+E2	Imipenem	1/12	Imipenem	acceptable	=

Table 1: Results for 12 patients. Meaning of abbreviations: Hosp: Hospitalisation stay and location;  $\geq 5$ daysW/IC:  $\geq 5$  days in hospital (at ward or ICU); Vent: duration of mechanical ventilation in hours; colonisation HI: colonisation by *H. influenzae*; Diff  $x/y$ : position  $x$  out of  $y$  different groups of drugs, ordered according to probability (from high to low) and grouped when probabilities were equal; Aug+Gen: Augmentin + Gentamycin; acceptable: advice of PTA is acceptable according to the expert; 2nd best: advice of PTA is second best choice of expert; C: coverage of recommendation by PTA; <: less broad spectrum; =: similar spectrum; >: broader spectrum.

the system as being either acceptable or second best choice. The results are shown in Table 1. It appears that in 8 cases the model prescribed a therapy that covers more pathogens than the therapy prescribed by the expert. This is caused by the fact that the present model does not take broadness of antimicrobial spectrum into account. All 12 prescriptions were considered acceptable or second best choice; none of the recommendations was considered unacceptable.

## 5 Discussion

Above, we have discussed the development of a decision-theoretic expert system that assists in exploring the diagnosis and treatment of ventilator-associated pneumonia. The knowledge incorporated in the model has been gathered from various sources: medical specialists, data extracted from the hospital information systems and from the local clinical information system at the ICU.

In the near future, we will design additional utility models, for example models that take into account antimicrobial spectrum and financial costs of antibiotics, to obtain a system that balances different costs and benefits of antibiotic drugs, to reach optimal treatment. We intend to embed the resulting system in the clinical information system of the ICU.

## References

- [1] J.G. Bartlett, *Management of Respiratory Tract Infections*, Williams & Wilkins, Baltimore, 1997.
- [2] M. Korver and P.J.F. Lucas, Converting a rule-based expert system into a belief network, *Medical Informatics*, 18(3) (1993) 219–241.
- [3] S.L. Lauritzen and D.J. Spiegelhalter, Local computations with probabilities on graphical structures and their application to expert systems, *Journal of the Royal Statistical Society (Series B)*, 50 (1987) 157–224.
- [4] P.J.F. Lucas, H. Boot and B.G. Taal, Computer-based decision-support in the management of primary gastric non-Hodgkin lymphoma, *Methods of Information in Medicine* 37 (1998) 206–219.
- [5] P.J.F. Lucas and L.C. van der Gaag, *Principles of Expert Systems*, Addison-Wesley, Wokingham, 1991.
- [6] J. Pearl, *Probabilistic Reasoning in Intelligent Systems*, Morgan Kaufman, San Mateo, California, 1988.
- [7] J.Q. Smith, *Decision Analysis*, Chapman and Hall, London, 1988.
- [8] S. Srinivas and J. Breese, *IDEAL: Influence Diagram Evaluation and Analysis in Lisp: Documentation and Users Guide*, Rockwell International Science Center, Palo Alto, 1993.