Neutrophil count prediction for personalized drug dosing in childhood cancer patients receiving 6-mercaptopurine chemotherapy treatment

Dalila Avdic, Michael Gallimore, Mike J. W. Riley, Chris Bingham
School of Engineering
University of Lincoln
Lincoln, U.K.
davdic@lincoln.ac.uk

Acute Lymphoblastic Leukaemia (ALL) is a common form of blood cancer that usually affects children under 15 years of age. Chemotherapy treatment for ALL is delivered in three phases viz. induction, intensification, and maintenance. The maintenance phase involves oral administration of the chemotherapy drug 6-Mercaptopurine (6-MP) in varying doses to destroy any remaining abnormal cells and prevent reoccurrence. A key side effect of the treatment is a reduction in neutrophil counts which can lead to a condition known as neutropenia. This carries a risk of secondary infection and has been linked to 60% ALL fatalities. Current practice aims to control neutrophil counts by varying 6-MP dosages on a weekly basis and is based upon clinical judgment and experience of the medical professionals involved. Conceived as a decision support aid for clinicians then, presented are the results of a machine learning technique that predicts neutrophil counts one or more weeks ahead using data from ALL blood test results and 6-MP dosing. In this work, a model is trained and validated using data from a single female ALL patient's maintenance phase. The prediction error is found to be typically within +/- 290/microL at one week and within +/- 820/microL for a 14 day prediction.

***Index Terms*—Leukaemia, Neutrophils, Artificial Neural Networks, Time Series Prediction**

I. INTRODUCTION

Acute Lymphoblastic Leukaemia (ALL) is the most common form of cancer in children and represents 80% of all leukaemia cases [1]. As the number of children with this type of leukaemia grows worldwide so does the demand for research into more effective treatment regimens. ALL is a form of leukaemia that affects lymphocytes; a type of white blood cell. ALL is an overproduction of immature lymphocytes called lymphoblast, or blast cells. These cells flood the bone marrow and prevent the body from producing the correct amounts of healthy cells in order to function normally. If untreated, ALL progression is rapid and requires aggressive chemotherapy treatment. The protocol for treating ALL typically spans 2 years for girls and 3 for boys. It consists of three phases of chemotherapy treatment viz. induction, intensification and maintenance - with the latter being the most protracted, lasting around 18 months for girls and 30 months for boys.

The induction phase is designed to achieve initial remission and involves intravenous administration of various drugs including vincristine, methotrexate and dexamethasone. Intensification is the most concentrated phase of treatment and is aimed at destroying remaining abnormal cells. The aim of the maintenance phase is to kill any remaining cells left over from the first two phases to minimize the chance of relapse. During this phase, patients receive doses of the chemotherapy drug 6-mercaptopurine (6-MP), typically administered orally on a daily basis [2]. 6-MP's effect is to damage the RNA or DNA thereby disrupting the natural division process, ultimately resulting in cell death. A primary side effect of this treatment, however, is a reduction in neutrophil counts which can lead to neutropenia (abnormally low level of neutrophils) and so presents a risk for acquiring secondary infection. Typically, neutropenia results in regular hospitalizations and/or unscheduled breaks in the treatment regime. Also, patients with neutropenia are at high risk of secondary infections with associated fatalities of 60% [3].

Decreasing drug dosages can reduce the risk of neutropenia, but can degrade treatment effectiveness. Conversely, increasing drug dosages where appropriate can be beneficial but dosing regimens are not always able to support such decisions with a satisfactory degree of accuracy. Medical professionals aim to control counts by varying 6-MP dosages on a weekly basis. Typically the aim is to achieve neutrophil counts between $[1,1.5] \times 10^9$ neutrophils per liter of blood. Patients with a count smaller than 0.5×10^9 are classed as neutropenic, i.e. no effective immune system at which point treatment is stopped until counts increase above the minimum threshold. Weekly blood counts are used to inform weekly dosing decisions

and typically patients are prescribed 100%, 75% or 50% of the calculated dose (per kg body mass).

Hence, the motivation of this work is to explore a machine learning technique that can predict, at least one week ahead, neutrophil counts. It is hoped that further research and development leads to supporting clinical decisions of 6-MP dose manipulations to reduce instances of neutropenia in children with ALL.

To the authors' knowledge, currently there is not wide reporting of methods for predicting neutrophil counts using a machine learning technique such as that presented herein. In the literature, such techniques have shown success when applied to other medical conditions.

An anti-diabetic drug failure prediction methodology for type 2 diabetes investigated by Kang, S. *et al.* [4] using support vector machines (SVM) proposed an ensemble of SVM for a large scale dataset, reporting a prediction accuracy of about 80%. Menden, M.P. *et al.* [5] also proposed a drug prediction methodology using analysis of variance (ANOVA) which aimed to predict patient response to a specific cancer therapy. Lin, C. *et al.* [6] carried out an investigation in applying neural networks to predict the likelihood of patient response to clozapine during the treatment of schizophrenia. The research showed that all clozapine responders and approximately 75% of non-responders were successfully predicted by the resulting artificial neural network (ANN) model. A similar study into the prediction of clozapine response was carried out by Khodayari-Rostamabad, A. *et al.* [7]. Machine learning techniques were applied to pre-treatment electroencephalography (EEG) data from schizophrenia patients in order to predict the likely response to clozapine therapy in adults suffering from schizophrenia. These techniques were able to predict, in advance of the first dose, whether a patient will or will not respond to the powerful but potentially toxic medication. The level of performance using the leave-one-out validation method was ≈85%. Yuan Li *et al.* [8] have developed a data-driven predictive system using machine learning techniques. The framework has been validated *in vitro* though experimental study with *Giardia lamblia* and the system categorizes the set of data using Fuzzy c-means clustering algorithm. It used a Probabilistic Suffix Automaton (PSA) to model the temporal state sequences. The accuracy of the system was 73% with four data points and 97.5% with nine data points.

This paper describes a time series prediction technique using an artificial neural network to predict future neutrophil counts based on measured blood count data and 6-MP dosing. Compared is the accuracy of the predicted neutrophil count with known (i.e. expected) counts.

II. BACKGROUND

Time series prediction, used for the current work, is a machine learning technique that is commonly used in weather forecasting, business planning, economic predictions and signal processing. Data from previously observed system states are used to create a model to predict future states. Although there are a large number of algorithms demonstrating high reliability in prediction, difficulties can arise when attempting to model systems that are highly non-linear, such as in drug prediction. In other fields, ANNs have been shown to provide accurate prediction results in such cases [9-12]. They represent a group of statistical algorithms inspired by biological neural networks that are able to learn system behavior from a sufficient number of inputs. In a biological system, the transmission of a signal from one neuron to another, through a synapse, releases specific transmitter substances. The outcome is to lower or raise the electrical potential inside the body of the receiving cell. Once the electrical potential reaches a threshold the neuron will fire. It is this underlying process that the artificial neuron tries to mimic [13], as depicted in Fig. 1 [9] [14].

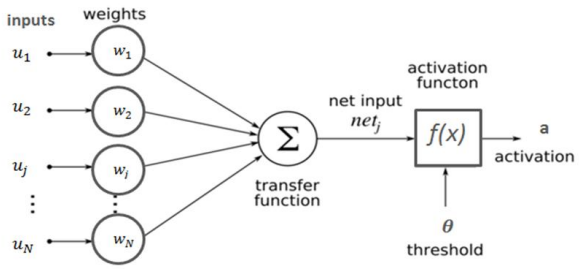

Fig. 1 Artificial Neuron Structure

The neuron has $x_1, x_2, x_j \ldots x_n$ inputs each attributed a weight $x_1, x_2, x_j \ldots x_n$. Input weightings are dependent on each individual inputs contribution to the output prediction and are analogous to the synaptic connections in biological neurons. The activation corresponding to the graded potential is given by [8], Eq. 1

$$a = \sum_{j=0}^{N} w_j u_j + \theta \quad (1)$$

where θ represents the threshold in the artificial neuron. Typically, the modelling of complex functions is not achievable through a single artificial neuron and so layers of artificial neurons are formed, where the outputs from many neurons are connected as inputs to the others and thus building a neural network. When making such a network,

the formula above can be modified by expressing the activation i^{th} neuron as Eg. 2,

$$a_i = \sum_{j=0}^{N} w_{ji} x_j + \theta_i \quad (2)$$

where, x_j is either the output of another neuron or an external input. Neural Networks are made up of a number of layers, the first being the input layer, the last the output layer and all layers between are the hidden layers. Each layer also carries a weight determined during the training phase.

In this paper we use Nonlinear AutoRegressive with eXogenous inputs of an ANN model (NARX) in Matlab 2012b. In a NARX network, the output signal of the network composes the input vector of the network using delay operators. A mathematical formulation of the output final response is expressed below:

$$y(n+1) = f[y(n), y(n-1), \ldots,$$

$$y(n-d_y+1);\ x(n), x(n-1), \ldots,$$

$$x(n-d_x+1)] = f[y(n); x(n); W] \quad (3)$$

Where x(n) and y(n) are the components of the input and output vector respectively. The delays being d_x and d_y, W is the matrix of the adjustable weights and f is the unknown nonlinear function.

III. METHODOLOGY

A clinical dataset from the maintenance phase treatment of one female ALL patient is used in the current work. Data consists of multiple full blood counts and blood differentials along with corresponding 6-MP dosages. Table I shows the ranges of the data in the medical records used.

TABLE I. INPUT PARAMETERS

Input	Range
Days of treatment	[1,588]
Hemoglobin (g/dL)	[78,169]
White Cell Count (10*9/L)	[0.4,11.9]
Platelets (10*9/L)	[87,507]
Red Cell Count (10*12/L)	[2.03,5.05]
Mean Cell Volume (fL)	[89,104]
Hematocrit	[0.134,0.496]
MCH (pg)	[28.9,34.6]
MCHC (g/dL)	[315,354]
Lymphocytes (10*9/L)	[0.25,2.83]
Monocytes (10*9/L)	[0,1.2]
Eosinophils (10*9/L)	[0,0.4]
Basophils (10*9/L)	[0,0.1]
6-marcaptopurine (mg)	[0,80]
Neutrophil counts (10*9/L)	[0,8.43]

The measured neutrophil count for the full 588 day treatment period is shown in Fig. 2 in which the graph markers indicate days when bloods were sampled.

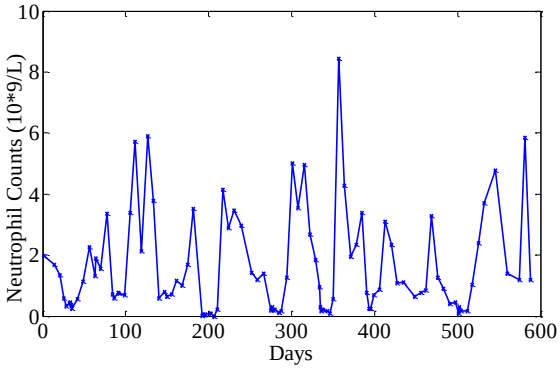

Fig. 2 Neutrophil counts during the Maintenance phase

To improve the accuracy of prediction, the dataset in Table I is first interpolated producing 588 samples over the treatment period. Data normalization, Eq. 4, is necessary prior to training

$$\bar{X} = \frac{X - X_{min}}{X_{max} - X_{min}} \quad (4)$$

where X is the actual value of the sample and X_{max} and X_{min} are the maximum and minimum values. Figure 3 shows a graph of the normalized and interpolated dataset.

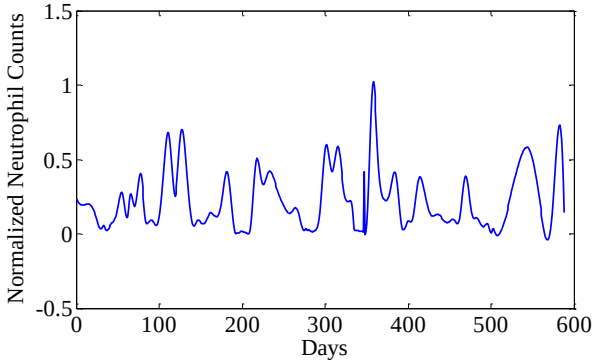

Fig. 3 Neutrophil counts after normalization and interpolation

The dataset is sub-divided as follows: 70% training, 30% for validation. Following training, the ANN is used to predict the patient neutrophil count one or more weeks ahead, and the resulting prediction compared with the actual count. The accuracy of two-week and three-week ahead predictions are also investigated. This opens the possibility of fortnightly blood tests and dosing should neutrophil counts be relatively high and if 14 day predictions from the ANN be consistently of sufficient accuracy.

IV. RESULTS AND DISCUSSION

A. *One week ahead neutrophil prediction*

Approximately four months of training data has been found to be the minimum threshold for training the NARX ANN. Training datasets of this size have been used to produce all the results in the following graphs and tables. Three different training datasets were created from the normalized blood test data in the 4 months preceding days 113, 233 and 352 and Figs. 4, 5 and 6 graph the predictions for the values in the next 7 days. The dashed lines display the close fit between the ANN's prediction of future values and the expected (known) values.

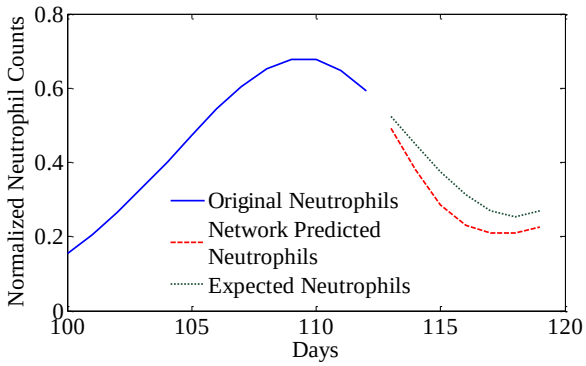

Fig. 4 Normalized neutrophil counts. Prediction period: 7 days

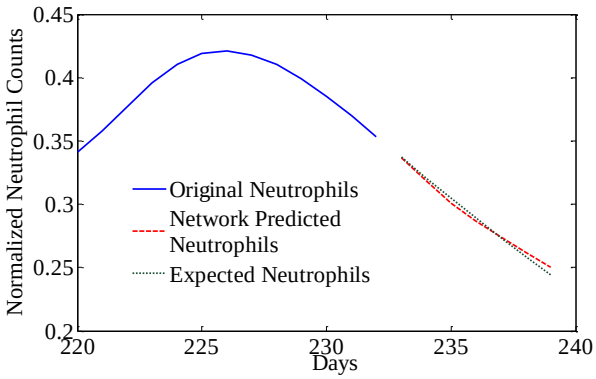

Fig. 5 Normalized neutrophil counts. Prediction period: 7 days

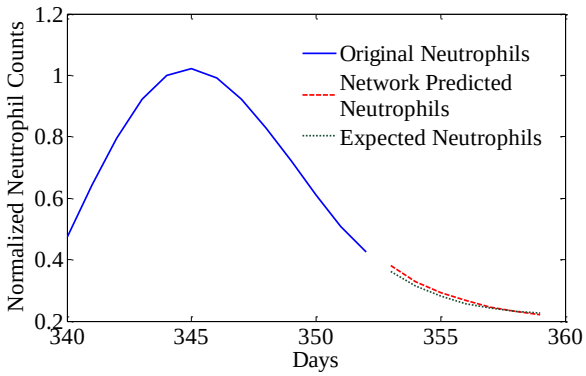

Fig. 6 Normalized neutrophil counts. Prediction period: 7 days

B. *Two weeks ahead neutrophil prediction*

Extending the prediction period to two weeks shows some degradation in the quality of the fit but Figs. 7, 8 and 9 do show qualitatively a reasonable agreement with the expected (known) value 14 days ahead.

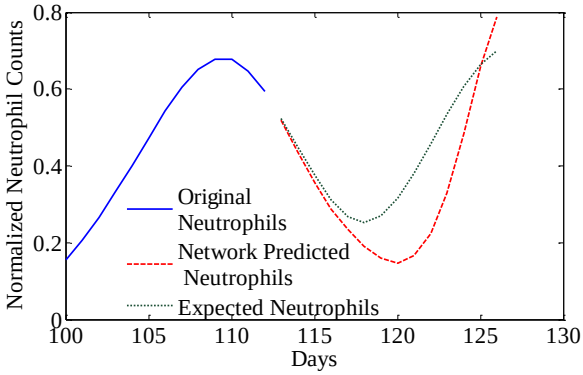

Fig. 7 Normalized neutrophil counts. Prediction period: 14 days

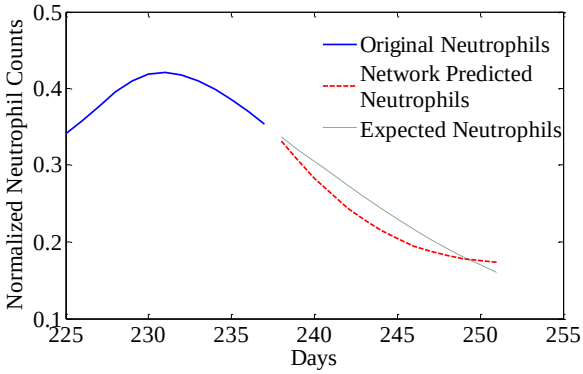

Fig. 8 Normalized neutrophil counts. Prediction period: 14 days

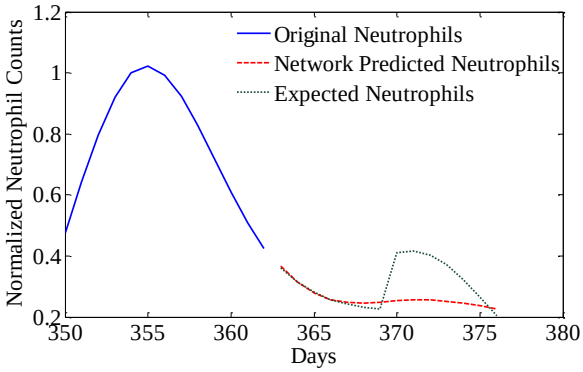

Fig. 9 Normalized neutrophil counts. Prediction period: 14 days

When analyzing the results, a reduction in accuracy was expected as the time period was increased – predictions of the future get more difficult in proportional to the size of look-ahead time period. There is also a degree of caution

necessary when utilizing a 'black-box' function such as neural network for prediction. This is illustrated by inspecting Fig. 7 and Fig. 8 and comparing with Fig. 9. Visible in Fig. 9 is a significant deterioration in the quality of the prediction after day 369.

For a quantitative analysis of the expected error, 10 unique predictions were made, each of three different time periods; 7, 14, and 21 days ahead. Those 30 predictions were then compared with measured blood test results - not interpolated or normalized data. The standard error was found and used to produce a 95% confidence interval (CI), Table II and Fig. 10. As expected, the bounds of our likely prediction error increase (≈ 2.8 times) in transitioning from a 7 to a 14 day prediction. We also judge from this analysis the 21 day prediction to be unreliable at the time of writing.

TABLE II. Errors in the ANN time series prediction

	Normalized error		**Error in Neutrophil prediction and 95% confidence intervals**		
Prediction (days ahead)	***10^3 MSE***	***10^2 RMSE***	***Lower***	***Mean (10^9 / L)***	***Upper***
7	2.87	5.36	-0.27	0.00523	0.285
14	22.5	15.01	-0.74	0.03620	0.812
21	147.0	38.38	-2.73	-0.88954	0.953

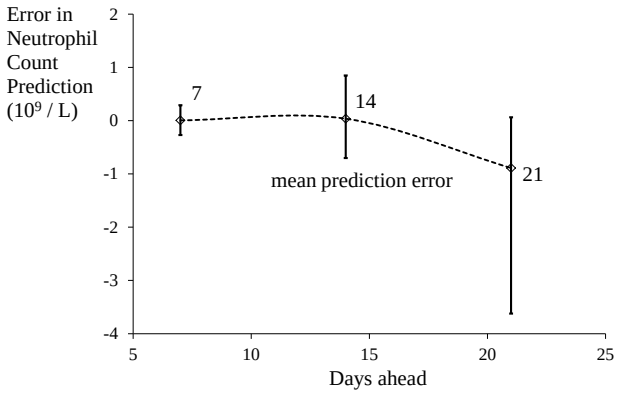

Fig. 10 Errors in 1, 2 and 3 week prediction (95% CI as error bars)

V. CONCLUSION AND FUTURE WORK

Predicting neutrophil counts with sufficient accuracy a week or more ahead is highly desirable. The potential is significant: aiding clinicians in reducing the risk of neutropenia thereby facilitating improved treatment success and reducing the number of Acute Lymphoblastic Leukaemia deaths brought about because of secondary infections.

In this paper, a description of an Artificial Neural Network tasked with predicting neutrophil count in Acute Lymphoblastic Leukaemia prediction has been presented. Blood tests results and 6-MP dosages from a 588 day treatment period for a female ALL patient were used for the training.

The results show effective prediction 7 days ahead using relatively modest training datasets (4 months historic data). Prediction accuracy degrades with increasing time period. The current work indicates that 21 day ahead predictions are not to be relied upon with the current implementation.

If machine-based learning techniques such as these can aid clinicians in decision making, then possibilities of time-off treatment and time spent in hospital, plus the associate costs, are reduced along with the potential from improved clinical outcomes for the patient.

In further work, we intend to improve the algorithm and gather more patient datasets to see if this method can be generalized with a view to improving its accuracy and performance. The authors also recognize that, not only could this work stream be significant in terms of enhancing the treatment of ALL, but that possibilities exist to extend this method for other drug dosing regimens where control is required, such as in the treatment of other types of cancer and clozapine treatment for schizophrenia.

REFERENCES

[1] A. Radaelli, BL. Laskin, JM/ Stephens, MF. Botteman and CL. Pashos, "A systematic literature review of the clinical and epidemiological burden of acute lymphoblastic leukaemia (ALL)," *Eur J Cancer Care (Engl).*, vol. 14, no. 1, pp. 53-62, 2005.

[2] "Chemo care," [Online]. Available: http://chemocare.com/chemotherapy/drug-info/6-mp.aspx. [Accessed 30 09 2014].

[3] G. P. Bodey, M. Buckley and Y. Sathe, "Quantitative relationships between circulating leukocytes and infection in patients with acute leukemia, " *Ann Intern Med*, vol. 64, no. 2, pp. 328-340, 1966

[4] S. Kang, P. Kang, T. Ko, S. Cho, S.Rhee and K. Yu, "An efficient and effective ensemble of support vector machines for anti-diabetic drug failure prediction," *Expert System with Applications*, vol. 42, no. 9, pp. 4265-4273, 2015.

[5] M. P. Menden, F. Iorio, U. M. M. Garnett, C. H. Benes, P. J. Ballester and J. Seaz-Rodriquez, "Machine Learning Prediction of Cancer Cell Sensitivity to Drug Based on Genomic and Chemical Properties," *Plos One*, vol. 8, no. 4, 2013.

[6] C. Lin, Y. Wang, J. Chen, Y. Liou, Y. Bai, I. Lai, T. chen, H. Chiu and Y. Li, "Artificial neural network prediction of clozapine response with combined pharmacogenetic and clinical data," *Computer methods and programs in biomedicine*, vol. 91, no. 2, pp. 91-99, 2007.

[7] A. Khodayari-Rostamabad, G. M. Hasey, D. J. MacCrimmon and J. P. Reilly, "A pilot study to determine wether machine learning methodologies using pre-treatment electroencephalography can predict the symptomatic response to clozapine therapy," *Clinical Neurophysiology*, vol. 121, no. 12, pp. 1998-2006, 2010.

[8] Y. Li, S. C. Lenaghan and M. Zhang, "A Data-Driven Predictive Approach for Drug Delivery Using Machine Learning Techniques," *Plos One*, vol. 7, no. 2, 2012.

[9] A. Kar, "Stock Prediction using Artificial Neural Networks," *Dept. of Computer Scienece and Engineering*, IIt Kanpur.

[10] E. Diaconescu "The use of NARX Neural Network to predict Chaotic Time Series," *WSEAS Tans. Comput. Res.* vol. 3, no. 3, pp. 182-191, 2008.

[11] M. S. Al-haik, M. Y. Hussaini, H. Garmestani, "Prediction of nonlinear viscoelastic behavior pf polymeric composites using artificial neural network," *International Journal of Pharmaceutics*, vol. 22, no.7, pp. 1367-1392, 2006.

[12] C. Xia, J. Xiu, "RBF ANN Nonlinear Prediction Model Based Adaptive PID Control of Switched Reluctance Motor Drive," *Neural Information Processing*, vol. 4234, pp. 626-635, 2006.

[13] G. M. Maggiora, D. W. Elrod and R. Trenary, "Computational neural networks as model free mapping device," *J. Chem. Inf. Comp. Sci.*, vol. 32, no. 6, pp. 732-741, 1992.

[14] J. Dayhoff and J. DeLeo, "Artificial neural networks: Opening the black box," *Cancer*., vol. 91. no. S8, pp. 1615-1635, 2001.