A HYBRID MODEL-BASED VENTILATORY DECISION SUPPORT SYSTEM

D. A. Linkens[1], H. F. Kwok[1,2], M. Mahfouf[1], G. H. Mills[2].

[1]Automatic Control and Systems Engineering, University of Sheffield, Sheffield S1 3JD, UK.
Address for correspondence and email: d.linkens@sheffield.ac.uk
[2]Academic Department of Anaesthesia, University of Sheffield and Intensive Care Unit, Royal Hallamshire Hospital, Sheffield S10 2JF, UK.

Abstract: A hybrid knowledge-and-model-based advisory system for intensive care ventilators has been developed. The system consists of two parts: a knowledge-based top-level module using neural fuzzy technology and a model-based lower-level module consisting of 4 sub-units. The system generates advice on four ventilator settings (the inspired fraction of oxygen (FiO_2), positive end-expiratory pressure (PEEP), peak inspiratory pressure (PINSP) and ventilatory rate) based on the patient's routine and cardio-respiratory measurements. The validation results of the top-level module were encouraging. Validation of the integrated system using retrospective clinical data is underway.

Keywords: decision support, fuzzy/neural control, model-based reasoning, biomedical systems

1. BACKGROUND

Artificial ventilatory support is an important component of intensive care therapy. Throughout the last three decades, various researchers have tried to develop different advisory systems to help clinicians adjust ventilator settings. This paper presents the design and development of such a system using a hybrid knowledge-and-model-based system, which takes advantage of the superb qualitative decision-making ability of humans and the precision of computer control. The background to the project is given in the first section. The second section describes the overall architecture and design of the system. The third section describes the development and validation of the top-level knowledge-based module while the fourth section describes the lower-level model-based sub-units and their validation. The final section deals with the way in which the system is implemented and integrated into a clinical decision support tool.

1.1. The need for decision support and automation in ICU

It has long been recognised that there is often a significant delay between the report of an abnormal blood gas result and the initiation of the appropriate treatment in intensive care units (ICU) (Simpson et al., 2000). With technological advances, the availability of an increasing number of monitoring systems and the increasing use of patient data management systems (PDMS), clinicians are often bombarded with a vast amount of information. A decision-support and closed-loop control system can reduce the delay in treatment and help clinicians handle the information.

1.2. Previous work

Early ventilator advisory systems used an algorithmic approach (Menn et al., 1973). Such an approach is inflexible and a very large program is often required to cater for the needs of patients with different lung pathologies. Later, a lot of work was done on the use of Artificial Intelligence. Systems such as VentEx (Shahsavar et al., 1985) and NeoGanesh (Dojat et al., 1997) used classical logic in the inference engine. Recent work concentrated on the use of fuzzy logic (Nemoto et al., 1999 and Schuh et al, 2000). Although a knowledge-based approach eliminates the need for an accurate mathematical model, it is prone to subjectivity in the domain experts' opinions and communication errors.

On the other hand, the model-based approach is more objective. However, it is often very difficult to accurately model the respiratory system of ventilated patient with different lung pathologies. The VentPlan (Rutledge et al, 1993) is such a system. The mathematical model used consists of 5 compartments: the alveolar compartment, the pulmonary compartment, the arterial compartment, the tissue compartment and the venous compartment. The equations are based on mass transport while a belief network was used to calculate the probability distributions of the physiological parameters from qualitative and semi-qualitative inputs. The system would then search the space of possible plans and the most optimal plan would be used as the recommendation. Although results were promising in patients with normal lungs, attempts to improve the mathematical model by separating the lungs into more compartments resulted in an unacceptably long computation time.

1.3. Hybrid knowledge-and–model-based systems

To date, most of the ventilatory advisory systems use only one approach in deriving the advice. For example, in the case of fuzzy systems, the fuzzy rule-base is often derived using a knowledge-based approach. The final advice depends solely on the rule-base. In order to develop a system that offers more objectivity than conventional knowledge-based systems and eliminates the need for an extensive and complicated mathematical model, we have adopted a combined approach. Although this combined approach has been proposed as part of the KUSIVAR prototype (Rudowski et al., 1989), the VentEx system subsequently developed by the same research group was a pure knowledge-based system.

When a clinical decision is made for a change in ventilator settings, the decision can be divided into two parts. The first part is the qualitative aspect of the change, for example, the FiO_2 should be increased. The second part is the quantitative aspect, which defines the amount by which the particular setting should be changed. To achieve a target blood gas level, there are often a few options available. For example, to increase the arterial partial pressure of oxygen (PaO_2), one can increase the FiO_2 or increase the PEEP or both. One can also start prone ventilation and prescribe other respiratory therapies. Humans are very good at pattern recognition and can often quickly come to a reasonable conclusion as to which ventilator setting should be changed. On the other hand, due to the complexity of the problem and to the lack of a comprehensive model, it is difficult and time-consuming to use a computer algorithm to find the optimal solution in the domain of possible solutions. However, compared to computers, humans are often less capable of making good quantitative decisions. In clinical practice, by how much a ventilator setting should be changed is very often arbitrarily determined. Although equations and formulae based on respiratory physiology are available, clinicians often find the calculation time-consuming. This is the reason why we have adopted the combined knowledge-and-model-based approach.

2. SYSTEM DESIGN AND DEVELOPMENT

2.1. System Specification

The advisory system will generate advice on 4 ventilator settings: FiO_2, PEEP, PINSP and ventilatory rate. The inputs to the system include the patient's demographic data, routine measurements, blood gases, ventilator settings and the respiratory measurements. In the initial prototype, the data are keyed in by the user. However, in the future, data will be automatically retrieved from the PDMS and the routine and respiratory measurements will be automatically logged into the system.

The system can be operated in a number of different modes. Firstly, in terms of the type of monitoring required, it is divided into the invasive mode and non-invasive mode. Operations in the invasive mode require data from invasive cardiovascular measurements, which are usually acquired through the pulmonary arterial catheter. In the non-invasive mode, the invasive cardiovascular measurements and the related parameters are estimated non-invasively. Secondly, in terms of the level of control, the system can be operated under full advisory mode or clinician-directed mode. In the full advisory mode, the therapeutic goals (target blood gases) and the type of ventilator setting(s) to be changed are determined by the system whereas in the physician directed mode, the clinician directed mode, the clinician can define the target blood gases and/or choose which ventilator setting(s) should be changed.

2.2. System Architecture

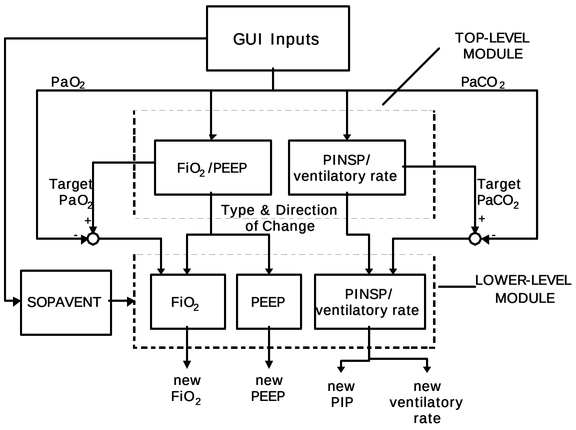

Fig. 1. The architecture of the advisory system.

In order to make the system easy to test and maintain, a modular approach has been adopted. The architecture is shown in Fig. 1. It is divided into two main parts: the top-level knowledge-based module and a lower-level model-based module. Each module is divided into a FiO_2/PEEP sub-unit which controls the oxygenation-related settings and a PINSP/Ventilatory rate sub-unit which controls the settings related to the minute ventilation. The top-level module will advise the type of the ventilator settings to be changed and the target PaO_2 and $PaCO_2$. The lower-level module will derive the amount of change required in each setting.

2.3. Implementation of the System

The system is implemented in LabVIEW 6.1 and MATLAB 5.3/ SIMULINK 3.0. The graphic user interface (GUI) is implemented in LabVIEW, while the neuro-fuzzy inference system and the mathematical models embedded in the lower-level module use the MATLAB fuzzy logic and neural network toolboxes, and SIMULINK block diagrams.

3. TOP-LEVEL MODULE DEVELOPMENT

The top-level module uses fuzzy inference systems to determine the type of ventilator settings to be changed. The rule-bases are derived using an observational approach with the help of a patient simulator. The use of a simulator allowed us to observe how different clinicians change the ventilator settings under the same condition and therefore, improves objectivity of the system.

3.1. Development of the simulator for knowledge acquisition

The patient simulator was implemented in MATLAB 5.3 (Kwok et al., 2001). The behaviour of the model patients were determined by a mathematical model of ventilated patients, SOPAVENT (Simulation of Patients under Artificial Ventilation) (Goode et al., 1998). The model will be described in more detail in Section 4.1. The demographic data and the routine measurements, cardio-respiratory measurements, ventilator settings and blood gases were retrieved from the PDMS of a general ICU. The data of 11 patients were retrieved with 260 sets of blood gases and patient measurements. These data were used to construct the simulated events of 11 simulated patient scenarios. In each simulated event, the model parameters of the SOPAVENT were derived from one set of patient/ ventilator measurements. One of the case scenarios was used as a test case to help the clinicians familiarize themselves with the GUI.

3.2. Architecture of the fuzzy rule-base

There are two sub-units in the top-level module and therefore, two fuzzy rule-bases were derived. The input variables were decided via discussion with the intensive care consultants. For the ventilator settings which primarily affect the oxygenation of the patient, i.e. FiO_2 and PEEP, the inputs include the past and present PaO_2, past and present FiO_2, and the PEEP. For the PINSP and ventilatory rate, the inputs include past and present pH, past and present arterial partial pressure of carbon dioxide ($PaCO_2$), the PINSP and the ventilatory rate.

In order to reduce the number of rules, the input variables were not directly input to the fuzzy inference system but grouped into three variables. For the FiO_2 and PEEP control, the inputs to the fuzzy inference system include the PaO_2, the patient's condition and the support level. The patient's condition is derived from the change in the hypoxemia index (PaO_2/FiO_2) and the support level is derived from the FiO_2 and PEEP.

For the PINSP and the ventilatory rate, the inputs to the fuzzy inference system include the previous $PaCO_2$, the metabolic status and the support level. The metabolic status includes 5 categories: metabolic acidosis, respiratory acidosis, normal, metabolic alkalosis and respiratory alkalosis. It is derived from the pH and $PaCO_2$. The support level is derived from the PINSP and the ventilatory rate.

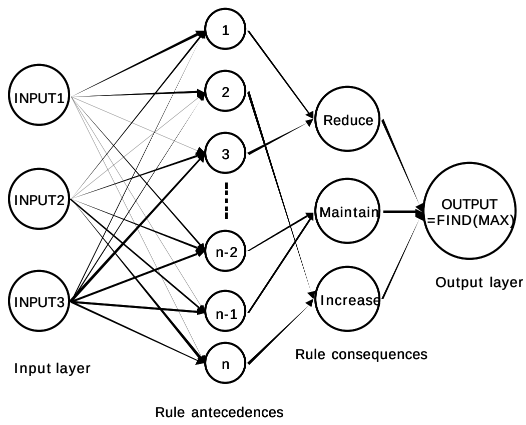

Fig. 2. The architecture of the fuzzy inference system used in the top-level module.

The structure of the fuzzy inference systems is shown in Fig. 2. Grid-partitions were used for the fuzzy rule-antecedents and 'multiplication' was chosen as the inference method. There are three output members for each ventilator setting: reduce, maintain or increase. The output member with the maximum membership value is chosen as the output.

3.3. Derivation of the initial rule-base

4 intensive care consultants were invited to take part in the simulations. The patient's demographic data, routine measurements, ventilator settings, blood gases, cardiovascular measurements and respiratory measurements of each simulated event was presented to each consultant via the GUI. The consultant gave an advice according to this information. The program then calculated the next blood gas measurements based on the model parameters of the next simulated event of the patient. The cycle continued until the consultant had given the advice on the last simulated event of the patient. The data from these simulations were then used to derive the initial fuzzy rule-bases.

Not all the intensive care consultants completed all the patient scenarios due to a lack of time. A total of 32 consultant/cases were completed resulting in a total of 788 simulated events. The data from the simulated events were used in the initial rule derivation. The set of data from each simulated event provided one training data set. However, the inputs to the fuzzy inference system include one past value and this resulted in 756 training data sets. 4 data sets had to be excluded for the training of the FiO_2/PEEP rule-base because the PaO_2 was out of range. 2 data sets had to be excluded for the training of the PINSP/ventilatory rate rule-base because the $PaCO_2$ was out of range. The rule-antecedents of the fuzzy rules were pre-determined via discussions with the clinical experts. The fuzzy memberships in the rule-antecedents were calculated for each training data set. We then examined the relationship between the relative frequency of each consequence (increase, maintain or reduce) and the fuzzy membership value of the antecedents. If there was a significantly positive correlation between the membership value in a rule-antecedent and the relative frequency of a consequence, we concluded that the rule-antecedent should result in that consequence. The initial rule-base derived was then reviewed and slightly modified by an intensive care consultant.

3.4. Tuning of the rule-bases

The validation results of the initial rule-base were not satisfactory for the settings: PEEP, PINSP and ventilatory rate. Therefore, tuning of the rule-base was needed. However, how can we adjust the parameters of the fuzzy rule-base using experimental data? Close examinations of the rule-based system will reveal that it is similar to a neural network. Indeed the inference mechanism is similar to a perceptron. Therefore, the perceptron training rule was used for tuning. The clinicians' simulation results were used as the training data. For each data set, the patient's data became the training inputs and the clinician's advice was used as the targets. In the initial rule-bases, the membership in the consequence associated with a rule-antecedent was either 1 or 0. This was adjusted during the training. The output by the fuzzy inference system was compared to the target during the training. If the output was output member i and the target was output member j, the weight of output member i was adjusted using the formula:

$$\mathbf{w}(n+1) = \mathbf{w}(n) - \mu_{\mathbf{A}} \cdot \mathbf{u} \cdot lr \quad (1)$$

The weight of output member j was adjusted using the formula:

$$\mathbf{w}(n+1) = \mathbf{w}(n) + \mu_{\mathbf{A}} \cdot \mathbf{u} \cdot lr \quad (2)$$

where $\mathbf{w}(n+1)$ is the new weight vector, $\mathbf{w}(n)$ is the old weight vector, $\mu_{\mathbf{A}}$ is the membership value in the antecedents for the particular data set, $\mathbf{u}$ is the input vector and lr is the learning rate. lr is always positive and is set to be between 0 and 1.

During the training, a performance index was introduced to monitor the progress. The performance index was calculated based on the number of exact match and mismatch between the targets and outputs for all the data sets. Conflicts (i.e. the target and the output were in opposition directions: one increases and the other reduces, etc.) between the target and output were penalized. The performance index used is:

$$Perf = n(exact\_match) - n(conflicts) \quad (3)$$

Table 1. The proportion of exact matching, partial matching and conflicting mismatch between the fuzzy inference system outputs and the targets from the training set before and after training.

Settings	Exact Match		Partial Match		Conflicts	
	Pre-training	Post-training	Pre-training	Post-training	Pre-training	Post-training
FiO_2	70.2%	71.4%	29.8%	28.5%	0.0%	0.1%
PEEP	64.6%	79.9%	34.4%	20.1%	0.9%	0.0%
PINSP	54.8%	63.0%	41.9%	35.0%	3.3%	2.1%
Ventilatory rate	82.6%	85.3%	17.2%	14.7%	0.1%	0.0%

The training stopped when a reduction in the performance index was detected. Table 1 shows that after the training, the matching between the fuzzy inference systems and the clinicians' advice outputs (targets) improved for all the ventilator settings.

3.5. Setting the targets

The target PaO_2 and $PaCO_2$ were determined by the outputs of the fuzzy inference system. By definition, the targets are the input blood gas values when no change in any ventilator setting is required. Therefore, the system searches within a range of PaO_2 and $PaCO_2$ where the fuzzy inference system output is 'maintain FiO_2 and PEEP' and 'maintain PINSP and ventilatory rate' respectively.

4. LOWER-LEVEL MODULE DEVELOPMENT

4.1. Physiological and mathematical model: SOPAVENT

The model adopted in the lower-level module is the SOPAVENT. The model equations were based on respiratory physiology and have been used by a number of researchers. There are two sets of equations: oxygen transport equations and carbon dioxide transport equations. They describe the passage of oxygen and carbon dioxide in the 5 compartments mentioned in Section 3.1. They are dynamic equations and solutions are difficult to obtain analytically. Moreover, the clinicians are often more interested in the steady-state blood gas level. Hence, we analyzed the equations at steady-state and used the resulting equations to derive the required ventilator settings.

4.2. Use of Newton's algorithm and non-invasive estimation of shunt to control FiO_2

By evaluating the oxygen transport equations at steady-state, one can derive the Jacobian, which is the first derivative of PaO_2 to FiO_2. This derivative, however, depends on the other patient parameters. The most important parameters include the respiratory shunt, the cardiac output and the oxygen consumption. All of these can only be accurately measured or derived with the help of a pulmonary artery catheter. However, from the sensitivity analysis, it was found that the relationship between the output PaO_2 and the FiO_2 is highly sensitive to the respiratory shunt and is moderately sensitive to the cardiac output and oxygen consumption. Moreover, in critically ill patients, the respiratory shunt can vary from 3 – 50% (a more than 15-fold difference) whereas the cardiac output and oxygen consumption vary to a much lesser extent (a 3-fold difference typically). Therefore, it is crucial to have a good estimation of shunt and it may suffice to use the population mean or median values for cardiac output and oxygen consumption.

The shunt could be derived in two ways. If the patient has a pulmonary artery catheter in situ, one can calculate the shunt from the arterial oxygen content and the mixed venous oxygen content. However, not all the patients have pulmonary artery catheter. Therefore, the shunt has to be estimated using non-invasive data. The respiratory index (the ratio between alveolar-arterial oxygen difference and the PaO_2) has been shown to correlate well with the shunt (Kwok et al., 2001b). The relationship has been shown to be linear over a large range of shunt values although the relationship becomes non-linear at extreme shunt values. The adaptive neuro-fuzzy inference system (ANFIS) performs a non-linear mapping between the inputs and outputs. Data from the ICU were used to provide training and validation data for the ANFIS and an ANFIS model of the relationship between the respiratory index and the shunt was then derived. This provides the method for shunt estimation in the non-invasive mode for the advisory system (Kwok et al., 2002).

The relationship between the PaO_2 and FiO_2 at steady-state can be represented by:

$$PaO_2 = f(FiO_2, \theta) \quad (4)$$

where f is the SOPAVENT model at steady-state and θ are the model parameters including shunt, cardiac output, oxygen consumption and haemoglobin. The model parameters differ for each patient. However, the values either are available from the patient's measurements or can be estimated. For the FiO_2 sub-unit of the lower-level module, once the patient's measurements are keyed in via the GUI, the program will create a patient-specific SOPAVENT model after calculating the model parameters. From this model, the FiO_2 required to achieve the target PaO_2, which has been defined by the top-level module, is then estimated using the Newton's method. The iteration formula used is:

$$FiO_{2,n+1} = FiO_{2,n} - \frac{g(FiO_{2,n},\theta)}{g'(FiO_{2n},\theta)} \qquad (5)$$

$$g(FiO_2,\theta) = f(FiO_2,\theta) - PaO_2 \text{ target} \qquad (6)$$

As a safety measure, the FiO_2 output from the lower-level module is limited to the range of 0.3 to 1.0. The lower level PEEP control sub-unit is not developed yet because a suitable PEEP model is not available. Therefore, the sub-unit only advises a fixed amount of change in PEEP if the top-level module directs it to do so. This amount was determined via consultations with clinical experts. When a suitable PEEP model becomes available, the necessary changes will be made to the sub-unit using model-based control algorithms.

4.3. Optimal Control of Peak Inspiratory Pressure and ventilatory rate

The target $PaCO_2$, which has been defined by the top-level module, can be achieved by altering the ventilatory rate or the tidal volume. In pressure-controlled ventilation, the tidal volume is determined by the PINSP.

At the steady-state, the SOPAVENT model can be simplified and the change in $PaCO_2$ ($\Delta PaCO_2$) becomes:

$$\Delta PaCO_2 = \frac{P_B \dot{V}_{CO_2}}{1000(1-K_D)}\left(\frac{1}{MV} - \frac{1}{MV_0}\right) \qquad (7)$$

where P_B is the barometric pressure, $\dot{V}_{CO_2}$ is the carbon dioxide production rate, K_D is the deadspace fraction, MV and MV_0 are the required minute volume ventilation and the initial minute volume ventilation respectively. The MV is given by:

$$MV = \frac{RR \cdot PINSP \cdot C}{1000}\left(1 - \exp\left(-\frac{60 t_i}{RR \cdot RC}\right)\right) \qquad (8)$$

where RR is the ventilatory rate, t_i is the inspiratory time, R is the airway resistance and C is the airway compliance.

Although one could calculate the required minute volume using equation (7); and if only one of the settings (PINSP or ventilatory rate) needs to be altered, one could derive the necessary change easily, the solution is not that straight-forward if two settings need to be changed concurrently. Moreover, both excessive PINSP and excessive ventilatory rate can exert adverse effects on the patient's lungs. Unlike FiO_2, the safety margins of PINSP and ventilatory rate are much smaller. Therefore, we introduce a cost function is the form of:

$$J = \left(\frac{\Delta PaCO_2 \text{ error}}{\text{target}\,\Delta PaCO_2}\right)^2 + \lambda_1\left(\frac{\Delta RR}{RR}\right)^2 + \lambda_2\left(\frac{\Delta PINSP}{PINSP}\right)^2 \qquad (9)$$

where λ_1 and λ_2 are positive constants between 0 and 1. When the top-level module advises a change in PINSP only, λ_1=0 and when it advises a change in ventilatory rate only, λ_2=0. Then the program will minimize the cost function and Fig. 3 shows how the advised PINSP varied with different λ_1 and λ_2. With a suitable selection of λ_1 and λ_2, one can balance the need to get closer to the target $PaCO_2$ and to avoid using excessive PINSP or ventilatory rate.

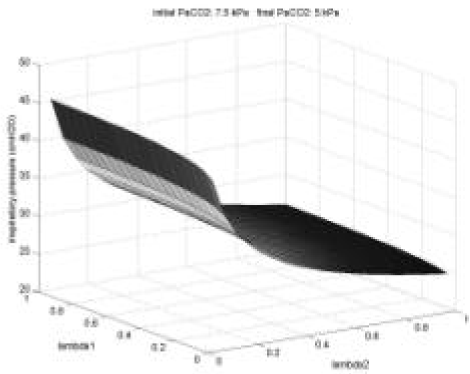

Fig. 3. The change in PINSP advised by the lower-level module to achieve a target $PaCO_2$ of 5kPa from 7.5 kPa with different values of λ_1 and λ_2.

5. SYSTEM INTEGRATION AND THE FUTURE

The top-level module and the sub-units of the lower level module are implemented in MATLAB scripts and SIMULINK. The LabVIEW provides the graphic user interface and the flow control of the program. The

MATLAB/SIMULINK programs for all the components are accessed from LabVIEW using MATLAB script nodes. The top-level module is completed and most of the lower level module is integrated into the system. As mentioned before, a good PEEP model is required for the PEEP control sub-unit of the lower-level module. Nevertheless, the validation using retrospective clinical data is already underway for the integrated system. In the future, the system will be revalidated after modification of the PEEP control sub-unit.

REFERENCES

Dojat M., F. Pachet, Z. Guessoum, D. Touchard, A. Harf, L. Brochard (1997). NeoGanesh: a working system for the automated control of assisted ventilation in ICUs. *Artificial Intelligence in Medicine*, **11**, 97-117.

Goode K.M, D.A. Linkens, P.R. Bourne, J.G. Cundill (1998). Development of a fuzzy rule-based advisor for the maintenance of mechanically ventilated patients. *Biomedical engineering: applications, basis & communication*, **10**, 236-246.

Kwok H.F., K.M. Goode, M. Mahfouf, G.H. Mills, D.A. Linkens (2001a). Development of a simulator for control assessment of PaO2 in ICU. In: *UKSim 2001, Conference of the United Kingdom Simulated Society*. (D. Al-Dabass. (Ed)), 161-167. The United Kingdom Simulation Society, Nottingham Trent University, UK.

Kwok H.F., D.A. Linkens, M. Mahfouf, G.H. Mills (2001b). An adaptive approach to respiratory shunt estimation for decision support in ICU. In: *Adaptive Systems and Hybrid Computational Intelligence in Medicine: Special Session Proceedings of the EUNITE 2001 Symposium, Tenerife, Spain.* (D.D. Georgios, D.A. Linkens (Eds.)), 59-64. University of Aegean, Greece.

Kwok H.F., D.A. Linkens, M. Mahfouf, G.H. Mills (2002). Adaptive FiO_2 advisor: use of non-invasive estimation of shunt. Submitted to the Artificial Intelligence in Medicine.

Menn S.J., G.O. Barnett, D. Schmechel (1973). A computer program to assist in the care of respiratory failure. *JAMA*, **223(3)** , 308 - 312.

Nemoto T., G.E. Hatzakis, C.W. Thorpe, R. Olivenstein, S. Dial, J.H.T. Bates (1999). Automatic control of pressure support mechanical ventilation using fuzzy logic. *American Journal of Respiratory and Critical Care Medicine*, **160**, 550-556.

Rudowski R., C. Frostell, H. Gill (1989). A knowledge-based support system for mechanical ventilation of the lungs. The KUSIVAR concept and prototype. *Computer Methods and Programs in Biomedicine*, **30**, 59-70.

Rutledge G.W., G.E. Thomsen, B.R. Farr, M.A. Tovar, J.X. Polaschek, I.A. Beinlich, L.B. Sheiner, L.M. Fagan (1993). The design and implementation of a ventilator-management advisor. *Artificial Intelligence in Medicine*, **5**, 67-82.

Schuh C., C. Zelenka, M. Hiesmayr (2000). FuzzyKBWean: A computer program for fuzzy knowledge-based weaning of ICU patients. In: *Fuzzy diagnostic and therapeutic decision support. Proceedings of the ERUDIT-Workshop, Austria, May 11-12,2000*. (K.P. Adlassnig, (Ed)), 38-44. Osterreichische Computer Gesellschaft, Austria.

Shahsavar N., U. Ludwigs, H. Blomqvist, H. Gill, O. Wigertz, G. Matell (1995). Evaluation of a knowledge-based decision support system for ventilator therapy management. *Artificial Intelligence in Medicine*, **7**, 37-52.

Simpson C.L., D.A. Linkens, G.H. Mills (2000). Response times to abnormal respiratory parameters during weaning and non-weaning phases of mechanical ventilation on the ICU. *American Journal of Respiratory and Critical Care Medicine*, **161**:A559.