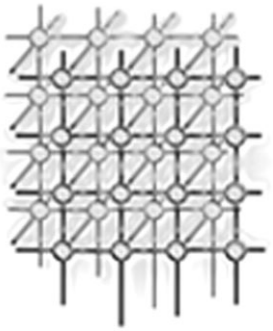
CONCURRENCY AND COMPUTATION: PRACTICE AND EXPERIENCE
Concurrency Computat.: Pract. Exper. 2007; **19**:251–263
Published online 11 May 2006 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/cpe.1045



e-Science and artificial neural networks in cancer management

S. D. Dolgobrodov[1], R. Marshall[1,*,†], P. Moore[1], R. Bittern[2], R. J. C. Steele[2] and A. Cuschieri[3]

[1]*School of Physics and Astronomy, The University of Manchester, Manchester M13 9PL, U.K.*
[2]*Department of Surgery and Molecular Oncology, The University of Dundee, Dundee DD1 9SY, U.K.*
[3]*Scuola Superiore S'Anna di Studi Universitari e di Perfezionamento, Piazza dei Martiri della Libertá n. 33, 56127 Pisa, Italy*

SUMMARY

We describe the origins of this project, its aims and its relevance to e-Science research. Particle physicists at the University of Manchester with experience of artificial neural networks (ANNs) have collaborated with clinicians at the University of Dundee to produce an ANN that is intended to predict survival rates and to indicate management profiles for cancer patients. Comparisons are made between typical data handling problems in particle physics and health care. The problems associated with data procurement, namely reliability and censoring are described, together with a discussion of how these problems were addressed. The inputs to the ANN and its decision output are discussed. The reliability of the ANN is assessed quantitatively. The prototype secure Web-based interface, which allows clinicians to input new patient data to the central node at the University of Manchester and to obtain prognoses from anywhere in the world is presented. For each topic, the e-Science relevance is described and underlined. Copyright © 2006 John Wiley & Sons, Ltd.

Received 21 January 2005; Revised 8 June 2005; Accepted 24 August 2005

KEY WORDS: cancer management; artificial neural network; e-Science

1. INTRODUCTION

This project was enabled by the 2002 e-Science initiative through the Medical Research Council (MRC) and the Particle Physics and Astronomy Research Council (PPARC) who awarded a joint

*Correspondence to: Professor Robin Marshall, School of Physics and Astronomy, The University of Manchester, Manchester M13 9PL, U.K.
†E-mail: robin.marshall@manchester.ac.uk

Contract/grant sponsor: Medical Research Council (MRC)
Contract/grant sponsor: Particle Physics and Astronomy Research Council (PPARC)

grant to the groups in Dundee and Manchester. At the time, PPARC interpreted e-Science more broadly than the original U.K. Government definition [1]. The Government defined e-Science as 'Science increasingly done through distributed global collaborations enabled by the Internet, using very large data collections, tera-scale computing resources and high-performance visualisation'. PPARC interpreted this definition more widely:

> '. . . to include computational and data Grid applications, middleware developments and essential hardware procurement. PPARC believes that e-Science, in the form of developing resources for data and information management and processing across local, national and global networks, will play a major role in the delivery of its future programme. Forthcoming developments in the scientific programme . . . will generate massive quantities of data that must be made available in a seamless way to scientists across the research communities, requiring advanced software tools.'

It was also expected that in 2002 most of the PPARC e-Science funding would be in support of people developing advanced informatics techniques, especially to meet the computing needs for experiments at the Large Hadron Collider (LHC) at CERN, Geneva. Furthermore, it was explicit in the proposal to MRC and PPARC that led to this project, that the research work was not directly Grid orientated but would address other issues of e-Science.

Others have widened or generalized the definition further. Thus, the National e-Science Centre in Edinburgh [2] states:

> 'In the future, e-Science will refer to the large scale science that will increasingly be carried out through distributed global collaborations enabled by the Internet. Typically, a feature of such collaborative scientific enterprises is that they will require access to very large data collections, very large-scale computing resources and high-performance visualization back to the individual user scientists.'

In this paper, the emphasis is placed more on size than in either the Government or PPARC definitions. In announcing its programme, the National Environment Research Council (NERC) were less specific [3]:

> 'In terms of research themes, NERC is not prescriptive and interprets the Government's definition of e-Science widely.'

These few examples suffice to demonstrate that e-Science is not confined to a straight-jacket. In this paper, we address e-Science issues such as the handling of distributed data, the linkage between distributed research teams, the tools and the objects that the tools operate on.

An Internet/Web based artificial neural network (ANN) has been developed for the use of practising clinical oncologists and medical researchers as part of an on-going programme aimed at the implementation of advanced neural networks for prognostic estimates and eventually for management and treatment decisions for individual patients with colorectal cancer. Historically, this would be carried out by clinicians relying on their skills, memory and conventional practice. The aim here is to provide an automated system, accessible globally, having perfect memory of all available, relevant clinical data and that provides an instant decision. In our project, an interdisciplinary team of academic oncologists and physicists has configured and implemented a partial likelihood ANN and trained it

DOI: 10.1002/cpe

to predict cancer-related survival in patients with confirmed colorectal cancer. They have worked together using the internet as an enabling tool. Comparisons of the use of neural networks in particle physics and this project are considered in this paper, as well the respective size and flow of data. A database of 58 220 patient records was obtained, from which 2778 were used for close scrutiny and analysis, with 1558 used eventually for training. With the necessary approval of the Privacy Advisory Committee, the database (Scottish Colorectal Cancer Audit Database) collected by the Scottish Cancer Therapy Network and Clinical Resource and Audit Group was used for this analysis. The predicted survival curves obtained as the output from the ANN showed close agreement with observed actual survival rates of a cohort of patients with various grades of risk of dying from the cancer. Although the data were partially incomplete, the Web-based ANN system accurately predicts survival after staging and treatment of colorectal cancer. The probability of making correct decisions has been investigated using the method of 'receiver operation curves'.

The aim of this project was to use some techniques that are common in particle physics decision making and analysis and apply them to health care. Specifically, the use of ANNs was considered appropriate by the physicists and clinicians working on this project. Of course, ANNs are used in health care but there is a significant body of opinion that they are not to be totally trusted, since the way they work is not transparent to the inexperienced user.

ANNs are used in particle physics for a variety of applications where a decision is needed on the basis of a varying and complex array of input parameters. The data flow in existing particle physics experiments is prodigious and will become extremely prodigious when the next accelerator—the LHC—becomes productive in 2007. A typical particle physics project, in which physicists from this ANN project participate, has a primary data flow of about 50 terabytes per second. It is impossible to store these data and so decisions have to be made, as the data flow in, whether to keep a particular data block or not. A variety of techniques, operating in real time, are used to filter out the unwanted background and one of these techniques is to pass the data through an array of processors that apply ANN computer code to the data patterns in order to reach a decision. Naturally, these decisions must be free from bias. Part of the rationale of these experiments is to observe rare, exciting physical processes, never before witnessed by humans. It is unacceptable that the ANN should discard the records of these phenomena, or distort them, before the scientists can observe and analyse them. Therefore they do not. The end product is data being stored at a rate of about 1 petabyte per year for subsequent analysis, a reduction by a factor of up to a million on the primary data flow. In the near future, decision making of this sort will be made within a local Grid at the site of the experiment before the data are distributed into wider global Grids.

There is no comparable circumstance in health care where ANNs are used like this. The health care data are static and if an error is made, the data can be processed again. The data are generally sparse, even unreliable, leading to problems in training the ANN. The data also usually involves human parameters, life and death and disease management where personal experience, opinions, emotions and faith might play a role. More importantly, the ANN might encroach into an area that is traditionally the domain of learned, expert clinical opinion—that of management of a particular patient. It is this key element that most distinguishes the particle physicists from clinicians: the particle physicist could not handle the data without devolving the microsecond decision making onto silicon chips, whereas the clinician might not wish to.

From the outset, the intention was to use the contents of databases held in Scotland as training data for suitably constructed ANNs. The e-Science challenges were as follows:

DOI: 10.1002/cpe

- to convert the data, much of it on paper, into a descriptive format, i.e. into the form of the components of a numerical vector;
- to satisfy the requirements of the database guardians that confidentiality will be maintained;
- to identify biases in the raw data caused by, for example, different reporting practices in the different regions of Scotland;
- to identify censoring and to devise an acceptable way to handle it;
- to identify errors in the data and to respond accordingly.

The clinicians in Dundee procured the data from various sources and prepared it for the ANN analysis; the physicists in Manchester produced the ANN to process them. The intention was to apply these analytical techniques to health care with special reference to patients suffering from common solid cancers, where there is increasing complexity in the staging of these cancers, requiring specialist knowledge and both multi-disciplinary and multi-professional management. We now believe that we can make such analytical systems more readily available to clinicians by exploiting Grid-secure technology, which has the potential to link large clinical and scientific data sets of cancer patients from various sources and institutions to a common analytic node. This takes the project beyond its original non-Grid scope.

2. ANNS IN HEALTH CARE

To date, ANNs of varying complexity and types have been used, mainly in clinical research rather than routine clinical oncology. Their usefulness has been in the diagnosis, monitoring the spread of the disease and for their prognostic value in breast, ovarian, gastrointestinal, bronchial and prostatic cancers [4–6]. In breast and colorectal cancers, ANNs have been shown to be significantly more accurate in predicting survival than in predicting spread from the primary cancer site [7]. To date, there have been no reported studies on the use of ANNs to formulate management plans for the treatment of patients with cancers and this remains a prime, eventual aim of this project. This aspect will assume greater importance as the proteomic characterization and biology of individual solid cancers becomes more detailed, having repercussions on the complexity of multi-modal treatment. There is evidence [5] that the results obtained with ANNs from one institution can be applied to other centres. Hence, the ANN approach should lead to better standardization and valid comparison of outcomes following specific therapies of solid cancers from different institutions, indeed from different countries.

So far, we have implemented the basic platform—a Web-based ANN that has been trained by exposing it to sets of existing data on one type of solid cancer (colorectal), where the clinical outcome of the patients included in the database is known over a five-year follow-up period.

3. WORKING METHODS

3.1. Meetings and video conferencing

The members of the two teams in Dundee and Manchester have only met physically on a few occasions. Most of the project meetings have been carried out by video conference. Initially this was done via an ISDN dial-up link, involving recurrent costs but this became rapidly obsolete. The Manchester team now has three alternative variants of video conference that are used regularly.

DOI: 10.1002/cpe

(1) A 'Polycom view station SP 128' that can enable both point-to-point meetings over IP and H.323 protocol (IP). This means that we can simultaneously connect with several institutes, although for this project, only a link between Manchester and Dundee was required.
(2) For Windows PCs, to we also have licenses to run PVX with a USB camera and head set microphone. This does exactly the same as Polycom but for one user only.
(3) All Mac users with OS X (Tiger) are equipped with iSight cameras that operate the new H.264 video codec, also known as MPEG-4 Part 10, as part of the iChat AV software. This allows up to four simultaneous users to hold a conference from any location in the world that provides wireless or fixed-cable access to the Internet. These facilities satisfy all the needs of the particle physics group at modest capital cost and have provided an extremely good project meeting infrastructure for this work.

4. PATIENT RECORD DATABASE AND METHODS

4.1. Data description

At the outset, it was anticipated that the bulk of the database of 58 220 patients records would be available and useable for the study. Approval was obtained from the Privacy Advisory Committee to use the database (Scottish Colorectal Cancer Audit Database) that contained information collected by the Scottish Cancer Therapy Network and Clinical Resource and Audit Group. This was obtained from the Information and Statistics Division of the NHS Scotland. The data used for the study are shown in Table I, which shows about a factor of 20 reduction compared to that anticipated. The reasons for this reduction in useful data were mainly due to incompleteness of the record with information recorded as 'unknown' or 'missing'. This can be rescued in some cases by inserting the average known value of a missing parameter and checking that the ANN makes more reliable predictions in known cases if this is done.

For the analysis, the significant event was defined to be a cancer-related death occurring within five years of clinical follow-up, beginning from the date of first diagnosis. Patients who fell in this category were designated as 'non-censored'. The rest were defined to be 'censored', i.e. a patient diagnosed before the start date of the dataset or one who died after five years was defined to be censored. Table II shows the 16 categorical variables that were selected and used for the ANN training. These are the input parameters to the ANN. The 'Age Group' variable has six attributes (less than 50, 50–59, 60–69, 70–79, 80–89 and over 90). Unfortunately, the database was unsatisfactory in the majority of cases, although it was satisfactory in a sufficient minority of cases. Hence, we decided to keep only those records with known values consisting of about 2778 records, of which 1558 were used for training. The most important variables were adjudicated by those clinicians involved in the study. These were: age; Duke's stage; number of positive nodes; vascular pedicle node identified; chest X-ray; liver US/CT scan; and laparoscopy. For the rest of the set, any missing values were set to the mean of the particular variable.

Regular video-conference discussions between scientists and clinicians were held throughout the development, training and implementation of the ANN, in order to establish the origin of the deficiencies and to address them. In those cases where important parameters were missing, there was little option but to discard the record. The only way to retrieve anything would have been

DOI: 10.1002/cpe

Table I. The statistics of useful records from the Scottish colorectal cancer audit databases (obtained from the Information and Statistics Division of the NHS in Scotland) used in the study.

	Sex	
Age group (years)	Male	Female
Less than 50	85	56
50–59	190	166
60–69	423	332
70–79	467	438
80–89	204	336
Over 90	19	62
Total	1388	1390

Table II. Variables recorded on the database and used for training the ANN.

N	Variable	Attributes
1	Age group	6
2	Duke's stage	A, B, C, D
3	Number of positive lymph nodes	1–10
4	Vascular pedicle node involvement	2
5	Chest X-ray	2
6	Liver abdomen US/CT	2
7	Laparoscopy	2
8	Operation intent	3
9	Weight loss	3
10	Radiotherapy	3
11	Chemotherapy	3
12	Tumour size	3
13	Tumour differentiation	2
14	Site group	4
15	Anastomotic leak	2
16	Clinical trial	2

DOI: 10.1002/cpe

to go back to the source. There are two virtually insurmountable obstacles in doing this. When the data are presented for analysis, they are anonymized. Re-identifying the source would have violated the terms on which the database guardians gave consent for the use. The second problem was one of the resources needed for retrieval. If, for example, five person-minutes were spent on each such faulty record, this would correspond to almost three years full-time work just to re-source the material. In most cases, it was our opinion that the search would reveal that the data were irretrievably faulty and so this huge resource, even if available, would have been wasted.

The data reduction experienced by particle physicists in their current massive data flow experiments corresponds to a reduction of a potential annual data store of a million petabytes down to a single petabyte of actual stored data. Most of the discarded data is background and noise. The medical counterpart here, 58 220 reduced by a factor 20 seems quite modest. However, each record of medical data is potentially useful and is only discarded reluctantly. Even so, the integrity of the analysis demanded that over 90% of patient record data be discarded for this analysis and this is a pointer for future data recording.

4.2. Modelling survival

We modelled the survival of patients using a partial likelihood ANN [8–10]. The base element of the ANN was incorporated into a multi-layer perceptron (MLP) with a sigmoid activation function. The input layer has units for all of the covariates (those in Table II) plus one bias unit and an additional unit for time. A single output unit denotes the conditional probability of failure (i.e. death). One of the advantages of such a model is the ease of incorporating time-dependent covariates, since each subject is represented, for each interval, by one input sector that can change across time intervals. The choice of this network configuration ensured maximal predictive capability. Only a brief outline of the procedure that we used is given here, since the details have little relevance to the e-Science. The ANN procedure itself is routine; it is the nature of the data that introduces challenges.

We also addressed problems associated with censoring and classification. The method used was found to have only a small bias due to censoring and it allowed the use of covariates that change dynamically with time. Our scheme classifies patients within each time period into either ‘alive’ or ‘dead’, provided that the data in the training set are well balanced and the distribution of both classes is uniform. However, we found a tendency for the lowest populated class at a specific time interval to be ignored; e.g. the class ‘dead’ at the beginning of the follow-up history and class ‘alive’ at the end of follow-up are both sparsely filled. This results in certain biases in the final classification, which are addressed by weighting the members of this class as in [11,12].

The process of choosing the optimal set of variables was based on the combination of the opinion of clinicians and a five-fold cross-validation procedure [10,13]. As a result, a total of 1558 records of patients referred between 1993 and 1998 with a follow-up of 60 months was chosen. The ANN was trained using the combination of simple regularization and the ‘early stopping’ technique [13].

The output prediction initially showed signs of instability due to the variance of the data used for training and the initial values of the weight matrix that were randomly generated. This instability was addressed by building an ensemble of ANNs and aggregating the results of the networks in the ensemble to produce reliable predictions via a bootstrap procedure: 200 separate ANNs were trained on randomly sampled subsets from the whole set of records for each ANN and the resulting survival curves were averaged at every point of the follow-up.

DOI: 10.1002/cpe

Finally the model was also calibrated against real survival, by comparing it with Kaplan–Meier (KM) [14] survival curves. This comparison is presented in future publications.

5. RESULTS

5.1. Prediction of colorectal cancer-related survival

The detailed discussion of the clinical results of this research will be presented in future publications. Here we are concentrating mainly on the e-Science factors and data quality. A critical data output is the so-called 'receiver operating characteristic' (ROC) curve [15]. This curve was first derived to quantify signal to noise identification in the reception of radio signals. In this application, it indicates the probability of a correct prediction (*POC*) versus the probability of a false positive (*POF*) for our data. In the case of radio reception, the situation was the probability of identifying a real signal against the probability of misidentifying noise as a signal. If it is accepted that all the available data have been used, that the data are unbiased and that the ANN has made the best decision, given this particular input data, then this curve represents the irreducible best case for decision making. In other words, there is no other input available that will improve this result.

The best case scenario corresponds to the case when the curve starts at the origin, moves vertically up the ordinate to the point ($POC = 1$, $POF = 0$) and then moves parallel to the abscissa to the point ($POC = 1$, $POF = 1$). In this case, the two classes, 'survival' and 'non-survival' are completely separated and identifiable. The worst case is given by the line $POC = POF$ where the prediction is always an indecisive 50:50. The reality in health care, especially here, is that one can never be completely sure either way and this curve quantifies the probabilities. There is of course the possibility that an analysis could give the perverse result $POC < POF$, in which case one would invert the final decision and search for the flaw in the analysis. For the best case, the average value of *POC*, $\langle POC \rangle = 1$, whereas the worst case has $\langle POC \rangle = 0.5$. The average value of *POC* is simply the area under the curve since both axes have a maximum value of 1. More generally, as shown by Hanley and McNeil [16] in radiological applications, the area under the ROC curve is, in our case, specifically a measure of the probability that the recorded characteristics of the patients in the database will allow a correct prediction.

Our ROC curves for survival to 10, 30 and 60 months respectively are shown in Figure 1. The three curves are essentially indistinguishable. For our analysis, the area under the curves, equal to the average probability, was found to have an average value of $\langle POC \rangle = 0.80 \pm 0.01$. This (standard) error was calculated using the method of Hanley and McNeil. Thus, there is an 80% probability that the database and the ANN will lead to a correct prediction.

The ANN, having been trained on a sub-set of patient records, was then used to predict the survival for patients whose actual survival was withheld from the ANN. Excellent agreement was found.

5.2. Safe computer access to the system

Another important e-Science element arose during the evaluation stage of the outputs from the neural network. It was important to allow the clinicians access to the software without them having to learn how to set up and submit the jobs. We therefore designed a secure Web-based interface between the

DOI: 10.1002/cpe

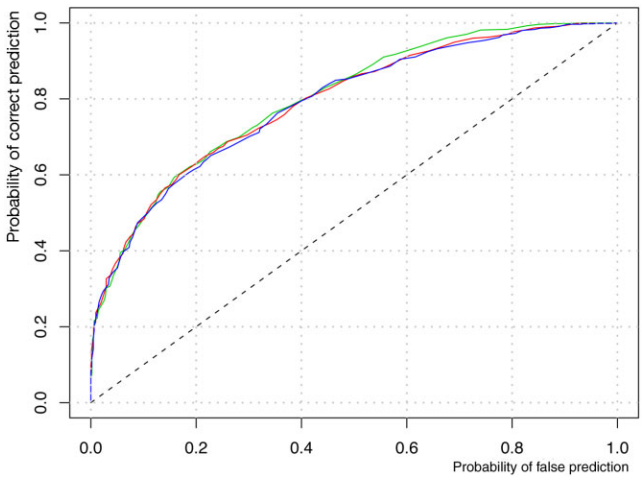


Figure 1. The ROC curves for various age groups. The curves are essentially identical and no attempt is made to distinguish them here. All curves were averaged to obtain the probability corresponding to the area under the curve.

clinicians and the associated kernel software. This will allow the eventual use of Grid technology for data collection from remote databases, training or re-training the ANN using distributed computer resources and finally, the provision of Web access interface for *bone fide* users.

A user logs into the system with secure ID and password. The security level at present is appropriate for anonymized data but would need strengthening when the system is extended to include manual intervention on the data by multi-users. The user is presented with an array of input boxes into which the specific data for a patient, new or existing, can be entered. It is not necessary to input all fields; indeed, some are still irrelevant but have been created for future exploitation. The user can select whether or not to retain previous results for comparison and then submits the request to the central processing node, which is currently on a server in Manchester. The clinician in Dundee then clicked the data field to Manchester where it was used as input to a specific run by the software. The ANN produced a survival curve corresponding to that particular request and transmitted it back to Dundee within the framework of the Web interface. In a system operating on truly randomized data, the clinician could now alter the management profile of the disease and study the effect of this change on the survival curve. Key factors here are the various types of surgical intervention: palliative, potentially curative and curative. The type of chemotherapy is also expected to be critical. One can visualize a scenario in clinical trials where new drugs are being used. If the data are fed to the ANN for ongoing training, such a network will provide instantaneous feedback on the trial while it is still in progress.

Figure 2 shows an example screen shot of the welcoming interface with known parameters for a hypothetical patient entered. The predictive survival probability for 60 months for this particular case is calculated by the ANN to be $P_{60} = 0.535$. If the operative plan is changed from 'potentially curative' to 'none curative resection' and then 'palliative', the probabilities become $P_{60} = 0.255$ and

DOI: 10.1002/cpe

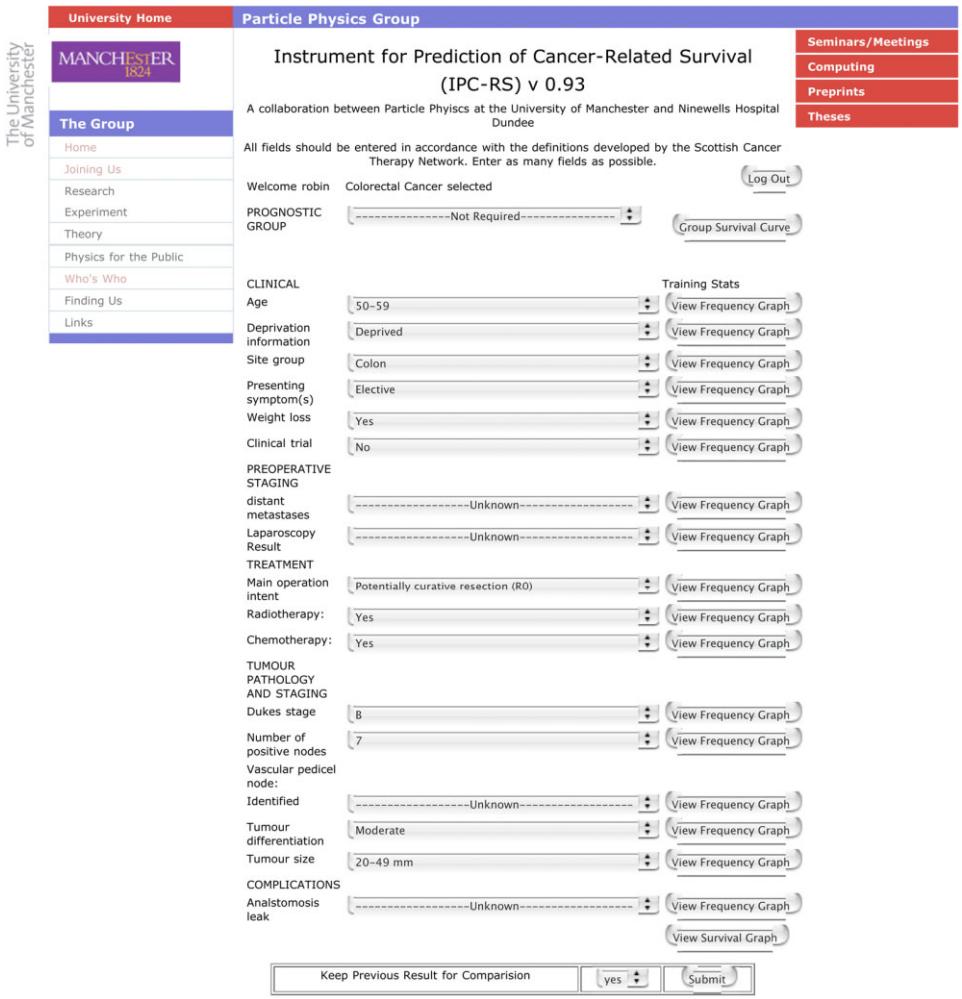


Figure 2. A screen shot of the user graphical Web interface for system access by clinicians.

$P_{60} = 0.242$, respectively. This indicated the ability of the ANN to make a quantitative prognoses of the effect of various types of surgery. Other parameters such as radiotherapy and chemotherapy can also be changed. Therefore, the clinicians will eventually be able to compare the curves for completely different management profiles.

Figure 3 shows two survival curves for two people with identical parameters except for age. The younger patient has a significantly better prognosis. The histogram is the distribution of real data (KM curves) and the error bars are statistical. The curve is the ANN prediction for an individual whose

DOI: 10.1002/cpe

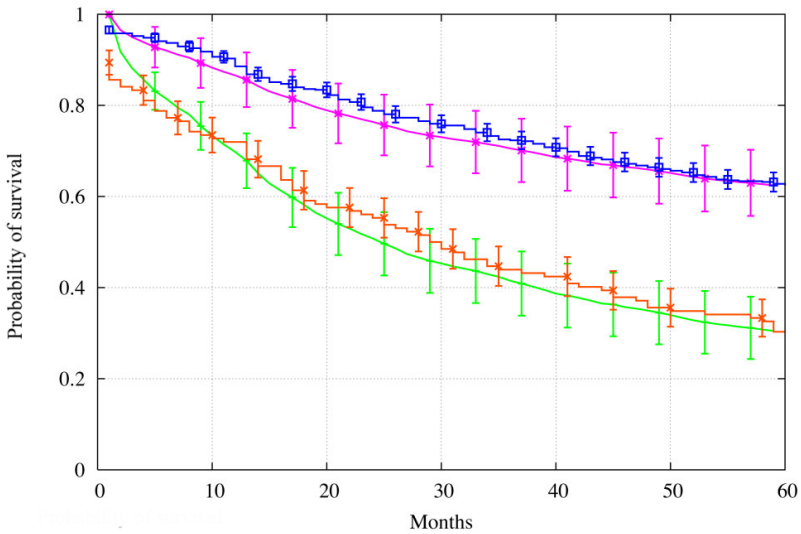


Figure 3. Typical survival curves returned by the ANN. For these two curves, both patients had identical parameters except for age. The upper curve is for a younger person, the lower curve for an older one.

personal parameters lie in the middle of the range of the population in this age and disease severity group. Here the error bars represent plus and minus the inner quartile. The use of these curves in general demography as well as individual patient management is clear from this graph.

The Web part of the system has been written using Perl and common gateway interface (CGI) scripts and uses 'GridSite' server technology [17], produced by the Manchester particle physics group. GridSite was originally a Web application developed for managing and formatting the content of the Web site for the Particle Physics Grid (GridPP). Since 2001, it has grown into set of extensions to the Apache Web server and a toolkit for Grid credentials, access control lists and http(s) protocol operations.

The full implementation on the Grid will enable access and linkage with remote other current and future databases anywhere. The kernel and interactive software was originally written using Linux based PCs. The full system—Web server interface and kernel software—is now placed on a 14 node, 14 TB Apple XServe system that can be accessed from anywhere in the world.

6. DISCUSSION AND CONCLUSIONS

The ANN model that has been implemented and trained works well in most cases, as demonstrated by the good agreement with actual survival of the patient cohort included in the study. In addition, we are able to account for instances where the initial output predictions obtained were not in perfect agreement with actual survival of the patients. The root cause of these problems is the inconsistency

DOI: 10.1002/cpe

and incompleteness of the data used to train the ANN, resulting in unduly large variance. Unfortunately, this problem is inherent to all existing clinical databases collected centrally over many years within the NHS system. If ANNs are to be used more extensively in the future, as suggested by the emergence of the Grid technology and associated middleware, this problem has to be resolved and the quality of data collection entry and management improved significantly beyond the existing level. In any modelling system, however advanced and sophisticated, the output cannot be more accurate that the data used for training. In completing the present work, we have come to the conclusion that the ideal databases for the development of high-level ANNs for complex clinical tasks are from completed randomized clinical trials. The present situation undoubtedly disadvantages any progress in ANN modelling and development and possibly explains why these and other mathematical models are not widely used in practice by clinicians [18]. At one end of the scale, they have a possibly unfair reputation for unreliability and are regarded with suspicion by some clinicians. At the other end of the spectrum of opinion, some clinical researchers believe that even ANNs are already old fashioned and should be superceded by more sophisticated models of artificial intelligence. Clearly, the ANNs themselves have to established as a solid procedure accepted by all before the more advanced frontiers of artificial intelligence can be brought into clinical practice.

The present stage in the ANN project can be regarded as a base platform, which we anticipate will develop into a system that will reliably stage, prognosticate and eventually outline an algorithm for the management of individual patients with cancer. So far, despite the problems we have encountered with the quality of the data set, we have produced a reliable, practical and easy to use system. As it is developed further, it should be able to meet the requirements for the other objective—the optimal management of cancer patients commensurate with the stage of their disease. Randomized data are needed for this stage.

We also believe that ANNs may also be developed and used for a variety of chronic disorders, e.g. cardiovascular disease, diabetes, etc., with the benefit of standardization that such systems bring to the treatment and management of patients. The project was carried out by two geographically separated teams with diverse sets of skills, working together using a few e-Science tools. The scope of the approach has been demonstrated and successful results obtained.

ACKNOWLEDGEMENTS

This work is supported by an e-Science grant from the MRC and the Particle Physics and Astronomy Research Council (PPARC).

REFERENCES

1. http://www.pparc.ac.uk/Rs/Fs/Es/e_science_first_call.asp [5 December 2005].
2. http://www.nesc.ac.uk/nesc/define.html [5 December 2005].
3. http://www.nerc.ac.uk/funding/escience/ESci_AO2_Framework.shtml [5 December 2005].
4. Dybowski R. Neural computation in medicine: Perspectives and prospects. *Proceedings of the ANNIMAB-1 Conference (Artificial Neural Networks in Medicine and Biology)*, May 2000, Malmgren H, Borga M, Niklasson L (eds.). Springer: Berlin, 2000; 26–31.
5. Lisboa PJG. A review of evidence of health benefit from artificial neural networks in medical intervention. *Neural Networks* 2002; **15**(1):11–39.
6. Baxt WG. Application of artificial neural networks to clinical medicine. *Lancet* 1995; **346**:1135–1142.

DOI: 10.1002/cpe

7. Naguib RN *et al*. Artificial neural networks in cancer research. *Pathobiology* 1997; **65**(3):129–157.
8. Biganzoli E *et al*. Feed forward neural networks for the analysis of censored survival data: A partial logistic regression approach. *Statistics and Medicine* 1998; **17**:1169–1174.
9. Lisboa PJG, Wong H, Harris P, Swindell R. A Bayesian neural network approach for modelling censored data with an application to prognosis after surgery for breast cancer. *Artificial Intelligence in Medicine* 2003; **28**:1–5.
10. Ripley BD, Ripley RM. Neural networks as statistical methods in survival analysis. *Clinical Applications of Artificial Neural Networks*. Cambridge University Press: New York, 2001.
11. Lowe D, Webb AR. Exploiting prior knowledge in network optimization: An illustration from medical prognosis. *Network* 1990; **1**:299–321.
12. Lisboa PJG, Vellido A, Wong H. Bias reduction in skewed binary classification with Bayesian neural networks. *Neural Networks* 2000; **13**:407–416.
13. Tarassenko L. *A Guide to Neural Computing Applications*. Wiley: London, 1998.
14. Collet D. *Modelling Survival Data in Medical Research*. Chapman and Hall: London, 1994.
15. Egan JP. *Signal Detection Theory and ROC Analysis*. Academic Press: New York, 1975.
16. Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology* 1982; **143**:29–64.
17. Gridsite. http://www.gridsite.org [5 December 2005].
18. Schwarzer G, Vach W, Schumacher WM. On the misuses of artificial neural networks for prognostic and diagnostic classification in oncology. *Statistics and Medicine* 2001; **19**:541–561.
19. Bishop CM. *Neural Networks for Pattern Recognition*. Clarendon Press: Oxford, 1995.

DOI: 10.1002/cpe