

A Bayesian Approach to Model-Development: Design of Continuous Distributions for Infection Variables

Logan Ward* Mads L. Mogensen** Mical Paul*** Leonard Leibovici*** Steen Andreassen*

*Centre for Model-based Medical Decision Support, Aalborg University, Aalborg, Denmark

**Treat Systems A/S, Aalborg, Denmark

***Sackler Faculty of Medicine, Tel-Aviv Univeristy, Tel-Aviv, Israel

Abstract: Bayesian networks can be used to build models of diseases for diagnosis, and, if complemented by decision theory and utility functions, can also suggest treatments. This paper presents a development framework for such a network that has been used to model sepsis and in particular focuses on incorporating knowledge from the literature and databases as well as expert opinion into the model. Two parameters are presented as examples of the methods used, and the model is validated for a cohort of patients suspected of infection.

Key Words: Model formulation, experiment design; Decision support and control; Quantification of physiological parameters for diagnosis and treatment assessment

1. INTRODUCTION

Bayesian networks can be used to build models of diseases for diagnosis (Alvarez et al. 2006, Sadeghi et al. 2006, Schurink et al. 2007, Kariv et al. 2011, Andreassen et al. 1996). If complemented by decision theory and associated utility functions, they can also be used to suggest treatments (Hejlesen et al. 1997, Andreassen et al. 1999, Leibovici et al. 2000).

A Bayesian network consists of nodes, representing stochastic variables, and arrows linking these nodes together. To provide plausible reasoning, these arrows must follow the flow of causality (Pearl 1988), which is why they are also referred to as Causal Probabilistic Networks (Andreassen et al. 1991). Numerically the arrows hold conditional probabilities and the task of constructing a CPN consists of specifying the graph of nodes and arrows and the associated conditional probabilities. In the CPN, some of the nodes will represent variables, such as diagnoses, which are of interest but unobservable and other variables, like body temperature or blood pressure, which are observable. The task of the CPN, once it is constructed, is to modify the probability distributions for the unobservable variables in accordance with the axioms of probability theory, when evidence, i.e. values for the observable variables, is inserted into the CPN.

During the construction of a CPN the conditional probabilities themselves may be considered stochastic variables. The initial specification of these variables can be seen as the specification of a priori distributions for the variables. These a priori distributions may come from any source of evidence, including expert opinion and knowledge derived from textbooks or the scientific literature. This is at best a semiformal approach, but its value has been

demonstrated empirically, for example through the clinical performance of the Treat decision support system (Paul et al. 2006a, Paul et al. 2006b, Paul et al. 2007, Leibovici et al. 2007, Leibovici et al. 2010), which was constructed this way (Andreassen et al. 1999, Leibovici et al. 2000, Andreassen et al. 2005). Treat provides decision support for diagnosis and antibiotic treatment of acute infections.

The Treat CPN is large with about 6000 nodes. A small part of the CPN, about 40 nodes, deals with the assessment of the patients' degree of illness, expressed on a sepsis scale going from "None" through "Mild", "Moderate" and "Severe" to "Critical". In the first version (Leibovici et al. 2000) the observable nodes were discrete stochastic variables, and accordingly we call this CPN the Discrete Sepsis CPN (D-Sepsis CPN). The structure of this CPN was largely determined by performing a factor analysis on a database of patients with different degrees of sepsis (Fig. 1, Phase I) and the conditional probabilities were based on expert and literature based opinions. The probability of a patient being bacteraemic is dependent on the patient's degree of sepsis and the clinical success of the D-Sepsis CPN was demonstrated by showing that it could provide a fair prediction of bacteraemia with an area under the ROC curve of 0.70 (Paul et al. 2006a).

Despite the clinical success of the D-Sepsis CPN it had some suboptimal properties. For example, temperature was represented by a discrete stochastic node with 5 states: <36.0, 36.0-36.5, 36.6-37.9, 38.0-38.6 and >38.6. This fairly crude discretization gave rise to undesirable jumps in the assessment of patient state, for example if the patient's temperature increases from 38.6 °C to 38.7 °C. This motivated a revision of the sepsis CPN into a Continuous

Sepsis CPN (C-Sepsis CPN) using continuous rather than discrete conditional probability distributions. Hugin (Hugin Expert A/S, Aalborg, Denmark), a tool for specification of CPNs, offers such a facility, provided the distributions are Gaussian. This revision uses the discrete sepsis CPN as its set of a priori conditional probability distributions to be modified (or learned). The revision itself is mainly based on expert and literature opinions (Fig. 1, Phase II) and thus its method of construction is essentially identical to the method used for the D-Sepsis CPN. The purpose of this paper is to provide insight into this semiformal process of construction.

Eventually these informal “manual learning” methods may be supplemented by learning from case databases, where methods such as Dirichlet learning or expectation-maximisation (EM) learning (Lauritzen 1995, Spiegelhalter, Lauritzen 1990) are employed (Fig. 1, Phase III). This is, however, outside the scope of this paper.

As examples of the method of development, ‘manually learned’ distributions for two of the observable variables in the sepsis CPN are included. The observable variables are C-Reactive Protein (CRP) and plasma albumin. They serve as extremes in terms of knowledge, because values of CRP during sepsis have been reported extensively in the literature, while literature on albumin during sepsis is sparser. We then validate the manual learning approach by testing the ability of Treat, including the continuous sepsis CPN, to predict bacteraemia.

2. METHODS

The purpose of the paper is to describe how the D-Sepsis CPN can be modified into the C-Sepsis CPN (Fig. 1, Phase II). For this purpose the D-Sepsis CPN serves as the specification of the a priori conditional probability distribution for the observable variables, given the state of sepsis.

The first step in the process is to apply expert knowledge to the structure of the CPN, adding or deleting nodes and rearranging arrows.

The next step is to make a literature search to locate papers containing information about these conditional distributions. A search of PubMed was performed using MeSH terms, followed by a staged elimination based on paper title, abstract and the full text. Information was extracted from the

remaining papers and an informal meta-analysis was attempted.

Finally the numerical results from the meta-analysis were modified in the light of the a priori distributions, i.e the D-Sepsis CPN.

The outcome of this process, i.e. the C-Sepsis CPN was then validated by comparing the ability to predict bacteraemia of Treat with the D-Sepsis CPN with the ability of Treat with the C-Sepsis CPN. This ability was benchmarked by the ROC curve for bacteraemia prediction. Of the two observable variables associated with severity we have available, bacteraemia and mortality, bacteraemia is chosen because of the greater number of cases, and consequently greater chance of showing a significant result.

3. RESULTS

To accomplish the first step (see Methods) a number of structural changes were made to modify the D-sepsis CPN into the C-Sepsis CPN (Fig. 2). Observable nodes were added for heart rate and oliguria to better model cardiovascular and renal failure, respectively. “Modifier” nodes were also added: these nodes are parents of observable nodes that modify the inference we make about sepsis from the observation made. For example, we make no inference about sepsis from the heart rate if the patient is receiving beta-blockers. Modifier parents were also added to the blood pressure node, where “age” was added because we expect higher blood pressure in older patients, and fluid resuscitation was added because it may partially alleviate reduced blood pressure. In addition, we chose to explicitly model the non-infectious aetiologies of the systemic inflammatory response syndrome (SIRS). This is accomplished through the addition of a diagnosis node for “other_SIRS” and a set of mapping nodes. These mapping nodes are a set of additional factors that define the distribution of severities for non-infectious causes of inflammation. Together with the other_SIRS node, they provide an alternative explanation for the observations in the sepsis CPN, which then reduces the strength of the evidence provided by the observations.

Previously non-infectious SIRS was modelled by small leaks in the conditional probability tables of the factor nodes. The section of the D-Sepsis CPN concerning respiratory effects was also remodelled to reflect that the lungs are affected both

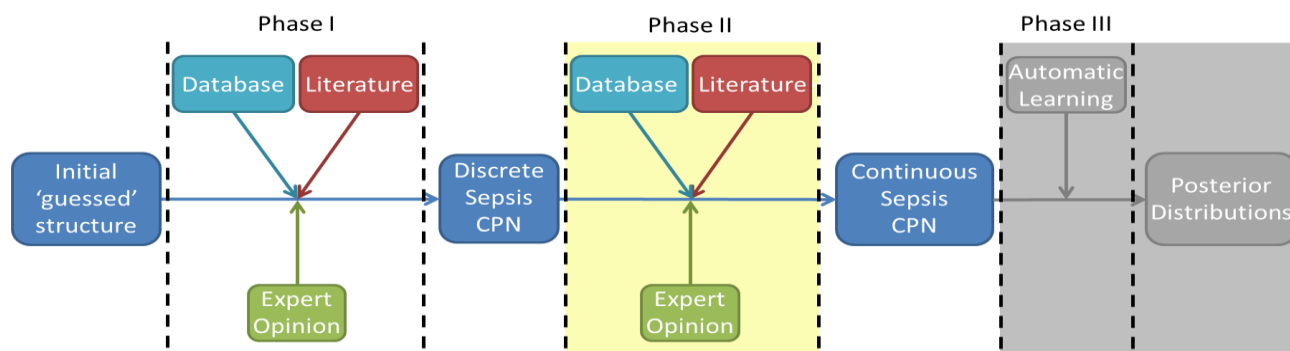


Fig. 1 Sepsis CPN development framework. Phase I describes the development process of the discrete sepsis CPN, phase II the development of the continuous sepsis CPN - the subject of this article, and phase III the future development of the sepsis CPN through formal learning methods.

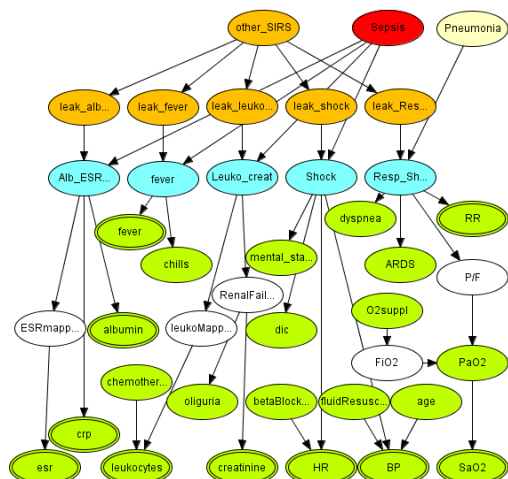


Fig. 2 The continuous sepsis CPN showing the nodes (ovals) and causal links (arrows). The colouring corresponds to the node type; red is the sepsis severity node, blue is an intermediate factor node, green defines an observed (or observable) node, orange the nodes associated with non-infectious SIRS, and white non-observed or unobservable nodes. Nodes with a double-ring are continuous nodes.

by pneumonia and by sepsis. All of these modifications were made based on expert opinion.

The second step (see Methods) in the development of the C-Sepsis CPN was to conduct a literature search and define distributions for individual parameters with respect to sepsis severity. For the purpose of this paper, two extreme cases are taken: one parameter which has been researched extensively as a sepsis biomarker, CRP, and one for which little literature is available, albumin.

3.1 C-Reactive Protein

CRP is an acute-phase protein, with rapid increases associated with the onset of infection. Although CRP levels are raised in the inflammatory condition in general, non-infectious causes of inflammation such as burns, trauma and surgery are usually associated with a lesser increase. CRP has been studied extensively as a potential biomarker for sepsis (Pierrakos, Vincent 2010). The normal range for CRP is well established, with serum concentrations >10 mg/L strongly indicative of an on-going pathological process (Vigushin et al. 1993). Attempts have been made to determine the validity of using CRP both as a marker to differentiate between those with and without an infection, the aetiology of the infection and the severity of illness for both septic and aseptic patients (sepsis, severe sepsis, septic shock and the non-infectious equivalents). Ideally, we would like to use this collection of knowledge about CRP to form a two-dimensional table of CRP distributions, with aetiology (non-infectious, viral, bacterial) on one axis, and severity (for example, mild, moderate, severe, critical) on the other. The severity gradations should also follow the consensus definitions for sepsis (Bone et al. 1992, Levy et al. 2003).

To identify a suitable set of literature, a systematic, although not exhaustive search was performed. Using a Boolean

strategy, and MeSH Terms, the PubMed database was searched using the query:

((c reactive protein[MeSH Terms] OR "CRP" OR "C-reactive protein") AND humans[MeSH Terms]) AND (adult[MeSH Terms] OR aged[MeSH Terms]) AND ("sepsis"[Title/Abstract] OR "SIRS"[Title/Abstract] OR "infection"[Title/Abstract]) AND ("Critical care" OR "icu" OR "hospital" OR "emergency department")

The search returned 1654 results. Inspection of titles and, if deemed necessary, abstracts reduced the set of results to a final list of articles for which full-text was inspected. One further article was identified through inspection of the reference lists of the articles for which full-text was inspected. A lack of data for subjects without infection or inflammation led to an additional search where three extra papers were found relating to healthy subjects to give a final list of 34 articles. 14 of these articles were then eliminated because they either failed to grade sepsis by severity, or the statistics were poor, that is the figures provided could not be converted to means and standard deviations.

Meta-analysis was then performed on distributions grouped according to nine classifications: 4 infectious; viral infection, sepsis, severe sepsis, and septic shock (Fig. 3, top panel) and 5 non-infectious; healthy, no SIRS (patients who may be ill, but do not meet the SIRS criteria), SIRS, severe SIRS and non-septic shock (Fig. 3, bottom panel). Median and interquartile range or median and range were converted to mean and standard deviation to allow pooling of data (Hozo et al. 2005). Initial inspection of the reported means and standard deviations suggested that log-normal distributions would be more appropriate than regular normal distributions. Data was converted into log-normal means and standard deviations (Higgins et al. 2008), and pooled together in a meta-analysis (Rudmin 2010).

The C-Sepsis CPN defines five severity states- no (healthy/normal patients), mild, moderate, severe and critical. Each CRP distribution should then be mapped to one of these states. Fig. 4 shows which distributions from Fig. 3 are

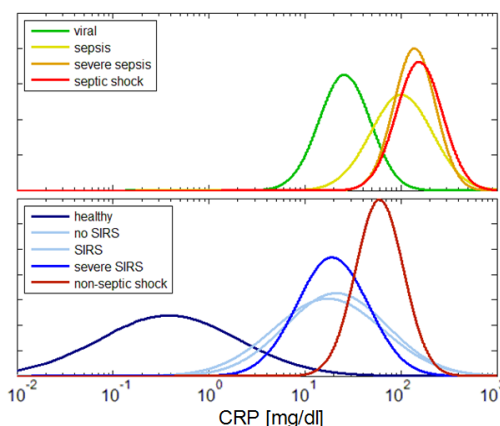


Fig. 3 Log-normal distributions for 4 infected (top panel) and 5 non-infected (bottom panel) patient groups. Individual distributions are the result of meta-analysis of literature studies, with reported distributions converted to log-distributions as required.

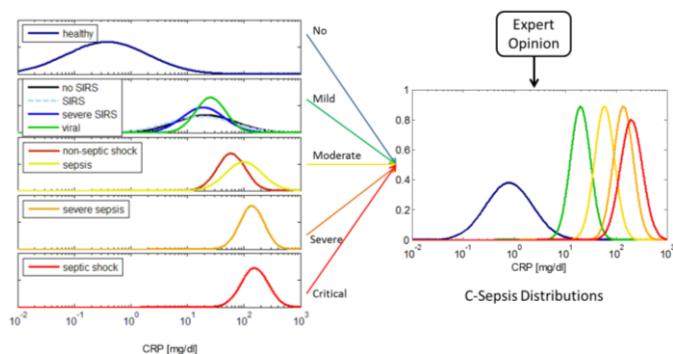


Fig. 4 Distributions from Fig. 3 are grouped together (literature distributions, left hand side) and assigned to one of the severity states as defined in the continuous sepsis CPN; no, mild, moderate, severe, and critical. The distributions were then modified based on expert opinion to give those that are used in the C-Sepsis CPN (right hand side).

grouped together, and mapped to each severity state (left hand panel) and the resulting C-Sepsis CPN distributions after input from experts (right hand panel).

The distributions used in the C-Sepsis CPN were based on the combinations of literature distributions and on the expert opinions of our clinical partners, which include the distributions in the D-Sepsis CPN. The mild distribution in the C-Sepsis CPN is an example of a modification made to the literature distributions. We define this distribution as covering viral sepsis, and other mild infections such as atypical pneumonias for which data was lacking, as well as several non-infectious aetiologies. These distributions, although formed from studies with relatively few cases, serve the purpose of defining an approximate mean and range to be covered by the resultant distribution and the C-Sepsis distribution (green) was chosen to cover this range.

Another expert-motivated decision was taken in forming the moderate curve. For this curve, we group non-septic shock patients with 'sepsis' patients. We would like to have that this state is used for patients with uncomplicated sepsis, that is, those that have not progressed to severe sepsis or septic shock. However, two problems arise here; not every paper explicitly defines sepsis diagnosis in this way according to the ACCP criteria, and the fact that the ACCP consensus definitions for sepsis have changed (Bone et al. 1992, Levy et al. 2003). This led to the decision that the moderate distribution should be shifted slightly lower than that formed from the literature. Coincidentally, this closely resembles the literature curve for non-septic shock.

It is important that there is significant overlap between the septic distributions because of the time taken for the CRP level to reach its peak, usually 36-50 hours after the original insult (Vigushin et al. 1993). Some studies have also suggested that the degree of increase of CRP does not reflect the severity of illness (Mitaka 2005). CRP typically has a maximum of 300-500 mg/l which can be reached even in non-severe cases, meaning that for individual patients, the sensitivity of CRP in differentiating between stages of sepsis is limited.

3.2 Albumin

Hypoalbuminemia is a commonly recognised marker of injury severity. Serum albumin concentration can be modified by several mechanisms; a change in production of albumin, a change in albumin consumption/degradation or a change in vascular permeability, allowing albumin to distribute from plasma into the interstitium. The rate of albumin synthesis and catabolism is many times smaller than the normal rate of albumin diffusion to the interstitium (5%/h), and is too slow to be responsible for the rapid drop in serum albumin seen in sepsis and other critical illnesses (Fleck et al. 1985), leaving capillary leakage as the accepted mechanism by which levels are reduced. A meta-analysis found serum albumin to be a dose-dependent predictor of poor outcome, with significant increases in mortality, morbidity and length of stay for each decrease of 10 g/L (Vincent et al. 2003).

Using a similar search procedure to that used for CRP, the following query was entered into the PubMed search facility:

```
(albumin, serum[MeSH Terms] AND humans[MeSH Terms]) AND (adult[MeSH Terms] OR aged[MeSH Terms]) AND ("sepsis"[Title/Abstract] OR "SIRS"[Title/Abstract] OR "infection"[Title/Abstract]) AND ("Critical care" OR "icu" OR "hospital" OR "emergency department")
```

The search returned 244 results, all of which were rejected for the purpose of defining explicit distributions for albumin stratified by sepsis severity and/or aetiology. Serum albumin has not, unlike CRP, been evaluated as a biomarker for sepsis severity, so there is a notable absence of literature in which measurements of albumin concentrations are recorded in relation to the stages of sepsis. The normal distribution used in our network has been taken from a large study of albumin as a predictor for mortality (Corti et al. 1994).

To design the severity-based distributions, data from the Treat study (Paul et al. 2006b) was used to set the range which needed to be covered. A set of distributions was defined, with means decreasing according to sepsis severity to reflect the trend noted in the meta-analysis of Vincent et al. (Vincent et al. 2003). The distributions implemented in the C-Sepsis CPN are shown in Fig. 5. Plasma hyperalbuminemia is rare (Quinlan et al. 2005), so it is not covered.

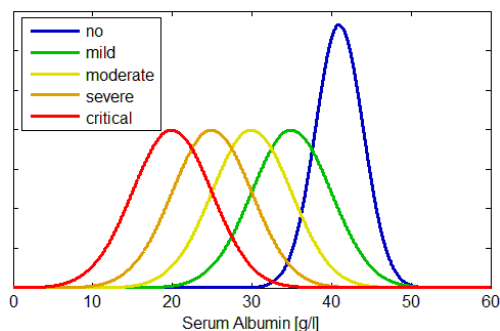


Fig. 5 Gaussian distributions for serum albumin for the severity states of the continuous sepsis CPN.

3.3 Predictive Performance

The final step of the development process, that is, phase II of our framework (Fig. 1), was the validation of the C-Sepsis network. We compare the respective abilities of the D- and C-Sepsis CPNs to predict bacteraemia, assessed by the area under the ROC curve.

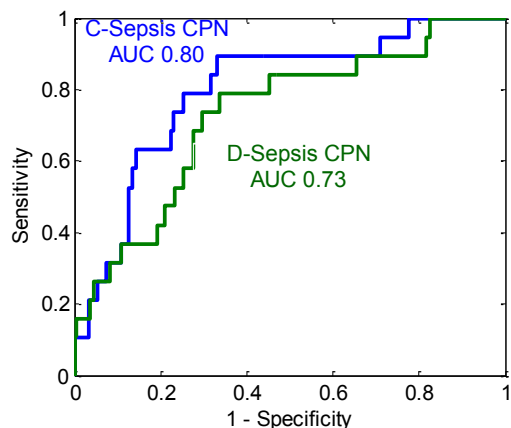


Fig. 6 ROC curves for the prediction of bacteraemia for the D- and C-Sepsis CPNs.

The data used was collected from 263 patients suspected of infection in the acute ward at Hvidovre Hospital, Copenhagen University Hospital, Denmark from November 2011 to May 2012. Of the patients admitted, 19 (7.2%) had bacteraemia. A comparison of ROC curves for the discrete and continuous models shows a non-significant improvement in bacteraemia prediction (two tailed, $p=0.3$). The area under the curve (AUC) for the continuous model is 0.80 (95% confidence interval 0.70-0.90) while the AUC for the discrete model is 0.73 (95% CI 0.62-0.85). The two ROC curves are shown overlaid in Fig. 6. Although the result here is not statistically significant, it is promising and provides motivation for continuing down this path of development.

4. DISCUSSION

In this paper we have described the manual part of the process of building the structure and numbers of a CPN from a mix of expert knowledge and the literature. Many systems have been built this way, but to our knowledge the process has not been described in any great detail – possibly because it is an unsatisfactory mix of partial knowledge based on literature or expert opinion and guesswork. However, despite any flaws in the process, the fact remains that it produces CPNs, such as the Treat CPN which not only work, but actually outperform clinicians (Paul et al. 2006b).

Learning, from a Bayesian perspective, can be considered as a stepwise updating of knowledge in the light of evidence. No formal restrictions are placed on the source or the format of evidence, leaving us free to combine expert opinion with knowledge gleaned from the literature. Adopting such a philosophy allows us to use all the information available, and transforms CPNs into an updateable repository for knowledge. Seen in that light it is less surprising that a manually constructed CPN can outperform clinicians. The CPN can be constructed after close scrutiny of the available literature. The CPN will then serve as a repository of the data

extracted from the literature and will continue to remember this. Any clinician will find it difficult to remember the exact statistics for findings, prevalences etc., which together make up the information going into the CPN. In addition the CPN can provide reasoning, strictly adhering to the axioms of probability theory.

Despite attempts to formalise the manual learning process through systematic searching of the literature and meta-analysis of the studies identified, the process remains difficult. For example, even the distributions for CRP, which was the best parameter in terms of literature availability, were flawed because of problems with the grouping of patients according to severity of sepsis. This is partially due to different strategies for grouping in different papers and partially due to a change in the consensus definitions of sepsis over the time that CRP has been investigated as a biomarker. Another flaw is incomplete data for some types of infection, such as viral infections or atypical pneumonias.

The process was also flawed for the other parameter presented in this paper, Albumin, although in this case it is due to a complete lack of relevant data, requiring that we instead base our guess for distributions on trends in mortality.

The conversion from discrete to continuous nodes for the observed parameters in the sepsis CPN has eliminated the “jumps” in diagnoses, previously seen with small changes of an observable parameter, e.g. fever, with the additional benefit of a modest, albeit not statistically significant, improvement in performance.

Another goal of the construction of the C-Sepsis CPN was to increase the transparency of the network, that is to say, to increase the ease with which the network can be interpreted. An example of this is that the non-infectious causes of SIRS are now explicitly represented by their own node.

We do not believe that substantial further improvement can be made by a continuation of the manual refinement of the probability distributions in themselves, although some benefit may be seen from the addition of new nodes, for example lactate or procalcitonin. To this end, further improvements must then come from a formalisation of the learning process (Fig. 1, Phase III), automatic learning from a database using Dirichlet or EM learning.

5. CONCLUSIONS

Initial results suggest that the manual construction of continuous nodes in the Treat sepsis network can improve the prediction of bacteraemia, although the results from the small test database are not statistically significant. The results obtained provide motivation for further testing using a larger database, and further development of the model through the addition of nodes and application of automatic learning methods to adjust the prior probability distributions.

Acknowledgements: The authors wish to thank Dr. Jenny Dahl Knudsen (Chief Physician, Department of Clinical Microbiology) for providing access to the data used to evaluate the performance of the sepsis CPN.

REFERENCES

- Alvarez, SM, Poelstra, BA & Burd, RS (2006), Evaluation of a Bayesian decision network for diagnosing pyloric stenosis, *J. Pediatr. Surg.*, **41** (1) 155-161.
- Andreassen, S, Leibovici, L, Paul, M, Nielsen, AD, Zalounina, A, Kristensen, LE, Falborg, K, Kristensen, B, Frank, U & Schönheyder, HC (2005), A probabilistic network for fusion of data and knowledge in clinical microbiology, in *Probabilistic Modeling in Bioinformatics and Medical Informatics*, 451-472, Springer, London.
- Andreassen, S, Riekehr, C, Kristensen, B, Schönheyder, HC & Leibovici, L (1999), Using probabilistic and decision-theoretic methods in treatment and prognosis modeling, *Artif Intell Med*, **15** (2) 121-134.
- Andreassen, S, Jensen, FV & Olesen, KG (1991), Medical expert systems based on causal probabilistic networks, *Int. J. Biomed. Comput.*, **28** (1) 1-30.
- Andreassen, S, Rosenfalck, A, Falck, B, Olesen, KG & Andersen, SK (1996), Evaluation of the diagnostic performance of the expert EMG assistant MUNIN, *Electroencephalogr Clin Neurophysiol*, **101** (2) 129-144.
- Bone, RC, Balk, RA, Cerra, FB, Dellinger, RP, Fein, AM, Knaus, WA, Schein, R & Sibbald, WJ (1992), Definitions for sepsis and organ failure and guidelines for the use of innovative therapies in sepsis. The ACCP/SCCM Consensus Conference Committee. American College of Chest Physicians/Society of Critical Care Medicine, *Chest*, **101** (6) 1644-1655.
- Corti, M, Guralnik, JM, Salive, ME & Sorkin, JD (1994), Serum albumin level and physical disability as predictors of mortality in older persons, *JAMA*, **272** (13) 1036-1042.
- Fleck, A, Raines, G, Hawker, F, Trotter, J, Wallace, PI, Ledingham, IM & Calman, KC (1985), Increased vascular permeability: A major cause of hypoalbuminaemia in disease and injury, *Lancet*, **1** (8432) 781-784.
- Hejlesen, OK, Andreassen, S, Hovorka, R & Cavan, DA (1997), "DIAS" the diabetes advisory system: an outline of the system and the evaluation results obtained so far, *Comput Methods Programs Biomed*, **54** (1) 49-58.
- Higgins, J, White, IR & Anzures-Cabrera, J (2008), Meta-analysis of skewed data: Combining results reported on log-transformed or raw scales, *Stat Med*, **27** (29) 6072-6092.
- Hozo, SP, Djulbegovic, B & Hozo, I (2005), Estimating the mean and variance from the median, range, and the size of a sample, *BMC Med Res Methodol*, **5** (1) 13.
- Kariv, G, Shani, V, Goldberg, E, Leibovici, L & Paul, M (2011), A model for diagnosis of pulmonary infections in solid-organ transplant recipients", *Comput Methods Programs Biomed*, **104** (2) 135-142.
- Lauritzen, SL (1995), The EM algorithm for graphical association models with missing data, *Comput Stat Data Anal*, **19** (2) 191-201.
- Leibovici, L, Fishman, M, Schönheyder, HC, Riekehr, C, Kristensen, B, Shraga, I & Andreassen, S (2000), A causal probabilistic network for optimal treatment of bacterial infections, *IEEE Trans Knowl Data Eng*, **12** (4) 517-528.
- Leibovici, L, Paul, M, Nielsen, AD, Tacconelli, E & Andreassen, S (2007), The TREAT project: decision support and prediction using causal probabilistic networks, *Int J Antimicrob Agents*, **30** 93-102.
- Leibovici, L, Paul, M & Andreassen, S (2010), Balancing the benefits and costs of antibiotic drugs: the TREAT model, *Clin Microbiol Infect*, **16** (12) 1736-1739.
- Levy, MM, Fink, MP, Marshall, JC, Abraham, E, Angus, D, Cook, D, Cohen, J, Opal, SM, Vincent, J & Ramsay, G (2003), 2001 sccm/esicm/accp/ats/sis international sepsis definitions conference, *Intensive Care Med*, **29** (4) 530-538.
- Mitaka, C (2005), Clinical laboratory differentiation of infectious versus non-infectious systemic inflammatory response syndrome, *Clin Chim Acta*, **351** (1) 17-29.
- Paul, M, Andreassen, S, Nielsen, AD, Tacconelli, E, Almasreh, N, Fraser, A, Yahav, D, Ram, R, Leibovici, L & TREAT Study Group (2006a), Prediction of Bacteremia Using TREAT, a Computerized Decision-Support System, *Clin Infect Dis*, **42** (9) 1274-1282.
- Paul, M, Andreassen, S, Tacconelli, E, Nielsen, AD, Almasreh, N, Frank, U, Cauda, R, Leibovici, L on behalf of the TREAT Study Group (2006b), Improving empirical antibiotic treatment using TREAT, a computerized decision support system: cluster randomized trial, *J Antimicrob Chemother*, **58** (6) 1238-1245.
- Paul, M, Nielsen, AD, Goldberg, E, Andreassen, S, Tacconelli, E, Almasreh, N, Frank, U, Cauda, R & Leibovici, L (2007), Prediction of specific pathogens in patients with sepsis: evaluation of TREAT, a computerized decision support system, *J Antimicrob Chemother*, **59** (6) 1204-1207.
- Pearl, J (1988), *Probabilistic reasoning in intelligent systems: networks of plausible inference*, Morgan Kaufmann.
- Pierrakos, C & Vincent, J (2010), Sepsis biomarkers: a review, *Crit Care*, **14** (1) R15.
- Quinlan, GJ, Martin, GS & Evans, TW (2005), Albumin: biochemical properties and therapeutic potential, *Hepatology*, **41** (6) 1211-1219.
- Rudmin, JW (2010), Calculating the Exact Pooled Variance, *arXiv preprint arXiv:1007.1012*.
- Sadeghi, S, Barzi, A, Sadeghi, N & King, B (2006), A Bayesian model for triage decision support, *Int J Med Inform*, **75** (5) 403-411.
- Schurink, CM, Visscher, S, Lucas, PF, Leeuwen, H, Buskens, E, Hoff, R, Hoepelman, AM & Bonten, MM (2007), A Bayesian decision-support system for diagnosing ventilator-associated pneumonia, *Intensive Care Med*, **33** (8) 1379-1386.
- Spiegelhalter, DJ & Lauritzen, SL (1990), Sequential updating of conditional probabilities on directed graphical structures, *Networks*, **20** (5) 579-605.
- Vigushin, DM, Pepys, MB & Hawkins, PN (1993), Metabolic and scintigraphic studies of radioiodinated human C-reactive protein in health and disease, *J Clin Invest*, **91** (4) 1351.
- Vincent, J, Dubois, M, Navickis, RJ & Wilkes, MM (2003), Hypoalbuminemia in acute illness: Is there a rationale for intervention?: A meta-analysis of cohort studies and controlled trials, *Ann Surg*, **237** (3) 319.