GRANT APPLICATION FORM NWO

1a) Project Title

Learning and Using *Understandable* Prognostic Models in Medicine

1b) Project Acronym

ProBayes

1c) Principal Investigator

Applicant	Dr Peter Lucas
Department	Institute for Computing and Information Sciences
Institute	University of Nijmegen
Address	Toernooiveld 1, 6525 ED Nijmegen, The Netherlands
	P.O. Box 9010, 6500 GL Nijmegen, The Netherlands
Telephone	+ 31 24 3652084
Telefax	+ 31 24 3653137
Email	lucas@cs.uu.nl

2) Summary

Although research on learning *Bayesian networks* from data started only about 10 years ago, significant theoretical progress has been made since that time. However, the community involved in this research has placed relatively little emphasis on gaining insight into the usefulness of this technology in solving real-life problems. In fact, the issue of tailoring Bayesian-network learning methods to the characteristics of problem domains has not even been addressed.

The ProBayes project's primary aim is to investigate whether the Bayesian-network formalism offers a suitable framework for *learning prognostic models from clinical datasets.* Prognostic models play an important role in oncology, and the quality of the clinical management of cancer in patients may profit considerably from deploying medical decision-support systems incorporating such models. Decision support in clinical oncology is therefore taken as an experimental setting for ProBayes.

Since it is expected to be essential to exploit *background knowledge* to guide data-mining and learning in the context of real-life problems, the major goal of the requested research is to obtain insight into the form of the required knowledge in constructing prognostic Bayesian networks from data. Within the empirical setting of the project, the problem of learning Bayesian networks will be studied along the spectrum from rare to common disorders, not simply by using publicly available datasets of unclear quality, but in its real-life context with considerable input from expert clinicians. There will be a major emphasis in the project on learning intuitive, understandable Bayesian networks, such as those that can be given a causal reading.

Finally, the clinical usefulness of the developed Bayesian-network models will be studied in the context of clinical management of cancer patients. Overall, the ProBayes project aims to extend Bayesian-network technology so that it comes nearer to fulfilling its promise as a practical technology for clinical decision support.

3) Classification

Datamining and Datawarehousing (2.2); Reasoning Systems (2.7); Heuristic Algorithms (5.5). NOAG-i: Data and Knowledge Systems.

4) Composition of Research Team

Institute for Computing and Information Sciences, University of Nijmegen (KUN), Toernooiveld 1, 6525 ED Nijmegen, The Netherlands		
Name	**Discipline**	**Paid by**
Dr P.J.F. Lucas	Computing Science (AI, Medical Informatics), Medicine	KUN
Prof.dr ir Th.P. van der Weide	Computing Science (IR, Information Systems)	KUN
OIO – requested	Computing Science (AI)	NWO
Institute for Information and Computing Sciences, Utrecht University (UU) Padualaan 14, 3584 CH Utrecht, The Netherlands		
Name	**Discipline**	**Paid by**
Prof.dr ir L.C. van der Gaag	Computing Science (Bayesian Networks)	UU
Dr S. Renooij	Computing Science (Bayesian Networks)	UU
Netherlands Cancer Institute/Antoni van Leeuwenhoek Hospital (NKI) Plesmanlaan 121, 1066 CX Amsterdam, The Netherlands		
Name	**Discipline**	**Paid by**
Dr B. Taal	Medicine/Gastroenterological Oncology	NKI
Dr H. Boot	Medicine/Gastroenterological Oncology	NKI
H. Zuetenhorst, MD	Internal Medicine/Oncology	NKI
OIO – requested	Health Sciences	NWO
Harvard Medical School, Harvard University, USA		
Name	**Discipline**	**Paid by**
Dr M. Ramoni (consultant)	Biomedical Informatics (Learning Bayesian Networks from Data)	Harvard
Academic Medical Centre (AMC), Meibergdreef 15, 1105 AZ Amsterdam		
Name	**Discipline**	**Paid by**
Prof.dr J.C. Wyatt (consultant)	Medical Informatics (Evaluation of Decision Aids)	AMC

5) Research School

Research School for Information and Knowledge Systems (SIKS); Oncology Graduate School Amsterdam (OOA).

6) Description of Proposed Research

Background: Prognostic Models in Medicine

Prognosis (pro: before; gnoscere: to know) literally means 'to know beforehand' or, as a noun, foreknowledge. *Medical* prognosis is defined here as [1]: *the prediction of the future* course *and* outcome *of disease processes, which may either concern their natural course or their outcome after treatment.* Prognosis is interwoven with the other aspects of clinical patient management, namely those of diagnosis, treatment and follow-up. The clinician views prognosis as a *decision tool* incorporating various amounts of diagnostic and treatment information [45]. Because of the increasing awareness within clinical practice of evidence-based medicine, the increasing availability of patient data, and the central role that prognosis plays in the management of patients, prognostic models are becoming more and more important. Examples of techniques used in building prognostic models are simple decision rules based on categorisations, and prognostic scores computed by using Bayes' rule, logistic regression, or Cox's regression if time is involved; their parameters are sometimes assessed from data especially collected for that purpose, and in many cases purely based on expert judgement.

Prognostic models are of great importance in oncology[1], as being able to predict treatment outcome in patients with cancer is the single most important issue in treatment selection. Having

---

[1]Oncology is the medical science of benign and malignant tumours.

available sound, practically useful prognostic models is a matter of life and death in this field. This thus also holds for the two clinical areas the proposed research will focus on: *colorectal cancer* and *gastrointestinal carcinoid tumours*. Colorectal cancer is among the major causes of death and disability within the western population. Carcinoid tumours are rare disorders, requiring specialised clinical knowledge in order to be treated effectively. For both disorders, various prognostic scoring schemes are in use to help clinicians in managing patients with these disorders. In colorectal cancer, Dukes' classification[2] is often used; the prognostic significance of this score has been extensively studied [15]. The Capella classification method [8] has been proposed as a prognostic method for carcinoid tumours; it has recently been revised for gastrointestinal carcinoid tumours [9]. These scores can be viewed as providing summaries of intermediate to long-term clinical outcome.

A limitation of such classification rules is that they summarise complicated medical conditions in terms of a single class or score, which does not do justice to the whole range of factors and mechanisms involved. In fact, such schemes give little or no insight into the way the various factors and mechanisms interact to influence the course or outcome of a disease in the individual patient or in groups of patients. They are *black-box* models, which, as such, can either be followed slavishly, or not followed at all. Neither a deep understanding of what is happening and likely to happen in the individual patient, nor any insight into the disease and how its natural course may be modified due to particular treatment modalities can be obtained from such rules. It is thus not surprising that many prognostic models have achieved at most a marginal role in the practice of medicine [47]; in oncology, however, the need for prognostic models is so large that having any 'model' available is considered to be better than having none at all.

Bayesian Networks as Clinical Decision-making Tools

Bayesian-network approaches to the development of medical decision-support systems offer substantial advantages in terms of flexibility and potential of knowledge reuse, mainly due to the declarative, i.e. task-non-specific, nature of the knowledge embodied in such systems [34]. Bayesian networks are examples of so-called graphical statistical models, and offer an encoding of a joint probability distribution on the variables in a problem domain in terms of local (conditional) probability distributions and a graph which represents the statistical dependences and independences among the variables concerned [13, 27, 39]. The Bayesian-network formalism can easily be augmented with decision theory, and is so attractive because it allows for explicitly representing the uncertainties and preferences that go with much of the knowledge used in medical decision making. A single Bayesian network can be used to deal with a large number of different what-if questions clinicians are likely to find important, such as optimal treatment selection, prognostic assessment for individual patients and generating profiles for specific treatment outcomes for one or more disorders [35]. Experience, both by the research team and others, suggests that clinicians find this computational formalism intuitive and appealing [24, 25, 35, 36].

A Bayesian network for a real-life clinical problem is usually constructed with the help of one or more expert clinicians [19]. However, building Bayesian networks using expert clinician knowledge, although by now known to be feasible, can be very time consuming. In the context of the current emphasis on evidence-based medicine, it seems inevitable to base the interactions represented in a Bayesian network as much as possible on observational clinical data. Within the Department of Medical Oncology of NKI high quality clinical data are being collected in the context of clinical trials and ongoing clinical research. In addition, the Comprehensive Cancer Centre Amsterdam (IKA), with which NKI has very close links, collects cancer registry data. NKI provides therefore a good real-life environment for research into clinical decision support and machine learning.

The combination of flexibility, understandability, and learnability renders Bayesian networks an ideal basis for an extensive study of the requirements for practically useful computer-based prognostic tools for clinicians. The selected subjects of colorectal cancer and carcinoid tumours ensure:

---

[2]This places a patient into one of the categories A, B, C and D, based on the presence or absence of invasion of the tumour into surrounding tissue and metastasis; A is the most favourable class with an associated 5-year survival rate of approximately 80%.

(1) practical relevance, as predicting outcome in patients with these very disorders is considered to be hard by clinicians; (2) that actual clinical data of patients are available so that learning Bayesian networks from data can be sensibly explored; (3) input from clinical experts, as there is a great need for sound prognostic models in oncology. The multidisciplinary research team has extensive experience in the research topic, as well as with collaborative research on Bayesian-network development for related clinical problems [35, 20]. No structure learning methods, however, were used in the context of our previous research.

Dr Lucas and Prof. van der Gaag have wide experience in building Bayesian-network based and other decision-support systems for clinical problems in collaboration with clinicians [29, 33, 35, 36, 19]. Dr Lucas has 18 years experience in dealing with the methodological and technical questions which arise in research in medical decision support; having been trained both as a computer scientist and an MD, he has the necessary background to lead this project. Prof. van der Gaag is an internationally recognised expert in the foundations of Bayesian networks, and was the first researcher to start with Bayesian-network research in the Netherlands in the 1980s. Dr Taal and Dr Boot are experienced research-oriented clinical oncologists, with an extensive research record covering more than 20 years of clinical research. There are excellent international contacts with other researchers in the area of Bayesian networks and medical decision support; in particular Dr Ramoni and Prof. Wyatt are prepared to offer advice during the course of the project.

Learning Bayesian Networks from Data

Learning a Bayesian network from data can be separated into two tasks: (1) *structure learning*, i.e. identifying the topology of the network, and (2) *parameter learning*, i.e. determining the associated joint probability distribution for a given network topology [41]. As the number of possible Bayesian-network structures for a given set of statistical variables is exponentially large, it is necessary to use heuristic methods to construct a Bayesian network automatically from data. A frequently used procedure for Bayesian network structure construction from data is the K2 algorithm, developed by Cooper and Herskovits [12]. Given a database D, this algorithm searches for the Bayesian network structure G which maximises the probability $\Pr(G \mid D)$, using simplifying assumptions. K2 is a greedy heuristic. It starts by assuming that a node lacks parents, after which in every step it adds incrementally that parent whose addition increases the probability of the resulting structure most. K2 stops adding parents to the nodes when the addition of a single parent cannot increase the probability. The K2 algorithm is a typical *search & scoring method*, i.e. heuristic search is guided by a scoring function, often called the Bayesian scoring function [12]. Other scoring methods include the MDL (Minimal Description Length) measure [32], and various variants of the Bayesian scoring function, such as the BDe measure [26]. The Bayesian scoring function and the MDL measure have been shown to be asymptotically equivalent [5]. The K2 algorithm needs a total order among nodes to start, which can be regarded as a form of prior domain knowledge, although not necessarily of the right type.

A learning method which takes an entirely different approach is *dependency analysis* [2, 6, 7, 10]. This method uses independence tests as part of a learning procedure. The collection of conditional independence statements represented in a Bayesian network imply other conditional independence statements using the independence axioms [39]; these can be utilised in learning. The independence axioms also imply that some Bayesian networks with different topologies may in fact be statistically equivalent. It has been shown that this feature can be favourably exploited by exploring the search space inhabited by equivalent class representatives rather than by the individual network models themselves [11, 28]. The scoring and dependency-analysis approaches have also been combined in the hope of obtaining a method that is more powerful than its constituents. For example, the CB algorithm (named after the two phases of the algorithm) does not require a total order among nodes as a start; it uses conditional independence tests for that purpose, and uses K2 in its second phase [43]. The BENEDICT learning method integrates the two approaches even further [2].

There is also experimental evidence that learning results can be improved by exploiting hidden structure within the conditional probability distributions, by taking advantage of the fact that they may incorporate context-specific independences [18]. This means that probabilities can be

the same for different values of one or more variables, given a fixed value of one or more other variables, formally: $\Pr(X \mid Y, c, Z) = \Pr(X \mid Y, c)$ for any value of X, Y and Z, and c a fixed value of context variable C. This implies that the probability distribution can be represented more compactly, and the learning procedure may then yield more accurate models.

In learning Bayesian networks from data, care must be exercised that the resulting models are not overly complex, as it has been shown that if Bayesian networks are used for classification purposes, those with simple topologies, such as naive Bayesian networks and tree-augmented Bayesian networks (TANs), may perform as well or even better than more sophisticated ones [14, 17, 37]. Although it is not our goal to merely construct models for patient classification purposes, finding the right balance between simple and complex networks is definitely difficult.

Even though the research results summarised above appears to be promising, very few researchers have actually tried to learn Bayesian networks for real-life problems. In many of the papers, experimental results are based on using datasets generated by Monte-Carlo simulation from a single hand-crafted Bayesian network, called ALARM. This network was, in fact, used in the first paper on learning Bayesian networks by Cooper and Herskovits [12]. This makes comparing results as reported in various papers straightforward, but results achieved with real-life datasets may still be substantially different from the results achieved with such generated datasets. In the ProBayes project the problem of learning Bayesian networks will be studied not simply by using generated datasets or publicly available datasets of unclear quality, but in its real-life context with considerable input from expert clinicians.

Before one uses a suggested prognostic model it is important to have an indication that the model would work well on a different population than the one which was used to develop the model. There are many studies reporting on model validation, certainly the lack thereof (e.g. [38, 47]). One may distinguish between laboratory evaluation and clinical evaluation [3, 48]. A *laboratory evaluation* usually focuses on the performance of the model. Relevant questions in laboratory evaluation are whether the model passes the appropriate statistical tests, usually on a new data set, and whether it is the best model given the available factors. In a *clinical evaluation* one is interested in the question whether the model is satisfactory for its *clinical purpose*. It is possible to have a statistically but yet not clinically valid model and vice versa.

Aims and Objectives

The project's primary aim is to investigate whether the Bayesian network formalism offers a suitable framework for learning prognostic models from clinical datasets, and for building computer-based medical decision-support systems (DSSs). Since it is usually necessary to use medical background knowledge in data-mining and machine learning applications to guide the learning process, an additional aim is to obtain insight into the appropriate form and mixture of data and medical expertise in constructing understandable prognostic Bayesian networks from data. Finally, we intend to investigate whether computer-based prognostic tools based on Bayesian networks are seen by clinicians as useful in patient management.

In order to see whether our methods are general enough to be used in related clinical areas with different characteristics, we have chosen two domains for experimentation: colorectal cancer and gastrointestinal carcinoid tumours. In the colorectal cancer area, the focus lies on learning prognostic Bayesian network models using much data with only a limited number of variables describing each patient; in the gastrointestinal carcinoid area the emphasis will be on coping with the situation that data of a hundred to a few hundred patients is available, but with many variables describing each patient. This corresponds to the distinction between registry datasets and clinical datasets that is quite typical for medicine as a whole.

Our specific objectives are:
- the development of new structure (and if time allows also parameter) background-knowledge enhanced learning methods and a workbench for learning Bayesian networks from data;
- the analysis of datasets with data of colorectal cancer and carcinoid tumours using the workbench;

- the development of causal models in collaboration with medical specialists in colorectal cancer and carcinoid tumours;
- the investigation of the role and place of medical background knowledge in guiding the construction of optimal network structures;
- obtaining insight into the clinical value of Bayesian-network based prognostic models in the context of decision support, for which actual decision-support systems will be developed.

Research Methods

There are various aspects that will be addressed in the proposed research.

A. *Role of background knowledge in learning prognostic Bayesian networks*

The existing algorithms for learning Bayesian networks from data offer a good starting point for the research; a number of software tools implementing these algorithms are available in the research community [10, 40]. Some sophisticated algorithms, such as simulated annealing and tabu search have also been studied in the past, but only using artificial data [5]. In particular tabu search, a tunable, partially informed search method, is considered to be suitable as a start for ProBayes. A limitation of current algorithms is that it is often unclear whether a particular topology makes medical sense. Of course, once a Bayesian network has been learnt, it can be evaluated, for example, by determining its performance rates, but it is still possible that the best network has not been discovered, or that there are networks which are probabilistically equivalent, but much easier to understand when given a causal reading. One of the hypotheses investigated in the project is whether starting Bayesian-network structure learning with fragments of *causal background knowledge*, based on strong clinical evidence, will improve both the quality and understandability of the networks. This research will combine ideas about informed search with the utilisation of background knowledge. The research will also undertake identifying other types of background knowledge that can be used in the context of structure learning, and if time allows, also parameter learning.

In close collaboration with the clinical oncologists causal models describing the relevant factors of the course and outcome of both disorders will be developed. This will not be easy, as we intend to model the information considered clinically important, including results from imaging and laboratory investigations. Experience in building Bayesian networks built up in the last decade has convincingly shown that causal knowledge can offer very useful guidance in developing Bayesian networks. A workbench will be developed which incorporates various search algorithms, including those of the more informed type, such as tabu search, and domain-specific methods for search guidance. This workbench will be employed in the process described above.

B. *Evaluation of prognostic Bayesian network models*

Evaluation of the prognostic models developed will be done by comparing them both quantitatively and qualitatively with existing prognostic models. A quantitative comparison is possible for simple classification rules, like Dukes' classification of colorectal cancer, as such rules are used in survival analysis. So, the true-positive and true-negative rates (TPR/TNR) of such rules when used predictively can be compared to those for a Bayesian network. Optimal classification rates for a Bayesian network can be determined using Receiver Operating Characteristics (ROC) analysis using a training set [44]. Evaluation can then be carried out using a test set, and a number of measures are available, in addition to the TPR and TNR, to get insight into the quality of the a posteriori distributions computed, such as the Brier score [22]. Decision-analytic assessment of the prognostic models will also be carried out [44]. A quantitative comparison is also possible between the expert-based and learnt Bayesian networks for colorectal cancer and carcinoid tumours. The extent to which the learnt networks are understandable and permit a causal reading involves qualitative analysis.

In additional to the standard evaluation measures discussed above, it is necessary to develop measures especially suitable for Bayesian networks in order to evaluate different aspects of the network models. The measure which comes closest to the intentions of this part of the research is

the Kullback-Leibler cross-entropy measure, which allows comparing two probability distributions [31, 46]. However, this measure also ignores important features of a Bayesian network. Hence, we will also undertake the development of evaluation measures which are able to reflect specific domain characteristics of network models.

C. *Clinical significance of prognostic Bayesian network models*

As previously developed prognostic models, and certainly computer-based models, have not been particularly successful in medicine [38, 47], establishing the clinical value and usefulness of the Bayesian-network approach is crucial. We will carry out a study of the clinical value of the colorectal cancer and gastrointestinal carcinoid tumour models. A collection of data of patients with stroke and colorectal cancer will be selected randomly. Medical doctors with varying amounts of expertise in the fields of colorectal cancer and carcinoid tumours are requested to draw up the disease management for these patients, and predict likely outcome, using a decision-support system incorporating the models for half of the patients and unaided for the other half. To save time, the basis for the decision-support systems will be the commercially available Hugin Bayesian-network shell, which will be augmented with an appropriate user-interface. The results will be analysed statistically. In both fields, the systems will be evaluated in terms of amount of assistance offered, user-friendliness, and whether clinicians would be prepared to use the systems as part of their every-day patient management. We have good contacts with Prof. Jeremy Wyatt (AMC, Amsterdam), an internationally recognised expert in the area of evaluation of DSSs, who will be consulted with respect to evaluation issue [16, 48].

Embedding of Research

The proposed research is in line with research being carried out by Dr Lucas and Prof. van der Gaag in the areas of Bayesian networks, (medical) decision support and model-based reasoning. Prof. van der Gaag holds an NWO Pioneer Grant in the area of Bayesian networks, and the computing-science oriented junior researcher will definitely profit from the collaboration between Nijmegen and Utrecht that will be established in this project. In addition, contact will be sought with the Department of Medical Physics and Biophysics at the University of Nijmegen, which also has expertise in Bayesian-network research.

The junior researcher at NKI will work in a environment known for its close links between research and clinical practice. NKI maintains a very strong research culture. Previous research by Dr Taal and Dr Boot has focussed on finding useful outcome predictors in cancer, and both have experience with Bayesian-network technology, obtained from their involvement in other Bayesian-network projects with Dr Lucas and Prof. van der Gaag. Hence, the junior researcher will be working in an environment with a background in the sort of research requested.

7) Work Programme

I. Start-up (months 0–3)
   1) collecting and studying relevant literature
   2) review of available software packages
   3) review of the datasets for colorectal cancer and carcinoid tumours available at NKI and IKA
II. Design of qualitative models (months 3–9)
   1) study of literature on Bayesian networks, machine learning (KUN); similarly for medical
      decision making, colorectal cancer, gastrointestinal carcinoid tumours (NKI)
   2) design of causal models of colorectal cancer and carcinoid tumours with clinical oncologists
III. Learning from background knowledge and evaluation methods (months 9–21)
   1) study of the literature on using background knowledge in machine learning (mostly inductive-
      logic programming literature)
   2) development of methods which exploit similar principles in the context of learning Bayesian
      networks from data
   3) development of methods for the evaluation of Bayesian networks

Table 1: Distribution of work; ++: major task; +: minor task; –: none.

Task	Computer Scientist	Health Scientist
Design of learning algorithms	++	—
Development of Workbench	++	—
Acquisition of background knowledge	—	++
Utilisation of background knowledge in learning	++	+
Comparison with other prognostic methods	—	++
Performance evaluation	++	++
Clinical evaluation	+	++

IV. Workbench development and initial model discovery (month 21–27)
  1) development of methods for learning Bayesian networks from data, employing background knowledge
  2) development of a workbench and initial model discovery
  3) initial DSS development using Hugin

V. Refinement (months 27–36)
  1) improving the original methods for background-knowledge guided learning
  2) refinement of the workbench and DSS based on previous experience

VI. Evaluation (months 36–42)
  1) establishment of parameters for evaluation
  2) cross comparison of learnt Bayesian-network models, expert-based models, and models from the literature
  3) study of understandability of Bayesian network models and clinical usefulness of the DSS

VII. Dissemination of results (months 42–48)
  1) writing of two PhD theses
  2) writing of scientific journal and conference papers
  3) the software that is developed during the project will be made publicly available

Table 1 indicates how the work load will be distributed among the computer scientist and the health scientist requested for this project.

8a) Expected Use of Instrumentation

Standard workstations will be used in the project.

8b) Required Observing Facilities

NA

9a) Selected publications by applicants

  1) Lucas PJF. Analysis of notions of diagnosis. Artificial Intelligence, 1998; 105(1-2): 293–341.
  2) Lucas PJF, Boot H, Taal BG. Computer-based decision-support in the management of primary gastric non-Hodgkin lymphoma. Methods Inform Med, 1998; 37: 206–219.
  3) Lucas PJF, De Bruijn NG, Schurink K, Hoepelman IM. A Probabilistic and decision-theoretic approach to the management of infectious disease at the ICU. Artif Intell Med, 2000; 19(3): 251–279.
  4) Taal BG, Peterse H, Boot, H. Clinical presentation, endoscopic features and treatment of gastric metastases from breast carcinoma. Cancer, 2000; 89: 2214–2221.
  5) Van der Gaag LC, Renooij S, Witteman CLM, Aleman B, Taal BG. How to elicit many probabilities. In: K. Laskey, H. Prade (Eds). Proc UAI99, Morgan Kaufmann, San Francisco, 1999: 647–654.

9.b) International literature

[1] Abu-Hanna A, Lucas PJF. Prognostic models in medicine: AI and statistical approaches. Methods Inform Med, 2001; 40: 1–5.

[2] Acid S, Campos LM de. A hybrid methodology for learning belief networks: BENEDICT. International Journal of Approximate Reasoning, 2001; 27: 235–262.

[3] Altman DG, Royston, P. What do we mean by validating a prognostic model? Statistics in Medicine, 2000; 19: 453–473.

[4] Asch D, Patton J, Hershey J. Knowing for the sake of knowing: the value of prognostic information. Medical Decision Making, 1990; 10: 47–57.

[5] Bouckaert R. Bayesian Belief Networks: from Construction to Inference. PhD Thesis, Utrecht University, 1995.

[6] Campos LM de, Huete JF. On the use of independence relationships for learning simplified belief networks. International Journal of Intelligent Systems, 1997; 12: 495–522.

[7] Campos LM de, Huete JF. A new approach for learning belief networks using independence criteria. International Journal of Approximate Reasoning, 2000; 24: 11–37.

[8] Capella C, Heitz PU, Hofler H, Solcia E, Kloppel G. Revised classification of neuroendocrine tumors of the lung, pancreas and gut. Virchows Archive, 1995; 425: 547–560.

[9] Capella C, La Rosa S, Ucella S, Billo P, Cornaggia M. Mixed endocrine-exocrine tumors of the gastrointestinal tract. Sem Diagn Pathol, 2000, 17: 91–103.

[10] Cheng J, Bell D. Learning Bayesian networks from data: an efficient approach based on information theory. Proc 6th ACM Intern Conf Inform and Knowl Man, 1997.

[11] Chickering DM. Learning Equivalent Classes of Bayesian-Network Structures. Technical Report, Microsoft Research, 2001.

[12] Cooper GF, Herskovits E. A Bayesian method for the induction of probabilistic networks from data. Machine Learning, 1992; 9: 309–347.

[13] Cowell RG, Dawid AP, Lauritzen SL, Spiegelhalter DJ. Probabilistic Networks and Expert Systems. Springer: Berlin, 1999.

[14] Domingos P, Pazzani M. On the optimality of the simple Bayesian classifier under zero-one loss. Machine Learning, 1997; 29: 103–1

[15] Kyrialos M. The president's cancer, the Dukes classification, and confusion. Arch Pathol Lab Med, 1985; 109: 1063.

[16] Friedman C, Wyatt JC. Evaluation Methods in Medical Informatics. New York: Springer, 1997.

[17] Friedman N, Geiger D, Goldszmidt M. Bayesian network classifiers. Machine Learning, 1997; 29: 131–163.

[18] Friedman N, Goldszmidt M. Learning Bayesian networks with local structure. In: Jordan MI (Ed.), Learning in graphical models. MIT Press: Cambridge, MA, 1999.

[19] Gaag LC van der, Renooij S, Witteman CLM, Aleman B, Taal BG. How to elicit many probabilities. In: K. Laskey, H. Prade (Eds). Proc UAI99, Morgan Kaufmann, San Francisco, 1999: 647–654.

[20] Gaag, LC van der, Taal B, Alleman, Renooij S. A Bayesian network for the management Oesopheal cancer. Artificial Intelligence in Medicine, 2002.

[21] Glymour C, Cooper GF. Computation, Causation, and Discovery. MIT Press: Cambridge, 1999.

[22] Hand DJ. Construction and assessment of classification rules. John Wiley, 1997.

[23] T. Hastie, R. Tibshirani, J. Friedman. The Elements of Statistical Learning: Data Mining, Inference, and Prediction. Springer, New York, 2001.

[24] Heckerman DE, Horvitz EJ, Nathwani BN. Towards normative expert systems: part I – The Pathfinder project. Methods Inf Med, 1992; 31: 90–105.

[25] Heckerman DE, Nathwani BN. Towards normative expert systems: part II – probability-based representations for efficient knowledge acquisition and inference. Methods Inf Med, 1992; 31: 106–116.

[26] D. Heckerman. A Tutorial on Learning with Bayesian Networks. Technical Report MSR-TR-95-06. Microsoft Corporation, Redmond, 1996.

[27] Jensen FV. Bayesian Networks and Decision Graphs. Springer: New York, 2001.

[28] Kočka T, Castelo R. Improved learning of Bayesian networks. Proc UAI2001, Morgan Kaufmann, San Francisco, 2001.

[29] Korver M, Lucas PJF. Converting a rule-based expert system into a belief network. Medical Informatics, 1993; 18(3): 219–241.

[30] Krause P. Learning probabilistic networks. Knowldge Engineering Review, 1998; 13(4): 321–351.

[31] Kullback S, Leibler R. On information and sufficiency. Annals of Mathematical Statistics, 1951; 22: 79–86.

[32] Lam W, Bacchus F. Learning Bayesian belief networks: an approach based on the MDL principle. Comp Intell, 1994; 10: 269–293.

[33] Lucas PJF. Knowledge acquisition for decision-theoretic expert systems. AISB Quarterly, 1996, 94: 23–33.

[34] Lucas PJF, Boot H, Taal BG. A decision-theoretic network approach to treatment management and prognosis. *Knowledge-based Systems*, 1998; 11: 321–330.

[35] Lucas PJF, Boot H, Taal BG. Computer-based decision-support in the management of primary gastric non-Hodgkin lymphoma. Method Inform Med, 1998; 37: 206–219.

[36] Lucas PJF, de Bruijn NG, Schurink K, Hoepelman IM. A Probabilistic and decision-theoretic approach to the management of infectious disease at the ICU, Artif Intell Med, 2000; 19(3): 251–279.

[37] Lucas PJF. Expert knowledge and its role in learning Bayesian networks in medicine: an appraisal, in: S. Quaglini, P. Barahona, S. Andreassen (eds.), *Proceedings of Artificial Intelligence in Medicine in Europe 2001*, Lecture Notes in Artificial Intelligence 2101, Springer-Verlag, Berlin, 2001, pp. 156–166.

[38] Neuberger J. Prognostic models – what is their future? Eur J Gastol & Hepatol, 1997; 9; 1145–1147.

[39] Pearl J. Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference. San Francisco: Morgan Kaufmann, 1998.

[40] Ramoni M, Sebastiani P. Discovering Bayesian Networks in Incomplete Databases. Report KMI-TR-46, Knowledge Media Institute (KMI): Open University, 1997.

[41] Ramoni M, Sebastiani P. Bayesian methods. In: Berthold M, Hand DJ (Eds.). Intelligent Data Analysis: an Introduction. Springer: Berlin, 1999, pp. 129–166.

[42] Salzberg SL. On comparing classifiers: a critique of current research and methods. Data Mining and Knowledge Discovery, 1999; 1: 1–12.

[43] Singh M, Valtorta M. Construction of Bayesian Network Structures from Data: a Brief Survey and an Efficient Algorithm. In: Proc UAI93. San Francisco: Morgan Kaufmann, 1993; 259–265.

[44] Sox HC, Blatt MA, Higgins MC, Marton KI. Medical Decision Making. Boston: Butterworths, 1988.

[45] Windeler J. Prognosis - what does the clinician associate with this notion? Statistics in Medicine 2000; 19: 425–430.

[46] Wu X, Lucas PJF, Kerr S, Dijkhuizen R. Learning Bayesian-network topologies in realistic medical domains, in: J. Crespo, V. Maojo, F. Martin (eds.). Medical Data Analysis (2nd International Symposium, ISMDA 2001, Madrid, Spain, Oct. 2001), LNCS2199, Springer, Berlin, pp. 302–307.

[47] Wyatt JC, Altman DGF. Commentary – prognostic models: clinically useful or quickly forgotten. BMJ, 1995; 311: 1539–1541.

[48] Wyatt JC, Spiegelhalter DJ. Evaluating medical expert systems: what to test for and how? Medical Informatics, 1990; 15(3): 205–217.

10) Requested Budget

Two junior researchers (OIOs) are requested: one will be appointed in Nijmegen and one in Amsterdam (NKI).

OIOs		Amount	
Salary	$2 \times 129{,}897$ Euro	259794 Euro	
Benchfee	$2 \times 4{,}538$ Euro	9176 Euro	
Software (Hugin)	$3 \times 3{,}000$ Euro	9000 Euro	
		————————	+
		277970 Euro	

We intend to buy 3 site licenses of the Hugin Bayesian-network package (`http://www.hugin.com`) in order to speed-up the research project, and to circumvent putting unnecessary software development load upon the junior computer scientist. This software will be used for expert-guided model development and DSS development, but not for the machine learning research, for which a separate workbench will be developed.