Annals of Biomedical Engineering, Vol. 33, No. 6, June 2005 (© 2005) pp. 811–820
DOI: 10.1007/s10439-005-2867-7

A Model for Detecting Balance Impairment and Estimating Falls Risk in the Elderly

MICHAEL E. HAHN[1,2] and LI-SHAN CHOU[1]

[1]Department of Human Physiology, University of Oregon, Eugene, Oregon 97403; [2]Present address: Department of Health and Human Development, Montana State University, Bozeman, Montana 59717

(Received 7 April 2004; accepted 18 January 2005)

Abstract—Traumatic falls are a prevalent and costly threat to elderly adults. Accurate risk assessment is necessary for reducing incidence of falls. The objective of this study was to test the feasibility of a balance impairment detection model using tasks of sample categorization and falls risk estimation. Model design included an artificial neural network and a statistical discrimination method. The first system produced an individual categorization value, which was then assessed in the second system for relative risk of falls, compared to a normative distribution of healthy elderly peers. Input data included leg muscle electromyographic amplitudes, temporal-distance measures of gait, and medio-lateral measures of whole body center of mass motion. These input data were compiled from a sample of healthy elderly adults ($n = 19$) and a sample with impaired balance ($n = 10$) to develop and test the model. Accuracy of sample categorization was assessed using a relative operating characteristic (ROC) value. For relative risk estimation, categorical delineation of risk level was adopted. Sample categorization results reached ROC values of 0.890. Relative risk was frequently assessed at high or very high risk for experiencing falls. Temporal-distance measures were most influential in categorization accuracy, producing the most consistent risk estimates. Combined inputs further improved model performance. This model shows potential for detecting balance impairment and estimating falls risk; thereby indicating need for referral for falls prevention intervention.

Keywords—Artificial neural network, Sample categorization, Relative Risk, Prevention.

INTRODUCTION

Traumatic falls in the elderly are prevalent, debilitating and costly, with over 35% of the elderly population experiencing falls[5] and approximately \$20.2 billion spent in treatment each year.[1] Accurate assessment of the risk of falls is critical to reducing the incidence of falls. Many studies have attempted to predict falls prospectively, with varying results. Studies by Topper *et al.*[31] and Maki *et al.*[17] used measures of static posturography to indicate risk of falls. Results from their work showed that control of medio-lateral sway may be a strong predictor of falls in the elderly. Another study by Graafmans *et al.*[6] reported that general mobility impairments are strongly associated with recurrent falls. No measures of diagnostic accuracy were reported in any of these studies.

More predictive models were reported by Shumway-Cook *et al.*[26] using logistic regressions that combined Berg Balance scores with a self-reported history of imbalance to predict risk of falls. Results from this retrospective study produced a sensitivity of 91% and specificity of 82%. Further results[27] indicated that the timed-up-and-go test was also accurate in assessing risk of falls (sensitivity = 87%; specificity = 87%). These two studies showed that relatively simple clinical measures could predict risk of falls with reasonable accuracy.

Recently, a few research groups have reported fall-risk screening in community-dwelling elderly using logistic regression models based primarily on a previous history of falls. Tromp *et al.*[32] relied on reports of visual impairment and urinary incontinence in addition to a history of falls, to estimate risk using the relative operating characteristic (ROC) measure;[30] producing overall accuracy of 0.71 for predicting recurrent falls. Another study[28] relied on history of falls, measures of postural sway, hand grip strength, and "a depressive state of mind" to estimate risk of recurring falls (ROC = 0.79).

Other groups investigated risk estimation using measures familiar to the in-patient hospital or nursing home setting. Halfon *et al.*[10] used a Poisson regression model, resulting in five variables which predicted risk: age, gender, morbidity predisposition, surgical procedure and length of stay. No values of accuracy were reported. Izumi *et al.*[14] reported that in institutionalized settings (long-term care) the primary predictor of falls was that of nurses' opinion/prediction. Interestingly, this study also reported that the primary indicator of falls risk in a general hospital setting was mobility (including independent ambulation). Results from their study were quite low in prediction accuracy; ranging from 34.8 to 45.2%.

Address correspondence to Li-Shan Chou, Department of Human Physiology, 1240 University of Oregon, Eugene, Oregon 97403. Electronic mail: chou@uoregon.edu

0090-6964/05/0600-0811/1 © 2005 Biomedical Engineering Society

Many of the previous studies restricted predictive variable selection to static measures of posture, or clinical estimates of overall health; rather than dynamic musculoskeletal measures which may better represent balance control during activities when falls are likely to occur. One limitation of the studies that did include dynamic or musculoskeletal strength measures is that the models did not allow for prospectively testing the efficacy of interventions. This limitation may be significant, in that application of risk-estimation results can not be effectively used in falls-prevention, unless the model allows interventions to be examined prospectively. Exercise and strength training have become more common in recent years as an intervention for maintaining functional mobility in the elderly. Although improvements in strength and walking function were observed, studies only revealed that general exercise and strength training had a beneficial effect on fall incidence and could not determine which type of exercise was most effective.[22,33] There is, therefore, an ultimate goal to develop a model that can not only accurately assess the risk of falling of an older individual but also have the ability to predict possible outcomes of the prescribed muscle strengthening intervention.

The first step is to establish a model that can provide a more accurate estimation of falls risk in the elderly by taking into account the level of muscular challenge (electromyography), dynamic gait performance (temporal-distance parameters), and measures of whole-body dynamic stability. A few approaches could be used to develop such a model; linear regression or principal component analysis for example. However, both of these techniques restrict the input predictors to only those which explain a high amount of variability in the system. Furthermore, both of these models rely on linear relationships between variables. One approach which allows inclusion of more variables and is tolerant of non-linear relationships is to use artificial neural network (ANN) theory to map dynamic measures of gait and whole-body stability onto estimates of individual risk. Recent years have seen an increase in the use of ANN analyses applied to human locomotion.[2] Most of the studies have used back-propagation learning algorithms[23] and relatively simple ANN designs.[11,15,20,21,24,25]

One of the strengths of multi-layer neural network models is their robust ability to map inputs non-linearly onto resulting outputs. For this reason, neural network modeling theory was chosen in this study as the method for, first assessing the level of functional mobility through gait measures, and secondly to estimate the individual level of risk for falls in each case. Once an individual's risk of falls is estimated, it is possible to prescribe interventions best suited to that individual. The architecture of a neural network model has the potential to allow for simulation of improvements in balance control as a result of changing muscular strength inputs.

In previous reports, quantitative measures of dynamic stability and neuromuscular challenge were identified and examined for age effects.[7,9] An ANN model was then designed to map the relationships between measures of muscular challenge and dynamic stability.[8] The purpose of this study was to test the feasibility of using the model to categorize two samples of older adults and to estimate relative risk.

METHODS

Description of Research Subjects

For the diagnostic network designed and implemented in this study, network training was performed based on the data from a mixed sample of elderly subjects ($n = 29$), composed of 19 healthy elderly and 10 'fallers.' Network training groups were selected randomly, with 70% of each subject sample being selected. Table 1 provides descriptive information for the two subject samples. Inclusion criteria for the healthy elderly subjects required no histories of significant head trauma, neurological disease (e.g. Parkinson's, post-polio syndrome, diabetic neuropathy), visual impairment not correctable with lenses, musculoskeletal impairments (e.g. amputation, joint replacement, joint fusions, joint deformity due to rheumatoid arthritis), or persistent symptoms of vertigo, light-headedness, unsteadiness. Healthy elderly subjects were noted to be active community members, with many of them currently involved in recreational sporting activities. Ten elderly subjects with complaints of imbalance during walking or a history of falls, were recruited from the local community. Imbalance for these individuals was self-reported. A history of falls was defined as an occurrence of two or more falls. A fall was defined as any event in which the individual lost their balance and made contact with the floor (i.e. did not simply fall back into a chair after trying to stand up). Subject responses to inclusion criteria were confirmed by a consulting Physiatrist. Three subjects with imbalance were diagnosed with either unilateral or bilateral vestibular weakness. All

TABLE 1. Subject information; age, height and mass; mean (SD).

Sample	Female	Male	Age (years)	Height (cm)	Mass (kg)
Healthy elderly ($n = 19$)	8	11	71.5 (5.6)	167.9 (9.8)	73.3 (14.1)
Elderly/imbalance ($n = 10$)	7	3	78.5 (4.7)	162.0 (9.8)	72.2 (14.4)

subjects with balance disorders were community-dwelling and able to walk more than 100 m without the use of gait aides at the time of testing. The experimental protocol was approved by the Institutional Review Board and experimental procedures were explained to all subjects prior to testing, with verbal and written consent obtained.

Empirical Data

Input data consisted of normalized electromyography (EMG) data of the lower extremities (gluteus medius, vastus lateralis, medial gastrocnemius), temporal-distance (T-D) measures of gait (gait velocity, stride length, stride time, step width), and medio-lateral (M-L) motion (displacement and peak velocity) of the whole body center of mass (COM). Input data were averaged within each subject, over three individual gait trials. Input data were collected using the following gait protocol and analysis procedures.[7,9]

The gait protocol included walking on level ground with no obstructions and stepping over an obstacle corresponding to 2.5% of each subject's height. All subjects performed the trials at a self-selected pace, while barefoot. Whole body motion data were collected using a six-camera ExpertVision system (Motion Analysis Corp., Santa Rosa, CA). The COM position was estimated throughout the obstacle crossing stride, using a weighted-sum approach with a 13-link anthropometric model.[3,13,19] Velocities were calculated using the generalized, cross-validated spline algorithm.[34] Ranges of M-L displacement and peak M-L velocity values were compiled due to previous findings indicating their ability to distinguish between individuals with and without balance impairment.[4]

Dynamic EMG measures were taken from pre-amplified surface electrodes (Motion Lab Systems, Inc., Baton Rouge, LA) placed bilaterally over the bellies of the gluteus medius (GM), vastus lateralis (VL) and medial gastrocnemius (GA). Activation magnitudes of each muscle during gait were normalized to values taken during maximal effort manual muscle testing (MMT). Maximal effort EMG was thus set at 100%, with all dynamic magnitudes calculated proportionally with respect to 100%. This method of EMG normalization is widely accepted in human movement literature. Maximal GM activation was tested in 30° of hip abduction, while side lying. For VL maximum, subjects were seated with the knee in 45° of flexion. Maximal GA activation was tested in neutral ankle position, with the subject fixed to a table in prone position. MMT procedures were performed by one examiner for each muscle group, bilaterally. Subjects were verbally encouraged to ensure maximal recruitment.

During gait trials, only the peak phases of activity were used as EMG inputs into the model. Peak phases of activity corresponded to periods of double support for the GM and VL; single support for the GA. Gait velocity, stride length (normalized to body height), stride time and step width were measured during the obstacle-crossing stride, and used as inputs into the ANN model. From the various measurements taken during the gait analysis protocol, a total of 12 variables were available for input into the ANN model (2 COM, 6 EMG and 4 T-D measures).

ANN Input Combinations

These various combinations of input data were tested for their ability to accurately categorize fallers: EMG, T-D and COM parameters alone, then EMG and T-D together, and finally EMG, T-D, COM parameters all entered as input predictors. The reason for solely entering the EMG and T-D data as "paired" input was that these data were shown in previous work to accurately predict COM motion variables when entered into a simple ANN structure.[8] Paired entry of EMG and T-D data therefore serves as a substitution method in the event that COM data cannot be compiled for analysis.

ANN Architecture

The model used in this study was designed to assess interactions between muscular demands, gait performance, and balance control during walking; it was comprised of two systems (Fig. 1). The first system, an ANN model, estimated the category of non-faller/faller (0 or 1, respectively), based on the empirical inputs listed earlier. The second system discriminately classified the relative risk that an individual will experience falls, using output from the first system. Used in combination, these two systems can demonstrate the relative distance from a healthy, normative distribution of age-matched subjects, thereby giving a scaled estimate of risk.

The ANN structure was a feed-forward, three-layer network with scaled error correction conducted via the Levenberg–Marquardt back-propagation algorithm.[16,18] The first layer consisted of a variable number of input units (specific to the data type being tested). The second layer held a variable number of 'hidden' units ($H = 5$, 10, 20). These hidden units are generally considered to be the information-processing level of artificial networks, allowing parallel distributed processing to take place.[23] The final layer of the ANN consisted of a sole output unit. This output unit reflected the network's estimate of the faller/non-faller category. The network was trained using a set of training error goals ($E = 0.01$, 0.001, 0.0001). The training proportion was set at 0.7 for all tests. Twenty tests (with randomly selected training samples) were made for each of the nine combinations of ANN settings (H and E) that were possible (resulting in 180 training attempts made for each combination of input parameters). The hidden and output units summed incoming weighted connections, and processed an outgoing activation signal with a sigmoidal transfer function in the hidden units, and a pure linear transfer function in the output units.[11] Figure 2 shows a detailed diagram of the ANN structure used.

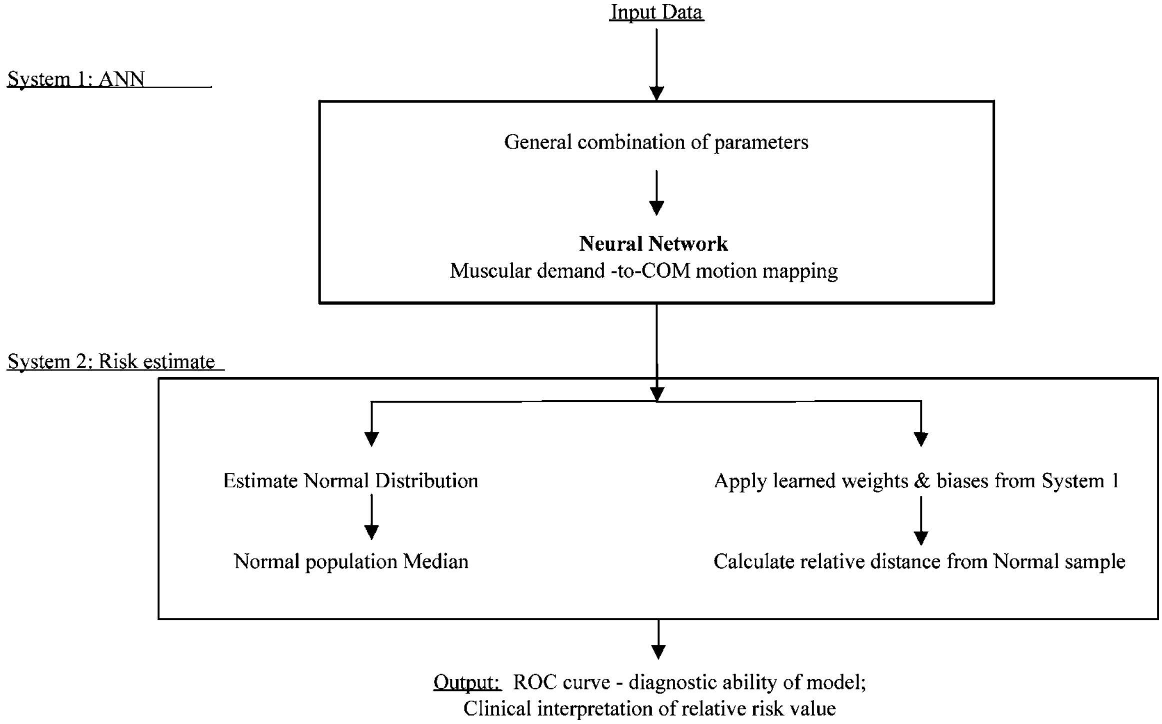

FIGURE 1. Flow chart describing the development of the two-system model.

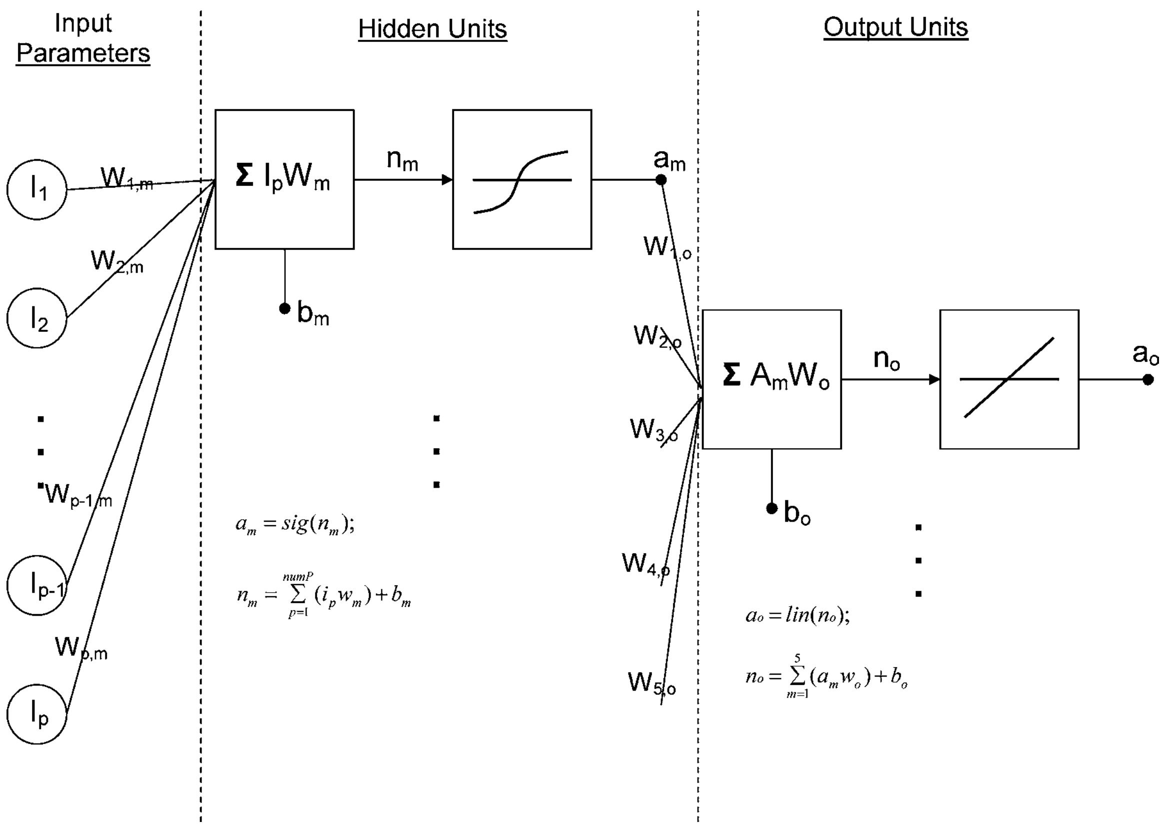

FIGURE 2. Diagram describing the neural network architecture. A tangential sigmoid activation transfer function was used in the hidden layer, and a pure linear activation transfer function in the output layer.

Analysis Procedures

The value of the ANN's output unit provided an estimate of the likelihood that an individual be categorized as healthy, or as a faller. For the output units, the result of the transfer function was the predicted category value for either 0, or 1. An output value of 0 indicated healthy, and a value of 1 indicated the faller category. Error correction was performed using the Levenberg–Marquardt learning algorithm,[16,18] known to provide rapid convergence and robust generality in small networks being trained on small-to-moderate sample subsets. Further details on the ANN model design have been reported previously.[8]

Sample Categorization

After training was completed in each attempt, cases from the testing set (30% of the initial subject sample) were input into the network with fixed ('learned') weights/biases. Each case then received an output value estimating their likelihood of being either a healthy subject, or a faller (0 or 1, respectively). A distribution function was generated from the ANN output for the healthy subject sample. Categorization required selection of a decision line at some value that was thought to best delineate between healthy elderly and fallers' output data. In this study, the median of the healthy subjects' output distribution was conservatively selected as the decision line (X_0). A new value for X_0 was calculated with each training session. A distribution function was also generated for the ANN output of the sample of elderly fallers, for initial validation of the categorizations made by the decision line. Figure 3 depicts how these distributions are shaped and how a decision line is used.

The effectiveness of categorization to either side of the decision line was measured by the ROC value calculated for the system. The ROC is a method derived from signal detection theory that has been described as providing "a precise and valid measure of diagnostic accuracy."[30] Specifically, the ROC value is calculated as the area under a curve which is generated from the specificity and sensitivity results of multiple diagnostic testing sessions. Specificity is defined as the proportion of false-positives (individuals categorized as being a 'faller,' when they are in fact a 'non-faller'). Sensitivity is defined as the proportion of true-positives (individuals appropriately categorized as 'fallers'). For each test, discrete values for specificity and sensitivity were calculated. A curve was thus shaped from the specificity (x-axis) and sensitivity (y-axis) values calculated from each of the multiple tests ($n = 20$) conducted for each type of input. The area under the resulting curve is interpreted as the overall diagnostic accuracy of the classification system (ROC value). This method of assessing diagnostic accuracy has been used widely in medical diagnostics,[30] and more recently in studies of fall-risk estimation.[28,32] A representative ROC curve is presented in Fig. 4.

Relative Risk Estimation

Network output data from the testing cases (which the network had not been trained on) were compared against X_0. If the testing case predicted value (X_1) was below X_0 it was inferred that the test case was similar to the normal, healthy population. If X_1 was greater than X_0, the inference was that the case was away from normal. The relative distance (D_r) of the predicted value from X_0 was used as an index for risk estimation, defined as:

$$D_r = \frac{1 - X_1}{1 - X_0} \tag{1}$$

Using this distance metric, it is possible for the difference between each case and the normal healthy sample to be quantified (i.e. relative risk of falling). For example, if one case receives ANN estimation identical to X_0, it will result in a D_r of 1.0 (identical to healthy sample median). However, a case resulting in ANN estimation of X_1 greater than X_0, would result in a lower value of D_r, indicating a substantial relative distance from the healthy comparison sample. In this way, as D_r approaches or exceeds 0.0, the risk of falling

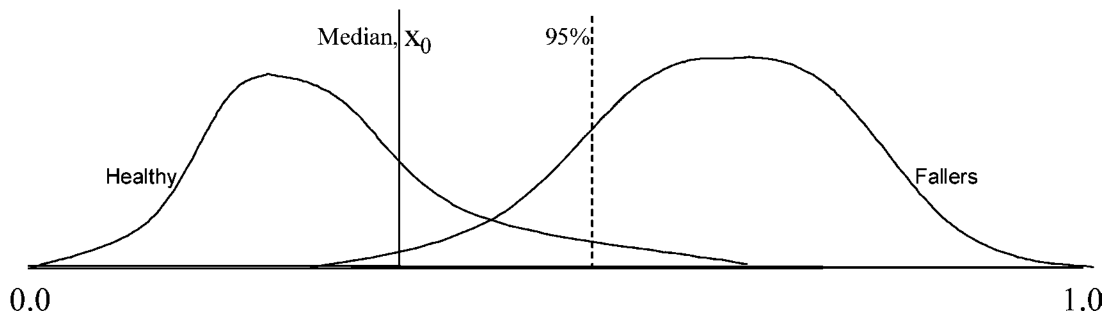

FIGURE 3. Two ANN output distribution curves representing healthy elderly subjects on the left, and those with reduced balance control on the right. Two viable decision lines are shown, at the median of the healthy subjects' distribution and at the upper 95% mark of the healthy distribution. Categorizations made using the 95% decision line would result in a lower proportion of false positives, at the cost of reducing the proportion of true-positives.

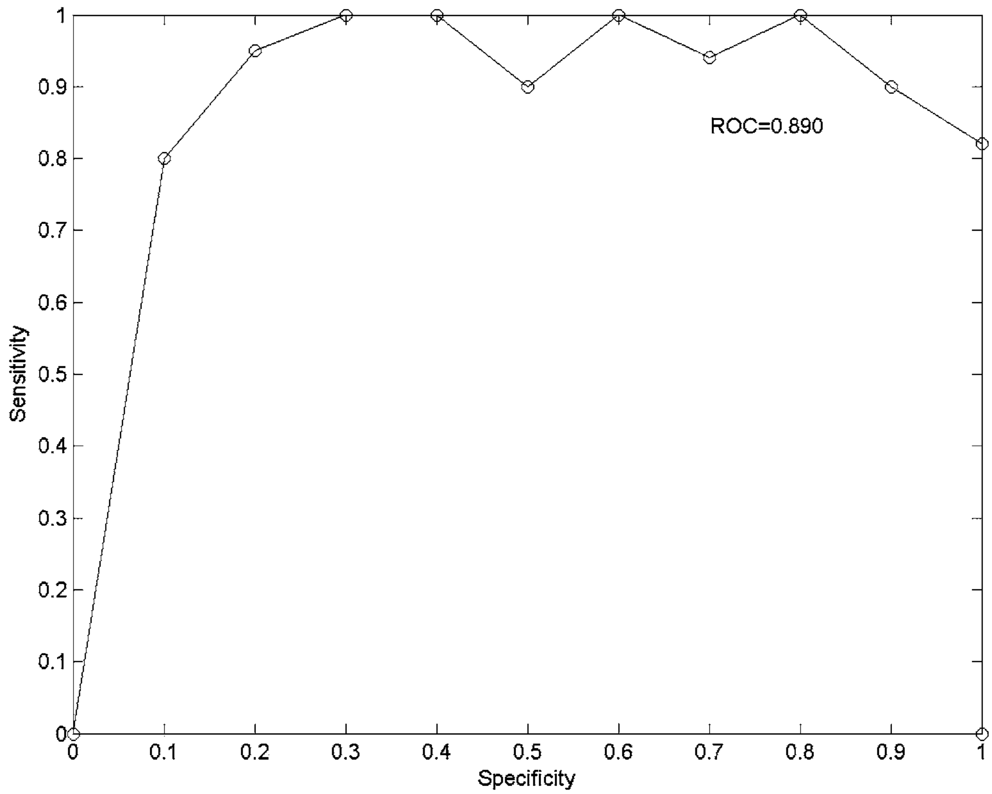

FIGURE 4. Representative ROC curve showing 89% overall classification accuracy. Sensitivity represents the proportion of true-positives; Specificity represents the proportion of false-positives. This curve was generated from ANN settings of $H = 5$, $E = 0.001$, and the combined input data of EMG and T-D.

increases, and as D_r approaches or exceeds 1.0, the relative risk of falling decreases. Therefore, the calculated D_r value is relatively inverse to the output of the ANN. A similar calculation of a distance metric has been used previously to categorize ankle function in patients with arthrodesis.[29] Figure 5 provides an example of how an individual case can be interpreted using this distance measure.

For delineation of the relative levels of risk, general category boundaries were set as follows:

Very low risk	$D_r \geq 1.00$,
Low risk	$1.00 > D_r > 0.50$,
Moderate risk	$0.50 > D_r > 0.25$,
High risk	$0.25 > D_r > -1.00$,
Very high risk	$D_r < -1.00$

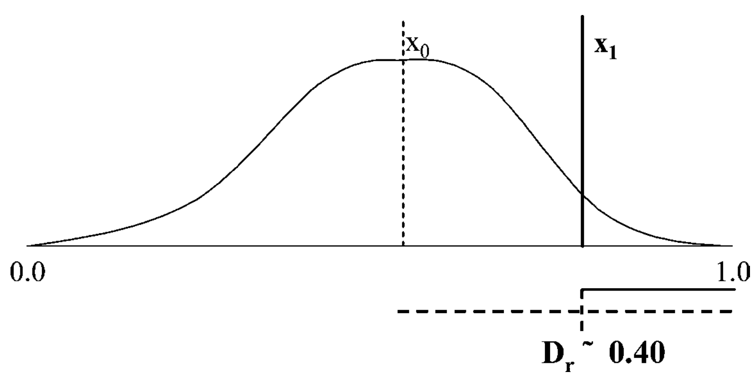

FIGURE 5. Example of risk estimation for an individual case, where the ANN output is ~0.8. In that case, the D_r value would be ~0.40, indicating a relatively moderate risk of falling. X_0 indicates the median value (decision line) of the healthy subjects' distribution; X_1 indicates the predicted value of an example test case.

TABLE 2. Overall accuracy results of the sample categorization (ROC value).

Input	Goal	No. of hidden units	ROC
EMG	0.01	5	0.689
T-D	0.001	5	0.807
COM	0.0001	5	0.702
EMG, T-D	0.001	5	0.890
EMG, T-D, COM	0.001	5	0.884

Note. For each input data type, the training goal and number of hidden units are reported as the settings that resulted in the highest accuracy.

RESULTS

Sample Categorization

Accuracy of the diagnostic network was moderately high, with top ROC values ranging from 0.689 to 0.890. Accuracy was lowest when COM or EMG data were used as the sole input. When T-D was the sole input, diagnostic accuracy increased substantially. Accuracy was further increased when the data types were combined, with overall best accuracy being achieved when the combined input of EMG and T-D data were used. Table 2 provides details of the categorization accuracy results. For each input data type, the training goal and number of hidden units are reported as the settings that resulted in the highest accuracy.

There was no discernible trend in the effect of training error on resulting ROC values. For most input data combinations the response of the ANN to more specific training (lower E value) resulted in no change, or very slight change to the positive or negative. The only combination of input data which resulted in improved diagnostic performance was the combination of EMG, T-D and COM. As the training error was reduced, the overall diagnostic accuracy increased for this combination of inputs. Table 3 provides descriptive data of these findings.

The effect of increasing the number of hidden units produced a very discernible trend (Fig. 6). For every combination of input data five hidden units produced the highest ROC value. As H increased to 10, there was a sharp decrease in diagnostic accuracy for all input combinations. Finally, as H increased to 20 a marked improvement in ROC values occurred, though still less than the original setting of $H = 5$. The combined inputs of EMG and T-D, and EMG, T-D and COM consistently produced the highest ROC values.

TABLE 3. Overall effect of the training error level (E) on diagnostic accuracy (ROC).

Input	$E = 0.01$	$E = 0.001$	$E = 0.0001$
EMG	0.532	0.532	0.503
T-D	0.634	0.612	0.637
COM	0.570	0.552	0.557
EMG, T-D	0.708	0.687	0.652
EMG, T-D, COM	0.653	0.661	0.714

Note. Values are averaged from the three different settings of hidden unit number (H).

Relative Risk Estimation

Risk estimation using the relative distance value (D_r) produced varied results across the subjects with balance disorders, and across the different types of input data. When EMG data were used alone as input, a majority of the subjects were estimated to be at moderate ($D_r < 0.50$) to high risk ($D_r < 0.25$); however, two of the subjects received D_r values of greater than 1.10 (indicating no greater risk of falls, compared with healthy peers). With T-D data as input, all 10 subjects were estimated to be at high risk of falls ($D_r < 0.25$), with 1 subject estimated at very high risk ($D_r = -1.86$). Entry of COM data resulted in an estimate that seven subjects were at high risk ($D_r < 0.001$), two at moderate risk ($D_r < 0.50$), and one subject at relatively low risk ($D_r = 0.70$). When EMG and T-D data were combined as input, nine subjects were estimated to have high risk of falls ($D_r < 0.20$) and one subject was estimated to be at very high risk ($D_r = -1.44$). With the combination of all three data types as input (EMG, T-D and COM), nine of the subjects were estimated to have high risk of falls ($D_r < 0.10$), while one subject was estimated to have relatively low risk ($D_r = 0.68$). Details of the risk estimation results are provided in Table 4.

DISCUSSION

This study sought to test the ability of an ANN model, recently developed to map muscular inputs and gait measurements onto balance control, in estimating the risk of falls in the elderly. The first task chosen in this study was to discriminate a random sample of healthy elderly adults and elderly with balance impairments into the categories of ‘healthy,’ or ‘faller.’ Secondly, an applied feasibility task was constructed in which the level of relative risk was estimated for each individual with balance impairment ($n = 10$).

In sample categorization the model performed respectably, with overall diagnostic accuracy reaching 0.89. Categorization in the present study indicates that 9 out of 10 elderly subjects with imbalance could be detected as being at risk of falls. Other studies that have reported faller categorization accuracy using the ROC measure indicated diagnostic accuracies of 0.71 [32] and 0.79.[28] These two studies relied on reports of visual impairment and urinary incontinence in addition to falls history,[32] as well as measures of postural sway, hand grip strength, and “a depressive state of mind”[28] to categorize older individuals as fallers or nonfallers.

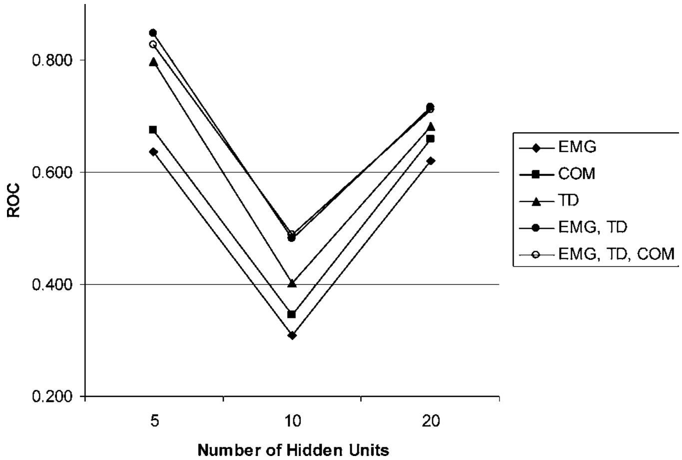

FIGURE 6. Overall effect of increasing the number of hidden units (*H*). The ROC value was highest with all data input types, when $H = 5$; decreased sharply when increased to $H = 10$; and then improved substantially when increased to $H = 20$. Values are averaged from the three different settings of training error (*E*).

Examination of model accuracy arising from each type of input and the ANN settings that produced higher accuracy yields some interesting discoveries. In a recent study by Chou *et al.*,[4] it was determined that variability of individual T-D parameters was enough to cause insignificance in the differences between healthy older adults and those with balance disorders. Interestingly, the present model showed improvement in sample categorization accuracy when it included multiple T-D parameters as input. When COM parameters were the sole input, categorization accuracy was at 0.702, lower than EMG or T-D parameter on their own. This finding indicates that while M-L COM motion may show less variability within samples (allowing significant differences to be detected);[4] it is perhaps not variable enough to allow robust generality in sample categorization through ANN modeling. To an extent, as dataset variability increases, the accuracy of ANN modeling also increases.

Results also showed that accuracy of the model's sample categorization improved when different types of data were combined as inputs. Combination of EMG and T-D data produced the highest ROC value (0.890), while combination of all data types (EMG, T-D, and COM motion) yielded a slightly lower value (0.884). Improvement in sample categorization with a more diverse dataset is not surprising in ANN models. It is thought that the more variety a neural network sees in data patterns, the more robust it will be in identifying similar patterns when new data is encountered. In the second task of estimating individuals' relative risk of falls, it appears that the relative distance index (D_r) provided adequately spaced values for delineation of risk. In fact, the D_r values were spread across a wider range than expected. Instead of being confined within the arbitrary boundaries of 0.0 and 1.0, index values reached beyond −1.25 and 1.25. As the ANN model used in this study converged to solution rapidly, it is not surprising that successful discrimination

TABLE 4. Relative risk values (D_r), for the 10 elderly subjects with balance impairment.

Balance impaired subject	EMG	T-D	COM	EMG, T-D	EMG, T-D, COM
1	−0.42 (*h*)	−0.68 (*h*)	0.70 (*l*)	−0.60 (*h*)	0.09 (*h*)
2	0.44 (*m*)	−1.86 (*vh*)	−0.42 (*h*)	−0.39 (*h*)	−0.73 (*h*)
3	−0.95 (*h*)	−0.32 (*h*)	−0.48 (*h*)	−1.44 (*vh*)	−0.74 (*h*)
4	−0.20 (*h*)	−0.66 (*h*)	0.00 (*h*)	0.10 (*h*)	0.68 (*l*)
5	−0.02 (*h*)	0.03 (*h*)	0.46 (*m*)	0.18 (*h*)	−0.34 (*h*)
6	1.47 (*vl*)	−0.85 (*h*)	−0.48 (*h*)	−0.48 (*h*)	−0.54 (*h*)
7	1.12 (*vl*)	−0.86 (*h*)	−0.42 (*h*)	−0.16 (*h*)	−0.16 (*h*)
8	0.00 (*h*)	−0.97 (*h*)	−0.42 (*h*)	−0.52 (*h*)	−0.77 (*h*)
9	−0.65 (*h*)	0.13 (*h*)	−0.80 (*h*)	−0.13 (*h*)	−0.50 (*h*)
10	−0.23 (*h*)	−0.28 (*h*)	0.27 (*m*)	−0.02 (*h*)	−0.44 (*h*)

Note. D_r values are given for each type of data input. Lower values indicate greater relative risk of falls. Risk category distinctions are indicated in parentheses; vl: very low, l: low, m: moderate, h: high, vh: very high.

between two categories necessitated a wide range of index values. If the network converged more slowly, it may be that the index values would fit a 0.0–1.0 range; however, this would likely result in 'overtraining' by the network. Overtraining often results in a reduced ability for the network to generalize. The wide range of index values simply requires wider category delineations for levels of risk (i.e. low, moderate, high, and very high).

Results from the relative risk estimation revealed a unique estimation pattern for subject #4 (Table 4). When the single data types were used as predictive inputs, the subject was estimated to have 'high' risk, but when all three data types were entered, the subject's risk was then estimated to be 'low.' This pattern is unique compared to the other subjects' estimations. One explanation of the low risk estimate when data were grouped is that the individual demonstrated conflicting trends with each type of input data, thus serving to cancel the predictive strengths of any one type of data. Having a subject display conflicting input data reveals the difficulty involved in applying a diagnostic technique universally. There is a small likelihood that some individuals will always fall outside of the detection pattern established by a diagnostic algorithm.

Ideally, a risk prediction model should not only yield a yes/no answer, but also provide an estimate of the severity of risk. One limitation of this study is that the relative severity of each subject with balance impairment was not quantified previously. If previous measures of risk severity had been available, the risk estimation from the present model could be sufficiently validated. However, it may be said that there are presently no 'gold-standard' methods for assessing relative risk of falls, and so direct validation of our model predictions to another technique is not currently possible. Given that many of our subjects reported previous falls and/or had demonstrated significant impairment in balance control, it is more likely that they would be categorized as moderate to high risk. It would be interesting to compare risk estimation levels of individuals who have not fallen previously, but are beginning to show signs of imbalance.

In this study, ANN modeling theory was chosen due to its ability to include a broad range of predictive variables and its strength in mapping non-linear relationships. The primary limitation in using an ANN model lies in the relative 'black-box' nature of its input/output relationships. This limitation is reduced to some extent by the *a priori* selection of input variables that have been reported in previous work to be influential on the overall function of balance control.

Linear regression models and principal component analysis could have been selected, but were passed over because they restrict variable inputs and rely on the assumption of linearity between input and output variables. Due to the ANN architecture used in this model, it may be possible to simulate the potential improvements that can arise in balance control from increasing the muscular strength inputs. Thus, targeted improvements at the muscular level could be investigated prospectively by varying its corresponding magnitude at the input layer. Furthermore, the ANN structure is adaptable, so that measures of sensory motor function could be entered into the prospective intervention set, allowing highly targeted predictions in balance improvement to be made. Further studies will be conducted to explore the use of this ANN model in predicting efficacies of different interventions of muscle strengthening in balance improvement. The ability to test the efficacy of various interventions before they are implemented is a primary strength of this model.

This study demonstrated the effectiveness of an ANN model in classifying individual cases of balance control as being within healthy ranges, or having high enough risk to be categorized as a 'faller.' It was further indicated that relative risk can be estimated for individuals with balance impairment. Findings of this study were based on a small sample of impaired subjects, compared to a relatively small sample of healthy peers. Therefore, further validation of both the sample categorization task ('healthy' vs. 'faller') and the relative risk estimation task is suggested in a larger, more diverse sample of healthy and balance-impaired elderly adults.

In conclusion, a model has been developed which demonstrates promise for accurate categorization of elderly adults and an estimation of relative risk at the individual level. Application of this model holds great potential for impacting the detection of balance impairment and further estimating the risk of falls, thereby reducing incidence of falls and enhancing quality of life in the elderly population. If level of impairment is well estimated, many elderly individuals who are at greater risk of falls might be identified prior to traumatic fall events. These individuals could then begin targeted balance-improving interventions to reduce their relative risk of falls.

ACKNOWLEDGMENTS

This study was supported in part by the National Institutes of Health (AG 022204-01; HD 042039-01A1), the Oregon Medical Research Foundation, and the International Society of Biomechanics Dissertation Matching Grant. We are grateful to Drs. Kai-Nan An and Art Farley for their critical review of the project design, and to Dr. Victor Lin for providing clinical evaluations.

REFERENCES

[1] American Academy of Orthopaedic Surgeons. Don't let a FALL be your last TRIP. 1998.

[2] Chau, T. A review of analytical techniques for gait data. Part 2: Neural network and wavelet methods. *Gait Posture* 13:102–120, 2001.

[3]Chou, L.-S., K. R. Kaufman, R. H. Brey, and L. F. Draganich. Motion of the whole body's center of mass when stepping over obstacles of different heights. *Gait Posture* 13:17–26, 2001.

[4]Chou, L. S., K. R. Kaufman, M. E. Hahn, and R. H. Brey. Mediolateral motion of the center of mass during obstacle crossing distinguishes elderly individuals with imbalance. *Gait Posture* 18:125–133, 2003.

[5]Coogler, C. E. Falls and imbalance. *Rehab. Manag.* (April/May), 53, 1992.

[6]Graafmans, W. C., M. E. Ooms, H. M. Hofstee, P. D. Bezemer, L. M. Bouter, and P. Lips. Falls in the elderly: A prospective study of risk factors and risk profiles. *Am. J. Epidemiol.* 143:1129–1136, 1996.

[7]Hahn, M. E., and L. S. Chou. Can motion of individual body segments identify dynamic instability in the elderly? *Clin. Biomech.* 18:737–744, 2003.

[8]Hahn, M. E., A. M. Farley, V. Lin, and L. S. Chou. Neural network estimation of balance control during locomotion. *J. Biomech.* 38:717–724, 2005.

[9]Hahn, M. E., H. J. Lee, and L. S. Chou. Increased muscular challenge in older adults during obstructed gait. *Gait Posture*, in press.

[10]Halfon, P., Y. Eggli, G. Van Melle, and A. Vagnair. Risk of falls for hospitalized patients: A predictive model based on routinely available data. *J. Clin. Epidemiol.* 54:1258–1266, 2001.

[11]Haykin, S. *Neural Networks: A Comprehensive Foundation.* New York: MacMillan College Publishing Co, 1994.

[12]Holzreiter, S. H., and M. E. Kohle. Assessment of gait patterns using neural networks. *J. Biomech.* 26:45–651, 1993.

[13]Jian, Y., D. A. Winter, M. G. Ishac, and L. Gilchrist, Trajectory of the body COG and COP during initiation and termination of gait. *Gait Posture* 1:9–22, 1993.

[14]Izumi, K., K. Makimoto, M. Kato, and T. Hiramatsu. Prospective study of fall risk assessment among institutionalized elderly in Japan. *Nursing Health Sci.* 4:141–147, 2002.

[15]Lafuente, R., J. M. Belda, J. Sanchez-Lacuesta, C. Soler, and J. Prat. Design and test of neural networks and statistical classifiers in computer-aided movement analysis: A case study on gait analysis. *Clin. Biomech.* 13:216–229, 1998.

[16]Levenberg, K. A method for the solution of certain nonlinear problems in least squares. *Quart. Appl. Math.* 2:164–168, 1944.

[17]Maki, B. E., P. J. Holliday, and A. K. Topper. A prospective study of postural balance and risk of falling in an ambulatory and independent elderly population. *J. Gerontol.* 49:M72–M84, 1994.

[18]Marquardt, D. W. An algorithm for least squares estimation of nonlinear parameters. *J. Soc. Ind. Appl. Math.* 11:431–441, 1963.

[19]Meglan, D. A. *Enhanced Analysis of Human Locomotion*, Ph.D. Dissertation, The Ohio State University, OH, USA, 1991.

[20]Prentice, S. D., A. E. Patla, and D. A. Stacey. Simple artificial neural network models can generate basic muscle activity patterns for human locomotion at different speeds. *Exp. Brain Res.* 123:474–480, 1998.

[21]Prentice, S. D., A. E. Patla, and D. A. Stacey. Artificial neural network model for the generation of muscle activation patterns for human locomotion. *J. Electromyo. Kinesiol.* 11:19–30, 2001.

[22]Province, M. A. The effects of exercise on falls in elderly patients: A preplanned meta-analysis of the FICSIT trials. *JAMA* 273: 1341–1347, 1995.

[23]Rumelhart, D. E., G. E. Hinton, and R. J. Williams. Learning representations by back-propagation errors. *Nature* 323:533–536, 1986.

[24]Savelberg, H. H., and A. L. de Lange. Assessment of the horizontal, fore-aft component of the ground reaction force from insole pressure patterns by using artificial neural networks. *Clin. Biomech.* 14:585–92, 1999.

[25]Sepulveda, F., D. M. Wells, and C. L. Vaughan. A neural network representation of electromyography and joint dynamics in human gait. *J. Biomech.* 26:101–109, 1993.

[26]Shumway-Cook, A., M. Baldwin, N. L. Polissar, and W. Gruber. Predicting the probability of falls in community-dwelling older adults. *Phys. Ther.* 77:812–819, 1997.

[27]Shumway-Cook, A., S. Brauer, and M. Woollacott. Predicting the probability of falls in community-dwelling older adults using the Timed Up & Go test. *Phys. Ther.* 80:896–903, 2000.

[28]Stalenhoef, P. A., J. P. M. Diedriks, J. A. Knottnerus, A. D. M. Kester, and H. F. J. M. Crebholder. A risk model for the prediction of recurrent falls in community-dwelling elderly: A prospective cohort study. *J. Clin. Epidemiol.* 55:1088–1094, 2002.

[29]Su, F.-C., and W.-L. Wu. Design and testing of a genetic algorithm neural network in the assessment of gait patterns. *Med. Eng. Phys.* 22:67–74, 2000.

[30]Swets, J. A. Measuring the accuracy of diagnostic systems. *Science* 240:1285–1293, 1988.

[31]Topper, A. K., B. E. Maki, and P. J. Holliday. Are activity-based assessments of balance and gait in the elderly predictive of risk of falling and/or type of fall? *J. Am. Geriatr. Soc.* 41:479–487, 1993.

[32]Tromp, A. M., S. M. F. Pluijm, J. H. Smit, D. J. H. Deeg, L. M. Bouter, and P. Lips. Fall-risk screening test: A prospective study on predictors for falls in community-dwelling elderly. *J. Clin. Epidemiol.* 54:837–844, 2001.

[33]Wolfson, L., R. Whipple, C. Derby, J. Judge, M. King, P. Amerman, J. Schmidt, and D. Smyers. Balance and strength training in older adults: Intervention gains and Tai Chi maintenance. *J. Am. Ger. Soc.* 44: 498–506, 1996.

[34]Woltring, H. J. A FORTRAN package for generalized, cross-validatory spline smoothing and differentiation. *Adv. Eng. Software* 8:104–113, 1986.