Artificial Neural Network to Predict Skeletal Metastasis in Patients with Prostate Cancer

Jainn-Shiun Chiu · Yuh-Feng Wang · Yu-Cheih Su · Ling-Huei Wei · Jian-Guo Liao · Yu-Chuan Li

Received: 31 January 2008 / Accepted: 23 April 2008

Abstract The application of an artificial neural network (ANN) in prediction of outcomes using clinical data is being increasingly used. The aim of this study was to assess whether an ANN model is a useful tool for predicting skeletal metastasis in patients with prostate cancer. Consecutive patients with prostate cancer who underwent the technetium-99m methylene diphosphate (Tc-99m MDP) whole body bone scintigraphies were retrospectively analyzed between 2001 and 2005. The predictors were the patient's age and radioimmunometric serum PSA concentration. The outcome variable was dichotomous, either skeletal metastasis or non-skeletal metastasis, based on the results of Tc-99m MDP whole body bone scintigraphy. To assess the performance for classification model in clinical study, the discrimination and calibration of an ANN model was calculated. The enrolled subjects consisted of 111 consecutive male patients aged 72.41±7.69 years with prostate cancer. Sixty-seven patients (60.4%) had skeletal metastasis based on the scintigraphic diagnosis. The final best architecture of neural network model was four-layered perceptrons. The area under the receiver-operating characteristics curve (0.88±0.07) revealed excellent discriminatory power ($p<0.001$) with the best simultaneous sensitivity (87.5%) and specificity (83.3%). The Hosmer–Lemeshow statistic was 6.74 ($p=0.08>0.05$), which represented a good-fit calibration. These results suggest that an ANN, which is based on limited clinical parameters, appears to be a promising method in forecasting of the skeletal metastasis in patients with prostate cancer.

Keywords Artificial intelligence · Computer assisted · Image interpretation · Radionuclide imaging · Prostatic neoplasm · Bone metastasis

J.-S. Chiu · Y.-F. Wang · L.-H. Wei · J.-G. Liao
Department of Nuclear Medicine,
Buddhist Dalin Tzu Chi General Hospital,
Chiayi, Taiwan

Y.-C. Su
Division of Hematology and Oncology, Department of Internal Medicine, Buddhist Dalin Tzu Chi General Hospital,
Chiayi, Taiwan

Y.-C. Li (✉)
Institute of Biomedical Informatics,
National Yang Ming University,
No. 155, Sec. 2, Linong St., Beitou District,
Taipei City 112, Taiwan
e-mail: jack.li@m2k.com.tw

J.-S. Chiu · Y.-F. Wang · Y.-C. Su
Department of Medicine, College of Medicine,
Tzu Chi University,
Hualien, Taiwan

Introduction

Skeletal metastasis in patients with prostate cancer causes considerable morbidity including severe bone pain, impaired mobility, hypercalcemia, leukopenia, pathological fracture, spinal cord or nerve root compression, and bone marrow infiltration [1]. Beside the significantly decreased quality of life from these complications, almost 50% of prostate cancer patients with metastatic bone disease die within 30 months [2]. In patients with prostate cancer, detection of osseous metastasis is important to selecting the best treatment and stratifying the prognosis [3]. Although bone histomorphometry is the gold standard for diagnosing metastatic bone disease, bone biopsy is a painfully invasive procedure. Moreover, if the patient is suspected to have multiple bone metastases, it is impossible to carry out several bone biopsies. In clinical practice, currently bone

scintigraphy is the modality for evaluating skeletal metastasis of prostate cancer. Bone scintigraphy not only has a high sensitivity derived benefit from the intrinsic nature of injected radiopharmaceutical, but also is a one-step procedure for whole body skeletal survey. However, the equipment for radionuclide imaging is not universally available due to concurrent requirement of the expensive and complex gamma scintillation camera system and a specialized team such as well-trained nuclear medicine physicians, radiological technologists, and professional nurses.

Artificial neural network (ANN) is an excellent presentation of artificial intelligence that is patterned with various computational algorithms after the structure of the human nervous system [4]. Every processing element in an ANN, usually called 'artificial neuron', is interlinked several weighted signals that emulate the human synaptic connections used to memorize, learn, and predict the subsequent interactions. An ANN allows identification of underlying associations among predictors and outcomes that may not be detected with classical statistical analyses [5]. Moreover, an ANN can improve its predictive ability through iterative learning algorithms. By virtue of these inherent advantages, ANN has been increasingly deployed as a forecasting method in clinical medicine and urology [6, 7]. In the specialty of urologic oncology, ANN has been applied successfully to the early diagnosis, screening, staging, and progression in prostate cancer [8]. However, no ANN model has ever been designed solely for the prediction of skeletal metastasis in prostate cancer patients. As a widely applied implementation of artificial intelligence in medicine, ANN holds promise as a tool to predict bony metastasis. Herein, our study is the first investigation to develop an ANN model to predict skeletal metastasis in patients with prostate cancer. We have validated its feasibility in comparison with whole-body bone scintigraphy as a reference method.

Materials and methods

Patients' enrollment

Between June 2001 and November 2005, consecutive patients with prostate cancer who underwent the technetium-99m methylene diphosphate (Tc-99m MDP) whole body bone scintigraphy were retrospectively analyzed. Their data were retrieved from respective records in our scintigraphic database at a regional teaching hospital. Patients whose radioimmunometric prostate-specific antigen (PSA) data could not be available within one month before and after the Tc-99m MDP whole body bone scintigraphy were excluded from this study. The study was conducted according to the guidelines of the Declaration of Helsinki and the Ethics Committee on Human Studies at our hospital (Buddhist Dalin Tzu Chi General Hospital, Chiayi County, Taiwan) approved the study; informed consent for enrolled patients was not required for clinical data collection from medical records according to the protocol of our institutional review board. To preserve patient confidentiality, direct patient identifiers were not collected. Data were reported only in aggregate form.

Image acquisition and biochemical data

All enrolled patients with prostate cancer received Tc-99m MDP whole body bone scintigraphy. Tc-99m MDP was commercially available and provided by Daiichi Radioisotope Labs, Ltd. (Tokyo, Japan). Following the Procedure Guideline for Bone Scintigraphy (version 3.0, approved June 20, 2003) announced by the Society of Nuclear Medicine, a whole body bone scintigraphy (from the toes to top of the head) was performed using whole-body moving camera technique (anterior and posterior) three to four hours after intravenous injection of 20 to 25 mCi Tc-99m MDP. All images were acquired using dual-head gamma camera scintillation system (DST-XL, General Electric Medical Systems, Buc, France) equipped with large field-of-view, low-energy, high-resolution collimators with a 20% energy window centered at 140 KeV. The scan speed was 16 cm/min and the matrix size was 512×2,048 pixels. One independent physician trained in nuclear medicine with 15 years experience interpreted all scintigraphic images blind to serum PSA levels and any available clinical findings. The reading and interpretation were used without reanalysis of images. The final results were elucidated as negative (Fig. 1) or positive (Fig. 2) for skeletal metastasis, according to the bone absorption of the radiopharmaceutical. Serum PSA concentrations of all registered patients were measured by radioimmunometric assay (Immunotech-PSA total IRMA kit; A Beckman Coulter Company; Fullerton, CA, USA), considering as normal values between 0 and 4 ng/ml. All patients were investigated in the same department by either scintigraphic images or radioimmunometric assay.

ANN construction

No well-established theoretic protocol exists for the determination of an ideal ANN configuration including numbers of hidden layers, numbers of neurons in each hidden layer, the optimal number of iterations, or activation functions [9]. At the beginning, the designer creates the structure and the best practice is typically established upon trial and error [4, 10]. We used Statistica 7.0 (StatSoft, Inc., Tulsa, OK, USA) to generate various formulations of ANN models. The patient's age of and serum PSA concentration

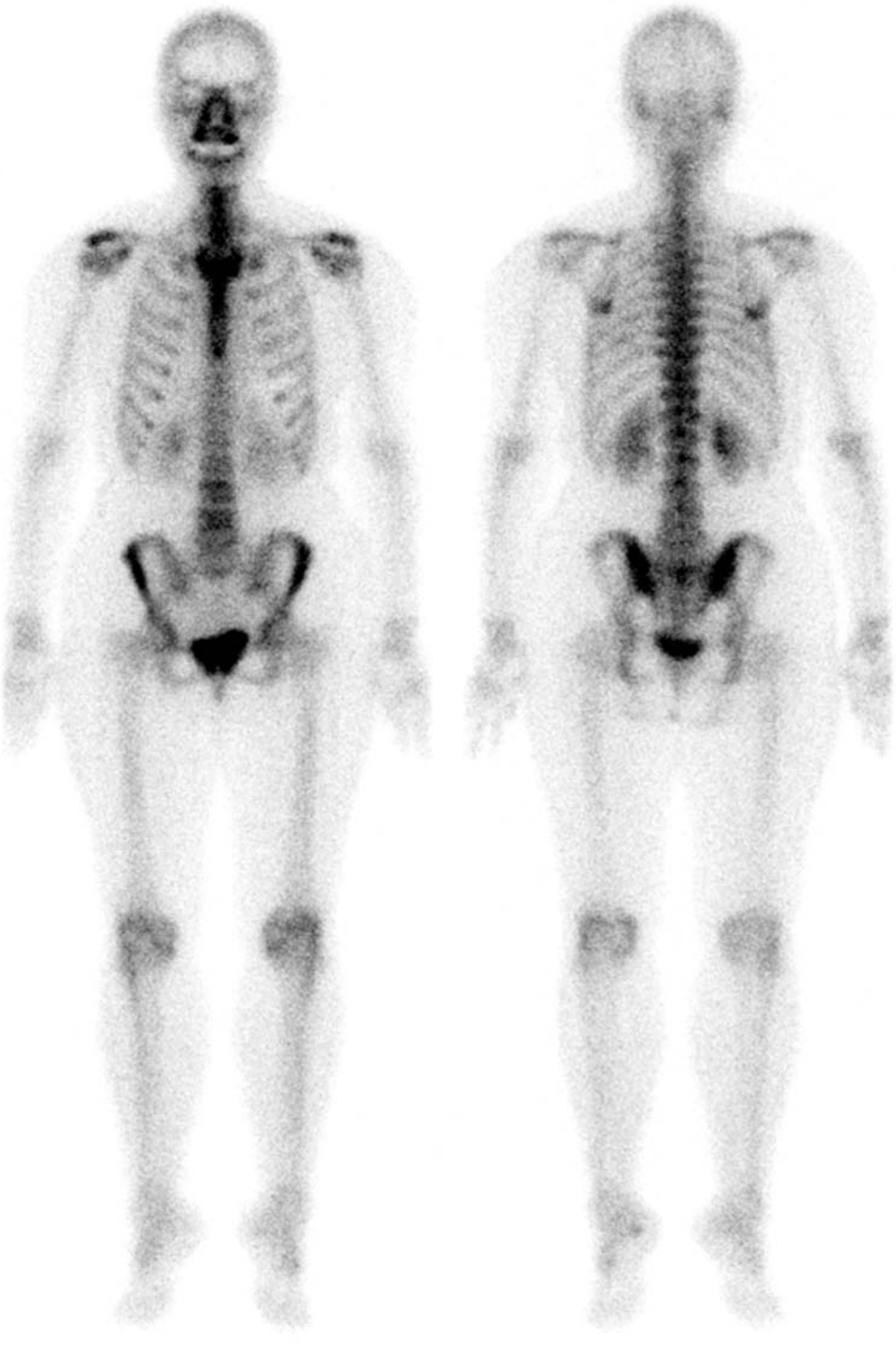

Fig. 1 Tc-99m MDP whole body bone scintigraphy revealed non-skeletal metastasis in a patient with prostate cancer

were entered as continuous input variables into ANN models. The outcome variable was dichotomous, either skeletal metastasis or non-skeletal metastasis, using the results of Tc-99m MDP whole body bone scintigraphy as interpreted by the nuclear medicine physician.

To treat the statistical problem resulting from the limited number of patients in our study, bootstrap resampling was done to mitigate the possible bias initiated by depending on any one particular part into train or test subsets [11]. This technique samples a data set with replacement, meaning a single case may be randomly sampled several times into the bootstrap set. The bootstrap can be enforced any number of times to raise accuracy. Compared with random sampling, the use of sampling with replacement can minimize the gen-

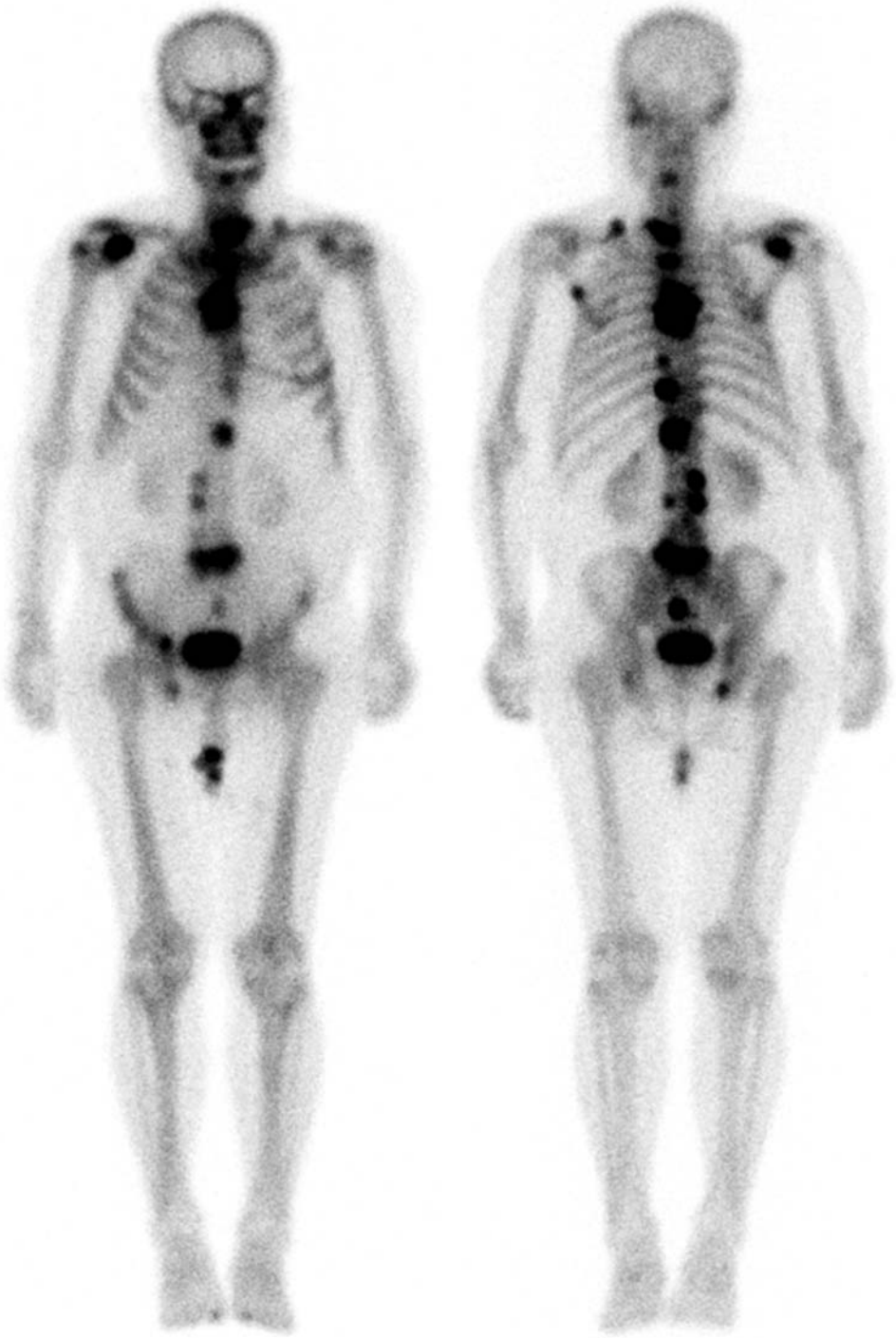

Fig. 2 Tc-99m MDP whole body bone scintigraphy displayed multiple bone metastases in a patient with prostate cancer

eralization problems caused by the dataset's finite size. We sampled the selection subset first without bootstrapping; subsequently, the training subset was bootstrapped from the remaining data. Although using a training subset optimizes a neural network, a selection subset is separately used to stop training to lessen overfitting and overtraining. Afterwards, a third subset known as a test subset is utilized to execute an unbiased estimation of the network's probable performance. Hence, we set the selection subset size as one-third of enrolled patients and the remaining cases to be bootstrapped into the training subset. The test subset is created from any cases left over after the bootstrap selection of the training subset.

During the processing protocol the intelligent problem solver instructed a large number of trials, which were used to settle the best architecture [12–14]. It could permit

concurrent comparison of different networks (linear network, three- and four-layer multilayered perceptron networks, radial basis function network, probabilistic, and generalized regression neural networks) using a combination of heuristic and optimal algorithms to choose the smoothing factor and the number of processing units for these networks [15]. For all types of networks, we set up the number of hidden units as one for a minimum and 14 for a maximum. To compare the performance of networks, the intelligent problem solver balanced error against type and diversity as criteria for selecting retained networks, in which case it preserved networks with a range of types and performance/complexity trade-offs. If the network file is plentiful and the new model is subordinate to the nominee for replacement, the network set will be intensified in utmost size to be suitable to the new networks. In a network that has not overfitted the data, the training subset error will generally be a good representation of test subset. After the network was permitted to run and a forecast was made, the predicted outcome was correlated with the observed outcome. If the network predicted the outcome incorrectly, by a process of back propagation, hidden weights within the network were readjusted until the predicted outcome was exact. Finally, the intelligent problem solver employed the best architecture of the network and the optimal set of input variables.

Statistical analyses

Statistical analyses were performed using MedCalc for Windows, version 9.1 (MedCalc Software, Mariakerke, Belgium) and expressed as mean ± standard error. The Mann–Whitney test was used to compare the differences of input variables (patient age of patients, serum PSA concentrations) between skeletal and non-skeletal metastatic groups. The statistically significant level for comparisons between two groups was defined as a p value less than 0.05. To assess the quality of a classification model in clinical investigation, discrimination and calibration should be assessed concurrently [16]. Discrimination is a measure of how well a model recognizes subjects correctly as two different classes; calibration, on the other hand, evaluates the degree of correspondence between the estimated probabilities produced by a model and the actual observation. The area under the receiver operating characteristic curve (AUC) with best sensitivity and specificity simultaneously were used as indicators to appraise the discriminatory power of an ANN model for prediction of skeletal metastasis in patients with prostate cancer [17, 18]. An AUC of 1.0 infers perfect discrimination, whereas an AUC of 0.5 is equivalent to a random model. An AUC between 0.7 and 0.8 was classified as "acceptable" and between 0.8 and 1.0 as "excellent" discrimination [19]. On the other hand, calibration was assessed using the Hosmer-Lemeshow goodness-of-fit statistic (H statistic) which divides subjects into deciles based on predicted probabilities and then computes a chi-square from observed and expected frequencies [20–22]. A statistically good fitness between a new model and the reference method is defined as p value more than 0.05.

Results

A total of 111 males (age: mean, 72.41±0.73 years; median, 73 years; range, 51–89 years) with prostate cancer were finally enrolled and their mean serum PSA concentration was 814.51±327.07 ng/ml (median, 46.20 ng/ml; range, 0.10–24567.00 ng/ml). Table 1 lists the various characteristics of patients between skeletal and non-skeletal metastatic groups. A comparison between two groups showed that patients with skeletal metastasis were older, a statistically significant finding. Patients with skeletal metastasis also had statistically significantly higher serum PSA concentrations than did the patients without skeletal metastasis.

Figure 3 portrays the schematic diagram of the final best ANN model, which was created to predict skeletal metastasis in patients with prostate cancer. The outline shows the number of artificial neurons in each of the four layers (one input layer of two neurons, first hidden layer of nine neurons, second hidden layer of six neurons, one output layer with one neuron), and it demonstrates that the network was fully connected in that each artificial neuron in a given layer was linked to every artificial neuron in the nearby layer. Two input variables were all adopted as significant features after training processes. The fact that the network was entirely interconnected meant that 78 weights had to be modified following the processing of each record in the training period. The ANN output is a single continuous variable with a scope of 0 to 1. A threshold in this interval was used above which all values were regarded as consistent with the skeletal metastasis in patients with prostate cancer. By varying this threshold, the discriminatory power of an ANN model for test subset is

Table 1 Characteristics of patients with skeletal and non-skeletal metastatic groups

	Skeletal metastasis (n=67)	Non-skeletal metastasis (n=44)	p value
Age (years)	73.90±0.97	70.14±1.03	0.003
PSA (ng/ml)	1334.27±533.87	23.05±9.47	<0.001

Data are expressed as mean ± standard error. The statistical p values were derived from the Mann–Whitney test.
Age age of the patient, *PSA* serum prostate specific antigen concentration

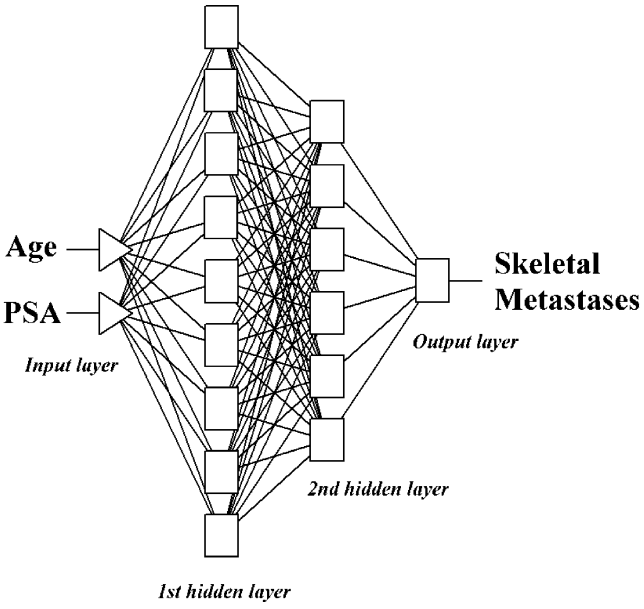

Fig. 3 Graphical representation of our multilayer perceptron artificial neural network model

depicted by the ROC curve (Fig. 4). The ROC analysis for the ANN model gave an AUC of 0.88±0.07, indicating that the ANN model had good diagnostic efficiency and represented a significantly excellent discriminatory power ($p<0.001$). The best simultaneous sensitivity and specificity were 87.5% and 83.3%, respectively, based on the threshold value of more than 0.64 for discriminating cases with skeletal metastasis from non-cases. On the other hand, the *H* statistic value of ANN model was 6.74 ($p=0.08$). The ANN model had a statistically good fit represented by statistically insignificant *H* statistics ($p>0.50$), which is a sensitive measure of differences between predicted and observed outcomes. These results suggested that the final best ANN model had excellent discriminatory power and good-fit calibration concurrently.

Figure 5 shows the three-dimensional plot indicating the relationship between input variables and output score for prostate cancer patients. In the plot, 1 is the score for the patients with prostate cancer who has the skeletal metastasis and 0 is the score for the patients with prostate cancer who does not have the skeletal metastasis. Patients with prostate cancer in the region with a score more than 0.64, derived from the best threshold value of the ROC analysis, belonged to a group with skeletal metastasis.

Discussion

In patients with prostate cancer, several forecasting methods were developed to aid clinicians in cancer staging, including the most widely accepted nomograms [23, 24] and state of the art ANN models [25–29]. Most of these investigations focused on the analysis in predicting positive lymph node involvement. No such predictive models exist to predict the skeletal metastasis for patients with prostate cancer. In this study, we not only found that the patients with skeletal metastasis were older and had higher serum PSA concentration but also we successfully used ANN-based modeling techniques in an attempt to individually predict the occurrence of skeletal metastasis in these patients.

To implement the application of artificial intelligence in future daily practice, the initial motivation for our ANN-

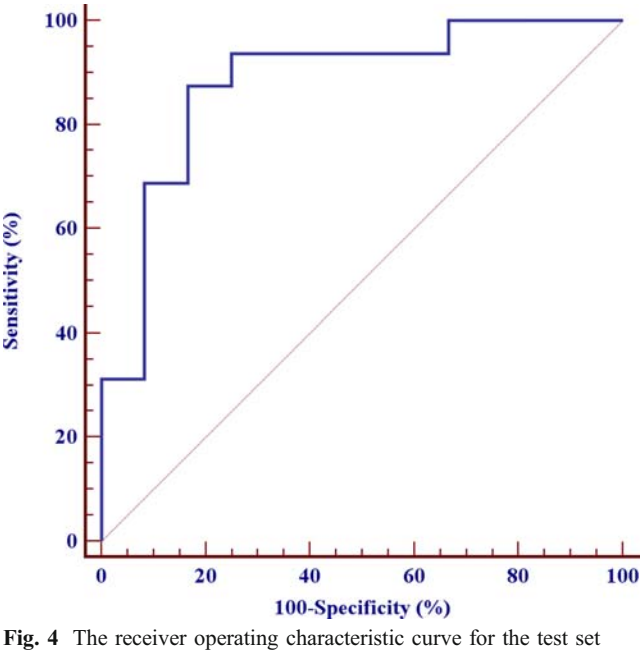

Fig. 4 The receiver operating characteristic curve for the test set

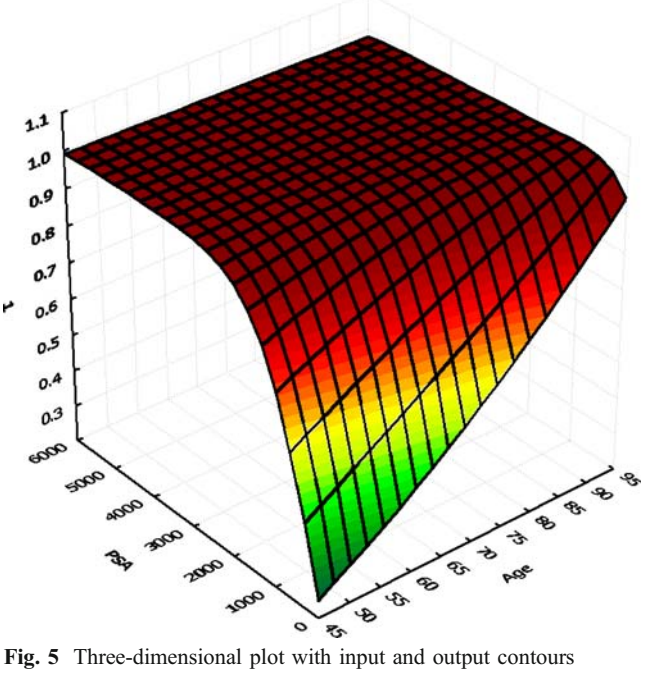

Fig. 5 Three-dimensional plot with input and output contours

based model is to exploit minimum features derived from clinical parameters; thus, imaging characteristics were not used. For support of any clinical decision to be effectively feasible, the predictors should be routinely available in the clinician's workflow and easily attainable at the point of care [30]. This is pertinent since the utilization of fewer features would permit acquisition and analysis of less data, simplifying the clinicians' work. By virtue of these commandments, we only selected two variables, age of the patient and serum PSA concentration, which are usually chosen as predictors in many ANN-based staging methods [28, 29]. Age is a basic patient demographic found in the medical record and easily calculated when the patient visits the hospital. In one study conducted by Carter and colleagues, their results showed that increasing age is positively associated with a higher probability of non-organ-confined prostate cancer [31]. Although our study had similar findings, we clearly displayed the statistical difference of age between skeletal and non-skeletal metastasis in patients with prostate cancer.

Serum PSA concentration is currently the most widely used tumor marker to diagnose and monitor prostate cancer, even though serum PSA level is not 100% specific and it may be affected by several confounding elements, including urologic manipulation, tumor volume, tumor differentiation, ejaculation, chronic prostatitis, or benign prostate hyperplasia. Recently, several biochemical markers, such as osteoprotegerin and tartrate-resistant acid phosphatase isoenzyme 5b, show promise in predicting skeletal metastasis in patients with prostate cancer [32]. Another pilot ANN study with novel inputs of macrophage inhibitory cytokine 1, human kallikrein 11, migration inhibitor factor, and prostate volume illustrated significant superiority compared with percent free PSA and total PSA to enhance the detection of prostate cancer [33]. While these biochemical markers might appear promise, they are not clinically available, which may impact the algorithms' applicability. Therefore, these biomarkers are not suitable for our ANN inputs. In our study, we were determined to utilize limited clinical variables to construct the ANN topology and we achieved good predictability in terms of either discrimination or calibration. Some medical applications of ANN models have been reported to present an excellent fit of the model to a given set of data. Results that were too imposing were usually derived from overfitted models, where too many inputs were entered as compared with their enrolled subjects [13, 34]. From the perspective of clinical practice, the challenge is to train an ANN model to recognize patterns without overfitting and avoiding input complexity for clinicians. Only simple clinical decision support can operate well since adding more redundant parameters may reduce the likelihood of success in implementing a computerized predictive model [35]. Simplification may expedite the employment of a decision support tool by clinicians into their regular clinical work, decreasing the chances of incompatibility and errors.

In addition, we also provide a practical visualization of input/output relationships by utilizing the three-dimensional plot achieved by using ANN modeling (Fig. 5). This plot can be applicable for rapidly determining the existence of skeletal metastasis according to the values of age and serum PSA concentration. When age or serum PSA concentration was raised the score became 1, implying that the patient would have skeletal metastasis. On the other hand, prognostication with extrapolation should be notified and the inferring territory should be inspected before clinical application since either overestimation or underestimation may lead to wrong conclusions with subsequent serious complications for patients with prostate cancer [36]. Taking advantage of this visualized graph, the interpreter can easily identify the predictable range from the data structure before the task of prediction. If the data of input variables are outliers, the interpretation should be cautious while using this three-dimensional plot derived from our ANN model.

The basic theory of the ANN is to imitate the processes of human decision-making using principles of adjustment and speculation. For example, an ANN model can learn in a manner comparable to the way nuclear medicine physicians learn: they are served with a large number of input scintigraphic images and output diagnoses, and learning occurs progressively [37]. Since a nonlinear phenomenon is a fundamental cornerstone in medicine, ANN has the ability to discover complex configurations of biomedical processes between input and output variables in a nonlinear pattern by learning algorithms and comprising more or less artificial neurons in hidden layers. For prostate cancer staging, predictions established upon ANN models are more precise than empirical rules or models based on univariate regression alone [38]. Nevertheless, ANN methodology is not without some controversial issues. The so-called black box phenomenon of the ANN hinders broad acceptance in clinical utility. Although ANN can reveal interconnected weights between individual artificial neuron during the analyzing processes, this digital fortress of "Da Vinci Code" is difficult for clinicians to interpret. As previously investigated, linear and logistic regression models also encounter the same explanatory shortfall [39]. However, the decode techniques of ANN's weights are accomplished by using sensibility analysis if many input variables are used in the ANN construction [14, 40, 41]. On the other hand, some clinical objectors argue that the engineer-based approach is a major obstacle to ANN being applied as an instrument for clinical decision support. Certain professional engineers indeed like to develop their own algorithm-based ANN by using programming language and fine tune the inside parameters to fit the specific demand for

solutions; hence, the original black box phenomenon will be another "deus ex machine". The only way to avoid this entrapment in clinical medicine is use commercially or publicly available software packages. The most influential value of these software packages is providing the capability to set various parameters in building the ANN topology. Thereafter, other investigators can validate the ANN study through the processes of repeatability and reproducibility. With the help of the friendly graphic user interface, the modern Windows-based software can ease the users' work for sophisticated computing without lacking the versatility of various ANN models. Some ANN creation software also offers the built-in function or advanced developer kit that allows packaging the trained ANN model into an executable file and making it available on the Internet for anyone to download [26].

The investigation of ANN to predict skeletal metastasis is very limited in clinical medicine. Only one study fostering an interest in its development and the attempted application of ANN to forecast bone metastasis was found in the PubMed database. Arana et al. used a feed-forward ANN with three-layer-perceptron configurations in the diagnosis of calvarial metastasis in 21 patients with different neoplasms [42]. Although their study purpose and design was different than our present study, there were some similar issues worthy. They used 19 input variables extracted from clinical data and radiologic features of computed tomography to get an excellent discriminatory power (AUC=0.93±0.04) with high sensitivity (97.9%) and specificity (95.8%). Of note, their ANN model might encounter the generalization problem and overfitted risk since they used too many input variables contrast to their participants in spite of the leave-one-out method they used. To contrast, our model used only two predictors to obtain comparable results even though the ANN model topology differed. In addition, we enrolled more patients and the ratio of enrolled patients to inputs in our study was also proper. Nevertheless, to make a confident diagnosis of metastasis, their results also emphasized the need to least includes the variable "age", which was one of two parameters in our study. On the other hand, their results were certainly impressive since they employed the radiologic features of computed tomography as part of their predictors. It is well acknowledged that computed tomography is the imaging technique of choice to detect cranial vault lesions because of its superiority in the illustration of cortical bone with internal characteristics visualization. We do not agree that the parameters derived from the reference method could be the part of predictors. Through these comparisons, we can confirm that our ANN model not only had appropriate design with adequate predictors but also had good performance to predict skeletal metastasis in patients with prostate cancer.

There are several limitations to our investigation that deserve comments. First, the number of enrolled patient number was small, but by taking advantage of bootstrap resampling with internal validation we overcame this difficulty. It is impractical to use new training, selection, and test cases selected from the sampled population since we usually have unsatisfactory data to conduct numerous training processes with independently separate training, selection, and test subsets in clinical study. To increase accuracy, the bootstrapping method which samples a dataset with substitution can be implemented any number of times. This method is not constrained by any particular classification rules. As long as the sample size remains small, that is a significant feature of costly clinical research with limited resources of substance or time; thus, a precise error estimation method that can be easily performed for small samples is probably beneficial. This bootstrapping method may be useful and it is appropriate for analyzing sample sizes as small as 16, where distributional hypothesis are vague, previous information is sparsely dispersed, and further data may be difficult to obtain [43]. Second, only one single nuclear medicine physician interpreted bone scintigraphic images. Although the interpreter is a senior physician with considerable experience, his subjectivity cannot be ruled in image interpretation. Furthermore, the imaging interpretation should use a quantitative scale or scoring system makes the scintigraphic diagnosis [44]. Accordingly, more engaged interpreters with quantitatively reading protocol will be more objective for the firmly scientific evidence. More recently the utilization of positron emission tomography/computed tomography (PET/CT), new combined modalities for imaging diagnostics, is rapidly growing. For detection of osseous metastasis in patients with high-risk prostate cancer, F-18 fluoride PET/CT is a highly sensitive and specific method whose performance is better than Tc-99m MDP bone scintigraphy [45]. Therefore, using PET/CT instead of bone scintigraphy will be considered as the next reference method for detecting skeletal metastasis.

Third, the serum PSA level was not drawn on the same day as administration of bone scintigraphy. To account for the connatural defects in a retrospective study, a shortcoming is that the criteria-matched subjects from our scintigraphic database would be too few to accurately predict whether we created data-collecting formulation that is too strict. Moreover, data from two procedures within one-month interval is reasonable since we believe the change in serum PSA levels within one month is considerable. In a rigid research design, it is certainly better to perform two procedures using a one-day protocol, even though this study is proof-of-concept investigation. Another design defect is that we did not stratify the enrolled patients into certain groups such as newly diagnosed, status post urologic operation, or follow up for treatment

response. Because the present protocol is a retrospective study, some effect might result when using serum PSA concentration as a predictor. Conversely, the unlimited participants' status might expand the usable range of the ANN model to tolerate more predictive scenarios. In a prospective study, blood sampling and bone scintigraphy should be done on the same day as well as correctly classifying the patient's status.

Lastly, our study was executed at one single center and the predictive model was not performed prospectively in other institutions. Promising performance of a predictive model for one group of patients does not guarantee its utility for other groups. To verify our findings, further longitudinal studies with a larger pool of randomly selected patients in different hospitals to decrease interinstitutional variation should be considered. Clearly, our current scheme was a proof of concept rather than an external validation of the technique. Despite these limitations, the preliminary findings are exciting and furnish the impetus for future validation studies in controlled experiments [46]. Future use of this technique may be expanded to develop the web-based software other than a stand-alone application, using an ANN model integrating it into a single-kernel engine for clinicians to perform real-time prediction.

To our knowledge, the comparative consequence of limited clinical variables has not been previously constructed by means of an ANN model in patients with prostate cancer who have skeletal metastasis. Our results show that an ANN model can accurately predict skeletal metastasis in patients with prostate cancer and that it might serve as a useful method to follow the response to all therapeutic interventions. We also clearly demonstrated the techniques for construction of an ANN model and that this kind of approach to ANN could help clinicians easily initiate the building procedures. We must emphasize that our ANN model is not intended to substitute for an experienced professional or any imaging modality such as bone scintigraphy, CT, magnetic resonance imaging, or PET; on the contrary, the ANN model can work as a supplementary tool to support clinical decisions or as an initial screening method, particularly where the diagnostic instruments are unavailable. Afterwards, the clinicians may use the information derived from the ANN prediction to decide whether further diagnosis and monitoring are necessary.

Acknowledgement This paper is in memorial of Mr. A-Tsai Lee, Dr. Jainn-Shiun Chiu's maternal grandfather, who died from prostate cancer with skeletal metastasis on 20 October 2006.

References

1. Coleman, R. E., Metastatic bone disease: clinical features, pathophysiology and treatment strategies. *Cancer Treat. Rev.* 27 (3):165–176, 2001.
2. Carlin, B. I., and Andriole, G. L., The natural history, skeletal complications, and management of bone metastases in patients with prostate carcinoma. *Cancer*. 88(12 Suppl):2989–2994, 2000.
3. Rigaud, J., Tiguert, R., Le Normand, L., Karam, G., Glemain, P., Buzelin, J. M. et al., Prognostic value of bone scan in patients with metastatic prostate cancer treated initially with androgen deprivation therapy. *J. Urol.* 168(4 Pt 1):1423–1426, 2002.
4. Rodvold, D. M., McLeod, D. G., Brandt, J. M., Snow, P. B., and Murphy, G. P., Introduction to artificial neural networks for physicians: taking the lid off the black box. *Prostate*. 46(1):39–44, 2001.
5. Forsstrom, J. J., and Dalton, K. J., Artificial neural networks for decision support in clinical medicine. *Ann. Med.* 27(5):509–517, 1995.
6. Wei, J. T., Zhang, Z., Barnhill, S. D., Madyastha, K. R., Zhang, H., and Oesterling, J. E., Understanding artificial neural networks and exploring their potential applications for the practicing urologist. *Urology*. 52(2):161–172, 1998.
7. Lisboa, P. J., A review of evidence of health benefit from artificial neural networks in medical intervention. *Neural Netw.* 15(1):11–39, 2002.
8. Anagnostou, T., Remzi, M., Lykourinas, M., and Djavan, B., Artificial neural networks for decision-making in urologic oncology. *Eur. Urol.* 43(6):596–603, 2003.
9. Miller, A. S., Blott, B. H., and Hames, T. K., Review of neural network applications in medical imaging and signal processing. *Med. Biol. Eng. Comput.* 30(5):449–464, 1992.
10. Penny, W., and Frost, D., Neural networks in clinical medicine. *Med. Decis. Mak.* 16(4):386–398, 1996.
11. Henderson, A. R., The bootstrap: a technique for data-driven statistics. Using computer-intensive analyses to explore experimental data. *Clin. Chim. Acta.* 359(1–2):1–26, 2005.
12. Das, A., Ben-Menachem, T., Cooper, G. S., Chak, A., Sivak, M. V. Jr., Gonet, J. A. et al., Prediction of outcome in acute lower-gastrointestinal haemorrhage based on an artificial neural network: internal and external validation of a predictive model. *Lancet*. 362(9392):1261–1266, 2003.
13. Banerjee, R., Das, A., Ghoshal, U. C., and Sinha, M., Predicting mortality in patients with cirrhosis of liver with application of neural network technology. *J. Gastroenterol. Hepatol.* 18 (9):1054–1060, 2003.
14. Wang, Y. F., Hu, T. M., Wu, C. C., Yu, F. C., Fu, C. M., Lin, S. H. et al., Prediction of target range of intact parathyroid hormone in hemodialysis patients with artificial neural network. *Comput. Methods Programs Biomed.* 83(2):111–119, 2006.
15. Guan, P., Huang, D. S., and Zhou, B. S., Forecasting model for the incidence of hepatitis A based on artificial neural network. *World J. Gastroenterol.* 10(24):3579–3582, 2004.
16. Dreiseitl, S., and Ohno-Machado, L., Logistic regression and artificial neural network classification models: A methodology review. *J. Biomed. Inform.* 35(5–6):352–359, 2002.
17. Linden, A., Measuring diagnostic and predictive accuracy in disease management: an introduction to receiver operating characteristic (ROC) analysis. *J. Eval. Clin. Pract.* 12(2):132–139, 2006.
18. Hanley, J. A., and McNeil, B. J., A method of comparing the areas under receiver operating characteristic curves derived from the same cases. *Radiology*. 148(3):839–843, 1983.
19. Chatzicostas, C., Roussomoustakaki, M., Notas, G., Vlachonikolis, I. G., Samonakis, D., Romanos, J. et al., A comparison of Child-Pugh, APACHE II and APACHE III scoring systems in predicting hospital mortality of patients with liver cirrhosis. *BMC Gastroenterol.* 3:7, 2003.
20. Lemeshow, S., and Hosmer, D. W. Jr., A review of goodness of fit statistics for use in the development of logistic regression models. *Am. J. Epidemiol.* 115(1):92–106, 1982.

21. Hosmer, D. W., Hosmer, T., Le Cessie, S., and Lemeshow, S., A comparison of goodness-of-fit tests for the logistic regression model. *Stat. Med.* 16(9):965–980, 1997.
22. Chen, C. A., Lin, S. H., Hsu, Y. J., Li, Y. C., Wang, Y. F., and Chiu, J. S., Neural network modeling to stratify peritoneal membrane transporter in predialytic patients. *Intern. Med.* 45 (9):663–664, 2006.
23. Partin, A. W., Kattan, M. W., Subong, E. N., Walsh, P. C., Wojno, K. J., Oesterling, J. E. et al., Combination of prostate-specific antigen, clinical stage, and Gleason score to predict pathological stage of localized prostate cancer. A multi-institutional update. *JAMA*. 277 (18):1445–1451, 1997.
24. Kattan, M. W., Stapleton, A. M., Wheeler, T. M., and Scardino, P. T., Evaluation of a nomogram used to predict the pathologic stage of clinically localized prostate carcinoma. *Cancer.* 79(3):528–537, 1997.
25. Murphy, G. P., Snow, P. B., Brandt, J., Elgamal, A., and Brawer, M. K., Evaluation of prostate cancer patients receiving multiple staging tests, including ProstaScint scintiscans. *Prostate*. 42 (2):145–149, 2000.
26. Batuello, J. T., Gamito, E. J., Crawford, E. D., Han, M., Partin, A. W., McLeod, D. G. et al., Artificial neural network model for the assessment of lymph node spread in patients with clinically localized prostate cancer. *Urology.* 57(3):481–485, 2001.
27. Han, M., Snow, P. B., Brandt, J. M., and Partin, A. W., Evaluation of artificial neural networks for the prediction of pathologic stage in prostate carcinoma. *Cancer.* 91(8 Suppl):1661–1666, 2001.
28. Tewari, A., and Narayan, P., Novel staging tool for localized prostate cancer: a pilot study using genetic adaptive neural networks. *J. Urol.* 160(2):430–436, 1998.
29. Crawford, E. D., Batuello, J. T., Snow, P., Gamito, E. J., McLeod, D. G., Partin, A. W. et al., The use of artificial intelligence technology to predict lymph node spread in men with clinically localized prostate carcinoma. *Cancer.* 88(9):2105–2109, 2000.
30. Bates, D. W., Kuperman, G. J., Wang, S., Gandhi, T., Kittler, A., Volk, L. et al., Ten commandments for effective clinical decision support: making the practice of evidence-based medicine a reality. *J. Am. Med. Inform. Assoc.* 10(6):523–530, 2003.
31. Carter, H. B., Epstein, J. I., and Partin, A. W., Influence of age and prostate-specific antigen on the chance of curable prostate cancer among men with nonpalpable disease. *Urology.* 53(1):126–130, 1999.
32. Jung, K., Lein, M., Stephan, C., Von Hosslin, K., Semjonow, A., Sinha, P. et al., Comparison of 10 serum bone turnover markers in prostate carcinoma patients with bone metastatic spread: diagnostic and prognostic implications. *Int. J. Cancer.* 111(5):783–791, 2004.
33. Stephan, C., Xu, C., Brown, D. A., Breit, S. N., Michael, A., Nakamura, T. et al., Three new serum markers for prostate cancer detection within a percent free PSA-based artificial neural network. *Prostate*. 66(6):651–659, 2006.
34. Oates, J. C., Varghese, S., Bland, A. M., Taylor, T. P., Self, S. E., Stanislaus, R. et al., Prediction of urinary protein markers in lupus nephritis. *Kidney Int.* 68(6):2588–2592, 2005.
35. Martich, G. D., Waldmann, C. S., and Imhoff, M., Clinical informatics in critical care. *J. Intensive Care Med.* 19(3):154–163, 2004.
36. Yamamura, S., Takehira, R., Kawada, K., Nishizawa, K., Katayama, S., Hirano, M. et al., Application of artificial neural network modelling to identify severely ill patients whose amino-glycoside concentrations are likely to fall below therapeutic concentrations. *J. Clin. Pharm. Ther.* 28(5):425–432, 2003.
37. Boone, J. M., Gross, G. W., and Greco-Hunt, V., Neural networks in radiologic diagnosis. I. Introduction and illustration. *Invest. Radiol.* 25(9):1012–1016, 1990.
38. O'Dowd, G. J., Veltri, R. W., Orozco, R., Miller, M. C., and Oesterling, J. E., Update on the appropriate staging evaluation for newly diagnosed prostate cancer. *J. Urol.* 158(3 Pt 1):687–698, 1997.
39. Hurwitz, G. A., Weingert, M. E., Silver, D. L., MacDonald, A. C., Finnie, K. J., Powe, J. E. et al., The usefulness of stress tests performed in the nuclear medicine department: mathematical methods to assess efficacy at various angiographic endpoints. *Nucl. Med. Commun.* 17(6):463–474, 1996.
40. Hunter, A., Kennedy, L., Henry, J., and Ferguson, I., Application of neural networks and sensitivity analysis to improved prediction of trauma survival. *Comput. Methods Programs Biomed.* 62 (1):11–19, 2000.
41. Heckerling, P. S., Gerber, B. S., Tape, T. G., and Wigton, R. S., Entering the black box of neural networks. *Methods Inf. Med.* 42 (3):287–296, 2003.
42. Arana, E., Marti-Bonmati, L., Bautista, D., and Paredes, R., Qualitative diagnosis of calvarial metastasis by neural network and logistic regression. *Acad. Radiol.* 11(1):45–52, 2004.
43. Fu, W. J., Carroll, R. J., and Wang, S., Estimating misclassification error with small samples via bootstrap cross-validation. *Bioinformatics.* 21(9):1979–1986, 2005.
44. Fujimoto, R., Higashi, T., Nakamoto, Y., Hara, T., Lyshchik, A., Ishizu, K. et al., Diagnostic accuracy of bone metastases detection in cancer patients: comparison between bone scintigraphy and whole-body FDG-PET. *Ann. Nucl. Med.* 20(6):399–408, 2006.
45. Even-Sapir, E., Metser, U., Mishani, E., Lievshitz, G., Lerman, H., and Leibovitch, I., The detection of bone metastases in patients with high-risk prostate cancer: 99mTc-MDP Planar bone scintigraphy, single- and multi-field-of-view SPECT, 18F-fluoride PET, and 18F-fluoride PET/CT. *J. Nucl. Med.* 47(2):287–297, 2006.
46. Cross, S. S., Harrison, R. F., and Kennedy, R. L., Introduction to neural networks. *Lancet*. 346(8982):1075–1079, 1995.