ORIGINAL ARTICLE

A dynamic Bayesian network for estimating the risk of falls from real gait data

German Cuaya · Angélica Muñoz-Meléndez ·
Lidia Nuñez Carrera · Eduardo F. Morales ·
Ivett Quiñones · Alberto I. Pérez · Aldo Alessi

Abstract Pathological and age-related changes may affect an individual's gait, in turn raising the risk of falls. In elderly, falls are common and may eventuate in severe injuries, long-term disabilities, and even death. Thus, there is interest in estimating the risk of falls from gait analysis. Estimation of the risk of falls requires consideration of the longitudinal evolution of different variables derived from human gait. Bayesian networks are probabilistic models which graphically express dependencies among variables. Dynamic Bayesian networks (DBNs) are a type of BN adequate for modeling the dynamics of the statistical dependencies in a set of variables. In this work, a DBN model incorporates gait derived variables to predict the risk of falls in elderly within 6 months subsequent to gait assessment. Two DBNs were developed; the first (DBN1; expert-guided) was built using gait variables identified by domain experts, whereas the second (DBN2; strictly computational) was constructed utilizing gait variables picked out by a feature selection algorithm. The effectiveness of the second model to predict falls in the 6 months following assessment is 72.22 %. These results are encouraging and supply evidence regarding the usefulness of dynamic probabilistic models in the prediction of falls from pathological gait.

Keywords Probabilistic models · Dynamic Bayesian networks · Elderly · Gait analysis · Risk of falls

G. Cuaya (✉) · A. Muñoz-Meléndez · E. F. Morales
Computer Science Department, Instituto Nacional de Astrofsica
ptica y Elctronia, Luis Enrique Erro 1,
72840 Tonantzintla, PUE, Mexico
e-mail: germancs@inaoep.mx

A. Muñoz-Meléndez
e-mail: munoz@inaoep.mx

E. F. Morales
e-mail: emorales@inaoep.mx

L. N. Carrera · I. Quiñones · A. I. Pérez · A. Alessi
National Institute of Rehabilitation, Human Motion Analysis
Laboratory, México-Xochimilco 289,
14389 Mexico, D.F., Mexico
e-mail: lnunez@inr.gob.mx

I. Quiñones
e-mail: iquinones@inr.gob.mx

A. I. Pérez
e-mail: albperez@inr.gob.mx

A. Alessi
e-mail: aalessi@inr.gob.mx

1 Introduction

According to the National Institute of Rehabilitation (INR acronym in Spanish), the risk of falls in elderly (over 65) is higher than in other groups of society. For the elderly, a fall can cause severe injuries and even death. Statistics from the INR rank falls as the first cause of death for the elderly. This research aims at getting a better understanding of the evolution and relationships of human gait parameters leading to a probabilistic model for the prediction of falls in elderly.

A gait analysis study is conducted by experts with knowledge regarding the role of specific gait variables in expressing gait degradation of people. Gait analysis depends on subjective observations made by experts, along with objective measurements of body movement, mechanics, and muscle activity [23, 36]. However, it is unclear what gait variables are ruling the risk of fall in the elderly and how their specific values modulate this risk [1, 12, 18, 35].

Computer scientists investigating gait analysis have developed solutions for individual recognition by means of image processing of videos of people walking [19, 32, 34], classification and/or estimation of the number of people walking in public spaces [4, 22, 30], and design of ambulatory technologies for gait monitoring [24, 25]. Fall detection has been attempted from people tracking [20].

Other research have analyzed parameters observed while a subject is walking to obtain information regarding his gait condition [5, 9, 28, 33]. In this sense, the expert appraises certain parameters that are believed pertinent for fall prediction.

Predicting falls from gait data has already been attempted [17, 38]. In [17], and the feasibility of a balance impairment detection model using tasks of sample categorization and falls risk estimation was tested. In this work, a first stage discriminates a random sample of a healthy elder from that of an elder with balance impairments into "healthy" or "faller" categories. Following, a second stage estimates the level of relative risk of falling for those individuals with balance impairments. The authors classify whether an elder exhibits balance impairments and assess his/her risk of falling based on the falls history of the people with balance impairments in the study. In contrast, we aim to estimate the risk of falling prospectively in a specific time interval (6 months).

In [38] the objective was (1) to determine whether stability and limb support play analogous roles in dictating slip outcome in gait-slip movement and in sit-to-stand-slip movement and (2) whether the prediction of slip could also be derived from measures of these variables during regular, unperturbed movements. The authors found that immediately before recovery step touchdown, stability, and limb support could together predict falls for gait-slip movement and sit-to-stand-slip movement. This study was not centered on elderly but instead assessed a cohort suffering unexpected slip-perturbation induced in gait or unexpected slip-perturbation during sit-to-stand movement. In contrast, we are interested in estimating the risk of falling by means of appreciating the normal degradation of human gait, that is, without perturbations, on the elderly.

In this paper we developed a model assimilating the degradation of the human gait for the assessment of risk of falls in the elderly. In particular, we propose a DBN to estimate risk of falls in elderly using spatio-temporal gait data and further explore the use of feature selection algorithms to automatically choose the gait parameters relevant for predicting falls.

The model should predict falls of the elderly at different time intervals in the future. Thus, the model is able to represent and evaluate the dynamics of gait parameters estimating the risk of falling based upon observed gait changes. Observations of gait parameters may be incomplete, inaccurate and at times contradictory, and therefore, it is critical for the model to handle the uncertainty of the process.

BNs are graphical representations of dependencies for probabilistic reasoning. A BN is a graph in which vertices represent random variables and edges represent direct-dependency relationships among variables. BNs have been extensively used in artificial intelligence (AI) to model probabilistic relationships among interacting variables.

DBNs are an extension of BNs that can represent the evolution of variables course over time and can handle uncertainty in the data. DBNs have been applied in medical applications such as diagnosis of pneumonia associated with the use of medical devices [10], as well as maintenance diagnostic of industrial equipment [3, 15].

The remainder of the paper is organized as follows: In Sect. 2, we describe the methods used in this work. In Sect. 3, we depict the experiments and results obtained with the DBNs, and finally in Sect. 4, we provide conclusions and discuss future research directions.

2 Methods

2.1 Gait analysis

Gait analysis is based on specialized equipment for 2D or 3D motion capture sensed using motion tracking multi-camera systems, ground surface pressure detectors, and force platforms, among others. Data from walking for this study was acquired using a GaitRite system (CIR Systems Inc, Havertown, PA, USA), a system for monitoring spatio-temporal patterns of human locomotion (for a detailed description of the GaitRite system, see [6]).

Researchers of the Human Motion Analysis Laboratory at the INR in Mexico City longitudinally acquired gait data from 18 women aged 70 ± 10 years diagnosed with osteoporosis. Osteoporosis is a condition resulting in brittle bones which may lead to unsteady gait and falls. The participants walked on a 3.0-m-long GaitRite walkway. The GaitRite system measures the temporal and spatial gait parameters declared in Table 1.

Gait data were obtained from each subject every 6 months during 3 years. In addition, occurrence of falls was also logged for the duration of the trial. The number of assessment sessions differed among patients because subjects occasionally failed to comply with appointments and there were subjects who dropped out from testing.

DBNs models were built utilizing data from 18 patients distributed as follows: 16 patients that fell and 2 patients that did not fall during the study. A total of 66 records were used: 18 records—0 months, 17 records—6 months, 17 records—12 months, and 14 records—18 months.

Table 1 Spatio-temporal gait parameters obtained with the GaitRite system

No.	Parameter (unit of measure)
1	Walking distance (cm)
2	Ambulation time (s)
3	Walking velocity (cm/s)
4	Mean normalized velocity
5	Number of steps
6	Cadence
7	Left and right step time differential (s)
8	Left and right step length differential (cm)
9	Left and right cycle time differential (s)
10	Left step time (s)
11	Right step time (s)
12	Mean step time (s)
13	Mean step cycle time (s)
14	Left step length (cm)
15	Right step length (cm)
16	Left stride length (cm)
17	Right stride length (cm)
18	Base of support left step (cm)
19	Base of support right step (cm)
20	Percentage of single support left (%GC)
21	Percentage of single support right (%GC)
22	Percentage of double support left (%GC)
23	Percentage of double support right (%GC)
24	Left swing percentage (%GC)
25	Right swing percentage (%GC)
26	Left stance percentage (%GC)
27	Right stance percentage (%GC)
28	Step length/extremity length ratio left
29	Step Length/extremity length ratio right
30	Left toe in/out angle (°)
31	Right toe in/out angle (°)

2.2 Relevant variables

BNs and DBNs, being rooted in classical statistics, benefit from a large number of observations and few variables from those observations to identify significant probabilistic relationships. A number of methods have been proposed to alleviate problems with high-dimensional data when only a small number of instances are available, such as for instance oversampling and reduction of dimensionality. Reduction of dimensionality can be achieved by feature selection strategies dropping a subset of variables considered to be less informative. In this work, we reduce the dimensionality of our problem in two different ways: first, by keeping only those gait variables picked by an expert, and second, by applying a well-known feature selection algorithm, namely forward sequential selection (FSS), to

Table 2 Relevant variables and their abbreviated names

CAD	Cadence
Variables selected by gait analysis experts	
LPI	Left step length (cm)
LPD	Right step length (cm)
BSI	Base of support left step (cm)
LPCI	Left stride length (cm)
LPCD	Right stride length (cm)
BSD	Base of support right step (cm)
Variables selected by FSS	
TPD	Right step time (s)
BSI	Base of support left step (cm)
BSD	Base of support right step (cm)
SDI	Left double support (%GC)
RPED	Right step/extremity ratio
TIOI	Left toe in/out angle (°)
TIOD	Right toe in/out angle (°)

filter the gait variables with low predictive value from the set of gait parameters acquired with the GaitRite system.

Accordingly, we have built two DBNs for prediction of fall from two different reduced sets of variables of human gait. The first model is expert-guided and capitalizes on the set of variables identified by gait analysis experts from the Human Motion Analysis Laboratory of the INR as pertinent for figuring out the risk of falling. These variables are shown in the top half of Table 2.

The second model is strictly computational and is founded on the set of relevant variables as automatically selected by the FSS algorithm [21]. The FSS algorithm was run taking into consideration all the records from the 18 patients, that is, all 66 records added up in Table 2. In this case, the FSS algorithm achieves a reduction of the original 31 variables in Table 1 to 7 relevant variables. These variables are shown in the bottom half of Table 2.

2.2.1 Forward sequential selection

Let X be a matrix whose rows correspond to points (or observations) and columns correspond to features (or predictor variables). And let Y be a column vector of response values or class labels for each observation in X. FSS starts with an empty feature or variable set. It then creates candidate feature subsets by adding each of the features not yet selected. For each candidate feature subset, FSS performs a tenfold cross-validation, by repeatedly calling an evaluation function, called *FUN*, that defines the *CRITERION* that FSS uses to select features and determines when to stop. See Eq. (1)

$$CRITERION = FUN(X_{TRAIN}, Y_{TRAIN}, X_{TEST}, Y_{TEST}) \qquad (1)$$

where $X_{TRAIN}, X_{TEST}, Y_{TRAIN}, Y_{TEST}$ are a partition in training and test subsets of X and Y, respectively. *FUN* uses X_{TRAIN} and Y_{TRAIN} to train or fit a model, then it predicts values for X_{TEST} using that model, and finally, returns measurements of distance or loss of those predicted values from Y_{TEST}. We applied ten-time cross-validation to obtain representative results for each candidate feature set. FSS sums the values returned by *FUN* across all test sets and divides that sum by the total number of observations. Then it uses that mean value to evaluate each candidate feature subset. Commonly used loss measures for *FUN* include the sum of squared errors for regression models and the number of misclassified observations for classification models.

Given the mean loss values for each candidate feature subset, FSS chooses the one that minimizes the mean value. This process continues until adding more features does not decrease the criterion value.

In our case, the *GLMFIT* function of MATLAB (MathWorks, MA, USA) was used to retrieve the *CRITERION* for FSS. *GLMFIT* fits a generalized linear model (GLM) using the predictor matrix X, response Y, and a distribution *DISTR*. GLM is a flexible generalization of ordinary linear regression for response variables that are not normally distributed. We use the values generated by DISTR, X_{TEST} and Y_{TEST}. The result is a vector of coefficient estimates. Acceptable values for DISTR are normal, binomial, Poisson, gamma, and inverse Gaussian. We set a binomial distribution. In most cases, Y is an column vector of observed responses, and it is common to use the binomial distribution when Y is a binary vector indicating success or failure for each observation.

2.3 Bayesian networks

BNs have become a popular representation in AI for encoding uncertainty in knowledge [7, 16]. Extensions of the basic BNs scheme further expand their capabilities for probabilistic representation and reasoning [3, 15]. BNs are directed acyclic graphs that represent the conditional independence relationships for joint probability distribution of a set of random variables. These models belong to the family of models referred to as probabilistic graphical models and have been successfully applied in many domains, as for instance, fault diagnosis due to their capability for handling uncertainty in data. BNs offer advantages over alternative AI approaches such as neural networks [27] or fuzzy logic [31] such as, for instance, the capability of representing dynamic and uncertain data.

BN nodes or vertices represent variables of the process and directed edges represent both relationships and
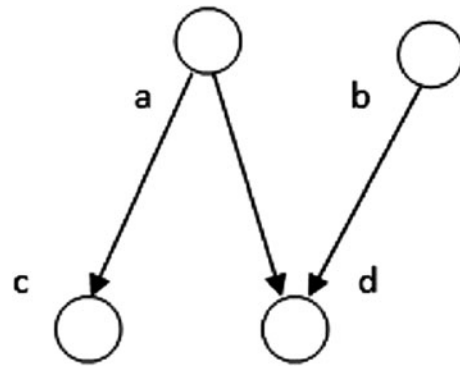

Fig. 1 Bayes network example

conditional independence-relationship among variables. A node is independent of other non-descendant nodes, given the values of its parents. Figure 1 illustrates a Bayesian network where node d is independent of node c, given node a.

The definition of a Bayesian network requires the estimation of conditional probability table (CPT) for each node. The total joint distribution for a BN is given by the following product:

$$P(X_1, \ldots, X_n) = \prod_{i=1}^{n} P(X_i | Pa(X_i)) \qquad (2)$$

where $X_1, \ldots, X_n$ is the set of n variables comprising the network and $Pa(X_i)$ represents the parent nodes of node X_i. There are several exact and approximate algorithms for inference of the posterior distribution of one or more observed variables in a BN. A BN can be manually defined or automatically induced from data.

2.4 Dynamic Bayesian networks

An ordinary BN is a static model representing a joint probability distribution at a fixed point or time interval. Instead, a DBN can model the evolution of the probabilistic dependencies within a system over time. In particular, variables are represented at multiple time points within the same network structure.

A number of ways to represent the passage of time in BNs exist, but perhaps the most popular is the method proposed by Dean [13]. In this approach, time is modelled discretely as in a discrete Markov chain. Each variable, X_t, has a time index subscript indicating the time slice to which it belongs. Each time slice in the DBN is an ordinary BN holding the static dependencies among the variables for a particular time interval. Additional temporal dependencies are represented in a DBN by edges between time slices. In many cases, it is only necessary to consider first-order time dependencies, in which case a two-slice network is sufficient to render all relationships. The two-slice
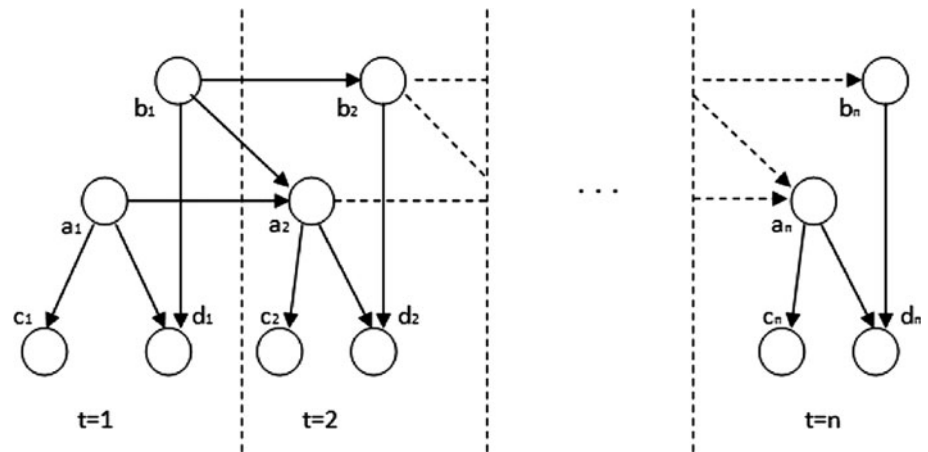
network can be regarded as being "unrolled" over the number of time slices required to solve the problem at hand. However, the actual number of slices on which inference is performed depends, of course, on the particular domain.

The definition of the conditional independence relationship is the same for both dynamic and static networks. An example of a DBN, with n time steps, is illustrated in Fig. 2. In this figure a_2 is independent of c_1 and d_1, given a_1 and b_1. The conditional probability relationship between time steps, such as $P(a_2|a_1, b_1)$, are referred to as the transition model. The task of a DBN is to make inferences about the posterior distribution, or belief state, of non-observable variables from previous time steps and from current observations.

DBNs are appropriate for monitoring and predicting values of random variables and are capable of representing the system state at any time.

2.5 Construction of DBN to gait model

For constructing a DBN-based gait model, the information from all 18 subjects across the four assessment sessions was used: first record (at study onset, or 0 months), second record (6 months from onset), third record (1 year from onset), and fourth record (18 months from onset), for a total of 66 records (as indicated in Sect. 2). Each record is originally a 31D vector which has been reduced in dimensionality as afore described by either expert selection or automatic feature selection. Each record is an observation, which can be labelled in a group with registered falls or "fallers", or in a group without registered falls or "non fallers". In addition, sets of up to four observations or records represent longitudinal assessment of a single patient. We used a portion of the data to build the model and the rest for evaluation as explained in Sect. 3.

In order to build a DBN, a static model of the phenomenon is generated. This requires learning the

relationships among the nodes of the static model (vertical structure). Then a copy of the model is created for each instant within the time range of interest. Finally, existing relationships among the nodes of two consecutive static networks are established (horizontal structure). The connections in the static model represent dependencies between nodes at a particular time-slot and the connections between consecutive static networks represent the dependencies between consecutive observations or time-slots.

Before training the DBN, the next step is qualifying how the nodes are related among them and also how their relationships change over time. To do so, it is necessary to characterize the possible values of the nodes or variables of the model.

Relevant variables express related but essentially different information of the observed phenomenon, i.e. the human gait, measured in different measurement units and scale. Hence, we expressed the values of the nodes of the model in terms of their implicit change. In this sense, values of relevant variables are binarized so to reflect the trend variations of a variable between two consecutive records, through the function *Change* defined as follows:

$$ChangePX = \begin{cases} 1 \text{ if } |PX_{rec2} - PX_{rec1}| > PX_{rec1} \times DetFac \\ 0 \text{ otherwise} \end{cases}$$

(3)

where, PX is a certain parameter of gait analysis PX_{rec1} is the value of the parameter in the antecedent record, PX_{rec2} is the value of the parameter in the subsequent record, and $DetFac$ is an experimental threshold value allowing for the "normal" deterioration factor of human gait.

In this paper, we set $DetFac = 0.05$, meaning that a difference greater or equal to 5 % between a given gait parameter at two consecutive times is necessary to consider it as an abnormal deterioration.

Previously, up to 8 % difference between the same elderly gait parameter has been considered normal [29, 37]. This study proposed a 5 % threshold since this value has

been widely accepted to represent clear deviations from normal deterioration. Also we wanted to obtain a 95 % of statistical confidence between the differences in gait values [2, 8, 14].

From Eq. 3, the values that any node of the model can take and the corresponding meanings are, 0 = no change or normal deterioration, 1 = change beyond normal deterioration.

Once the nodes of the model and their possible values are defined, the next step establishes how these nodes are proxy for the risk of falling. The K2-learning algorithm [11] was applied to learn the structure of our DBNs. This algorithm is available in MATLAB.

The K2-learning algorithm is a Bayesian algorithm that finds probabilistic dependencies among the variables of a dataset. Basically, the algorithm searches for a Bayesian network that has a high posterior probability given the dataset and outputs the mentioned structure and its probability. In other words, this algorithm learns the vertices or relationships among variables from data survey.

At this point, we have the nodes of the model, the values of these nodes, as well as both, horizontal and vertical, relationships among these nodes. To complete the DBN models, we still have to compute the CPTs: listing the probabilities of the values of a child node given the values of its parents. For the computation of the CPTs, an algorithm that counts the combination of each value for each node and its parents and estimates the conditional probabilities from there has been implemented in MATLAB.

This algorithm is first run over the whole set of 66 observations to establish the CPTs at any given time t. The result is represented in Fig. 3a. This vertical structure was replicated twice, i.e. representing a given time t and its subsequent time $t + 1$, because the goal is to forecast of risk of falling [within the first 6-month period (that is at time t)] that will be referred to for the rest of the paper as *imminent* and the risk of falling [within the second 6-month period (at time $t + 1$)] that will be referred to for the rest of the paper as *6 months*. Then, the algorithm to learn the CPTs is again applied to relate variables relevant at time t with the variables relevant at time $t + 1$, corresponding to consecutive assessments of the patients, thus sketching the dynamics of variables over time as illustrated in Fig. 3b. For this second round, 55 observations of the available 66 were employed. Those not used were the cases in which the subsequent assessment at $t + 1$ was missing for a given antecedent assessment at time t. In other words, registers lacking subsequent information were discarded.

Finally, Hugin (Hugin Experts, Aalborg, Denmark) [26], a specialized software package for designing and building BNs, was used for testing the models. The structures of both DBNs, as well as the CPTs and the dataset, were provided to Hugin.
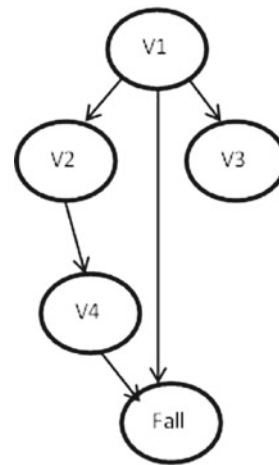
(a) Vertical structure.
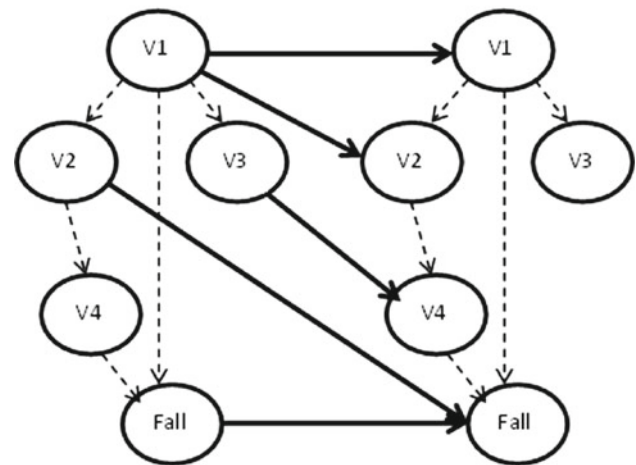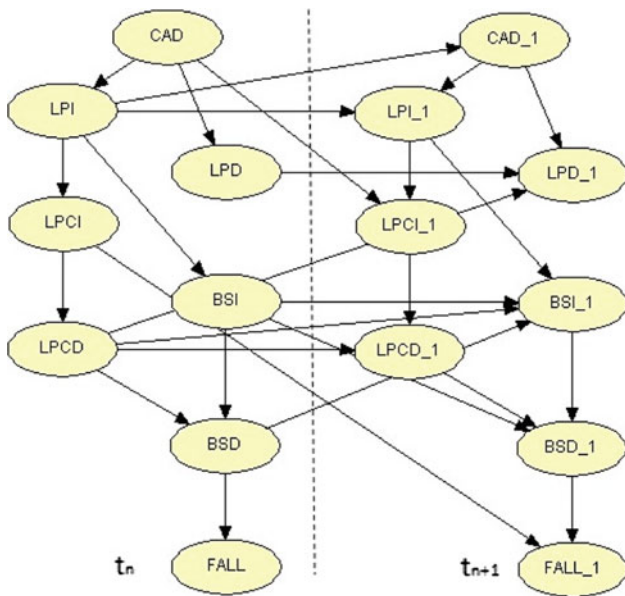

(b) Horizontal structure.

Fig. 3 **a** Probabilistic dependency of variables V_i at times t learned with K2. **b** Probabilistic dependency of variables V_i between two times, learned with K2

The aforedescribed procedure was repeated for the set of relevant variables picked by the gait analysis experts and the set of relevant variables picked by the FSS algorithm. In both cases, information of falls recorded from patients during the study was incorporated in addition to data of the GaitRite system. In this regard, an additional node called fall was added to the models. This new node fall can take one of two values, 0—*no fall*; and 1—*fall*.
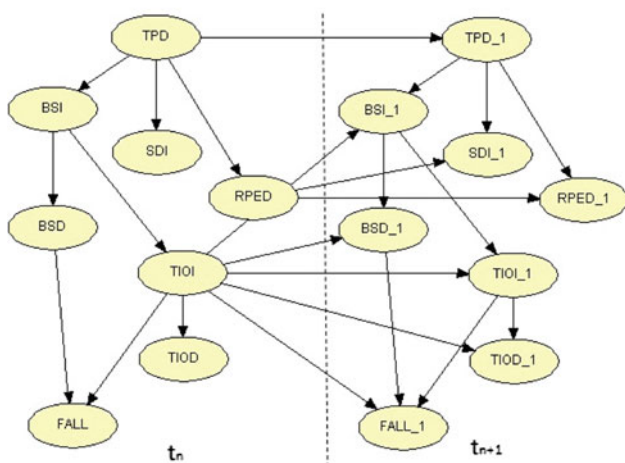
The evaluation of the two DBN models is described in the next section.

3 Results

Two conceptual models, one based on the expert's chosen relevant variables, and another based on the relevant variables automatically picked out, were evaluated with real data from elderly. The models achieved an average

(a) DBN based on expert's variables.

(b) DBN based on automatic selected variables.

Fig. 4 a DBN based on gait parameters selected by experts, **b** DBN based on gait parameters selected by the FSS algorithm. See Table 2 for the definition of abbreviations

Fig. 5 Comparison of the prognosis results of the two models

power of the model is considered to be the average of all the model instances.

Thereby, for each of the two conceptual models (the expert selection based and the FSS selection based), 18 different model instances were built using data from 17 patients and tested on the remaining patient on the most probable value of the node fall, 0 = *no fall* or 1 = *fall*. Note that here we exclude all observations corresponding to the longitudinal assessments for a given subject at a time, rather than dropping every single observation separately. Figure 4 shows the DBNs built by Hugin using human gait data.

The effectiveness of the forecasting of *imminent* falls and *6 months* risk of falling is defined as the ratio of the number of falls predicted by the model divided by the number of falls that actually occurred in the corresponding time interval.

The effectiveness of the forecasting of *imminent* falls and *6 months* risk of falling were 72.22 % in both cases for the DBN model built upon the variables selected by the domain experts, and 72.22 and 66.66 % for the DBN built upon the variables that were automatically selected with FSS algorithm. These results are summarized in Fig. 5.

The effectiveness of both models is presented without confidence intervals due to the small number of samples. Tables 3 and 4 detail he confusion matrices for the two conceptual models at the two time intervals considered; *imminent* and *6 months*, respectively. The confusion matrices presented have been built from the agglomerate results during the LOOCV.

Finally, the *sensitivity* and *specificity* of the conceptual models have been estimated. Sensitivity (also called recall rate in some fields) measures the proportion of actual positives which are correctly identified over the total real positives (i.e. the percentage of elders who are correctly

precision of 70 % in predicting both *imminent* falls and *6 months* risk of falling.

Both models were evaluated using leave-one-out cross validation (LOOCV). The LOOCV is a technique that facilitates estimation of performance of a predictive model when the available number of samples is low, as it occurs in this case where data from 18 patients only are available. During LOOCV, observations are left out one at a time and a model instance is trained with the remaining observations and tested on the observation left out. This procedure is repeated for all of the observations and the predictive
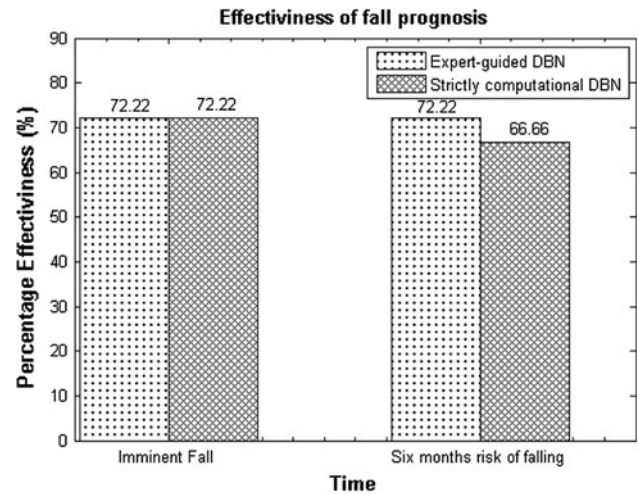
Table 3 Confusion matrix to prognostic of *imminent* falls

	Negative	Positive
Predicted using information get with computational techniques		
Actual		
Negative	31	2
Positive	13	9
Predicted with information get of experts		
Actual		
Negative	30	3
Positive	12	10

Table 4 Confusion matrix to prognostic of *6 months* risk of fall

	Negative	Positive
Predicted using information get with computational techniques		
Actual		
Negative	34	4
Positive	14	3
Predicted with information get of experts		
Actual		
Negative	36	2
Positive	13	4

Table 5 Sensitivity and specificity measures

	Imminent fall		Six months risk of fall	
	FSS	INR	FSS	INR
Sensitivity	0.952	0.930	0.902	0.952
Specificity	0.704	0.714	0.708	0.734

identified as fallers). Specificity measures the proportion of negatives which are correctly identified over the total real negatives (i.e. the percentage of elders who are correctly identified as not fallers). The sensitivity and specificity of both conceptual models are presented in Table 5.

4 Discussion

This research aims at generating a model for prediction of fall in the elderly that capitalizes on quantifying probabilistic dependencies among gait variables. Our objective has been to develop an intelligible model for non-experts in probabilistic reasoning that afford them with complementary information. Distinctly from related work [17, 38], our model incorporates people's gait degradation as captured by assessment instruments and learns the probabilistic

relationships among the variables recorded by such instruments to forecast the risk of fall of an elder.

A strictly computational model based on selected information extracted from gait assessments was compared with an analogous model based on information drawn from the gait assessments by experts. For estimating the risk of fall, the computational feature selection method picked out variables of human gait that were not considered relevant by the experts. Nevertheless, both conceptual models yielded comparable performances.

Although the model based upon automatic feature selection exhibits lower performance for the *6 months* prediction, this model emerges from a very limited dataset and still afforded competitive performance. In this sense, the experts base his/her selection on an experience certainly spanning more observations than those in the dataset. In general, the larger the training dataset the more reliable model can be expected.

This work demonstrates the feasibility of employing probabilistic models such as DBNs to estimate the probability of a fall. We have presented two probabilistic models of fall risk assessment that were developed from actual records of spatio-temporal gait parameters. To our knowledge, these are the first probabilistic models of fall risk assessment that exploit relationships among relevant variables. Since a gait model is now available, it can be applied by human gait experts and clinicians to obtain additional elements to enrich their decisions about treatment and therapies to be prescribed to elders with different degrees of gait impairment.

Note that both, the strictly computational model and the expert-guided model achieve comparable performances as measured by sensitivity and specificity outcomes. This can be indicative that there is a further reduced set of variables of human gait relevant for predicting falls, perhaps in the intersection of the subsets of variables considered by both models.

We are currently bettering the models by enlarging the dataset. Models for predicting falls within 12 and 18 months will be developed as we grow our cohort size and expand follow-up times. We will also incorporate information from people with normal gait to compare their assessments with those from individuals with pathological gait.

References

1. Allan LM., Ballard CG, Rowan EN, Kenny RA (2009) Incidence and prediction of falls in dementia: a prospective study in older people. Public Libr Sci 4(5):e5521

2. Armitage P, Berry G, Matthews JNS (2001) Statistical methods in medical research, 4th edn. Wiley-Blackwell, New York, pp 88–89

3. Arroyo-Figueroa G, Sucar LE (2005) Temporal bayesian network of events for diagnosis and prediction in dynamic domains. Appl Intell 23(2):77-86

4. BenAbdelkader C, Cutler R, Davis LS (2006) View-invariant estimation of height and stride for gait recognition. Biometric Authentication 2359:155-167

5. Bhatt T, Espy D, Yang F, Pai YC (2011) Dynamic gait stability, clinical correlates, and prognosis of falls among community-dwelling older adults. Arch Phys Med Rehabil 92(5):799–805

6. Bilney B, Morris M, Webster K (2003) Concurrent related validity of the gaitrite walkway system for quantification of the spatial and temporal parameters of gait. Gait Posture 17(1):68–74

7. Borsuk ME, Stow CA, Reckhow KH (2004) A Bayesian network of eutrophication models for synthesis, prediction, and uncertainty analysis. Ecol Model 173(2–3):219 - 239

8. Box GE, Hunter WG, Hunter JS (2005) Statistics for experimenters: design, innovation, and discovery, 2nd edn. Wiley-Interscience, New York, pp 19–43

9. Callisaya ML, Blizzard L, Schmidt MD, Martin KL, McGinley JL, Sanders LM, Srikanth VK (2011) Gait, gait variability and the risk of multiple incident falls in older people: a population-based study. Age Ageing 40(4):481–487

10. Charitos T, Gaag L, Visscher S, Schurink K, Lucas P (2009) A dynamic bayesian network for diagnosing ventilator-associated pneumonia in ICU patients. Expert Syst Appl 36(2):1249–1258

11. Cooper GF, Herskovits E (1992) A bayesian method for the induction of probabilistic networks from data. Mach Learn 9(4):309–347

12. Dabiri F, Vahdatpour A, Noshadi H, Hagopian H, Sarrafzadeh M (2008) Ubiquitous personal assistive system for neuropathy. In: Proceedings of the 2nd international workshop on systems and networking support for health care and assisted living environments. Breckenridge, pp 171–176

13. Dean T, Kanazawa K (1989) Persistence and probabilistic inference. IEEE Trans Syst Man Cybern 19(3):574–585

14. Devore JL (1999) Probability and statistics for engineering and the sciences, 5th edn. Duxbury Pr, pp 155–182

15. Galán SF, Arroyo-Figueroa G, Dez FJ, Sucar LE (2007) Comparison of two types of event bayesian networks: a case study. Appl Artif Intell 21(3):185–209

16. Gu T, Pung HK, Zhang DQ, Pung HK, Zhang DQ (2004) A bayesian approach for dealing with uncertain contexts. In: Proceedings of the 2nd international conference on pervasive computing, vol 176, Vienna

17. Hahn ME, Chou LS (2005) A model for detecting balance impairment and estimating falls risk in the elderly. Ann Biomed Eng 33:811–820

18. Haworth JM (2008) Gait, aging and dementia. Rev Clin Gerontol 18(1):39–52

19. Kale A, Sundaresan A, Rajagopalan AN, Cuntoor NP, Roy-Chowdhury AK, Kruger V, Chellappa R (2004) Identification of humans using gait. Image Process 13:1163–1173

20. Kangasa M, Konttilaab A, Lindgren P, Winbladad I, Jamsa T (2008) Comparison of low-complexity fall detection algorithms for body attached accelerometers. Gait Posture 28(2):285–291

21. Kohavi JG (1997) Wrappers for feature subset selection. Artif Intell 97(1–2):272-324

22. Lam T, Lee R, Zhang D (2007) Human gait recognition by the fusion of motion and static spatio-temporal templates. Pattern Recognit 40(9):2563–2573

23. Lee HJ, Chou LS (2006) Detection of gait instability using the center of mass and center of pressure inclination angles. Arch Phys Med Rehabil 87(4):569–575

24. Lopez-Meyer P, Fulk GD, Sazonov ES (2011) Automatic detection of temporal gait parameters in poststroke individuals. Inf Technol Biomed 15(4):594–601

25. López-Nava IH, Muñoz-Meléndez A (2010) Towards ubiquitous acquisition and processing of gait parameters. In: Proceedings of the 9th Mexican international conference on artificial intelligence, Pachuca, pp 410–421

26. Madsen AL, Lang M, Kjrulff UB, Jensen F (2003) The Hugin tool for learning bayesian networks. In: Proceedings of the 7th European conference on symbolic and quantitative approaches to reasoning with uncertainty, Aalborg, pp 594–605

27. Madu EO, Stalbovskaya V, Hamadicharef B, Ifeachor EC, Van S, Timmerman D (2005) Preoperative ovarian cancer diagnosis using neuro-fuzzy approach. In: Proceedings of the European conference on emergent aspects in clinical data analysis, Italy, pp 1–8

28. Maki BE (1997) Gait changes in older adults: predictors of falls or indicators of fear. J Am Geriatr Soc 45(3):313–320

29. Menz HB, Lord SR, Fitzpatrick RC (2003) Age-related differences in walking stability. Age Ageing 32(2):137–142

30. Meyer D (1997) Human gait classification based on hidden markov models. In: Proceedings of the 3D image analysis and synthesis conference, Erlangen, pp 139–146

31. Papageorgiou EI, Papandrianos NI, Karagianni G, Kyriazopoulos G, Sfyras D (2009) Fuzzy cognitive map based approach for assessing pulmonary infections. In: Proceedings of the 18th international symposium on methodologies for intelligent systems, Prague, pp 109–118

32. Phillips PJ, Sarkar S, Robledo I, Grother P, Bowyer K (2002) The gait identification challenge problem: data sets and baseline algorithm. In: Proceedings of the 16th international conference on pattern recognition, Qubec, pp 385–388

33. Rogers ME, Rogers NL, Takeshima N, Islam MM (2003) Methods to assess and improve the physical parameters associated with fall risk in older adults. Prev Med 36(3):255–264

34. Romero-Moreno M, Martínez-Trinidad JF, Carrasco-Ochoa JA (2008) Gait recognition based on silhouette, contour and classifier ensembles. In: Proceedings of the 13th Iberoamerican congress on pattern recognition. Havana, pp 527–534

35. Saboune J, Charpillet F (2005) Markerless human motion capture for gait analysis. In: Proceedings of the 3rd European medical and biological engineering conference, Prague

36. Whittle MW (2001) Gait analysis: an introduction, 3rd edn. Butterworth-Heinemann, pp 127–160

37. Winter DA, Patla AE, Frank JS, Walt SE (1990) Biomechanical pattern changes in the fit and healthy elderly. Phys Therapy 70:340–347

38. Yang F, Bhatt T, Pai YC (2009) Role of stability and limb support in recovery against a fall following a novel slip induced in different daily activities. Biomechanics 42:1903–1908