Automatic gait classification in children with Cerebral Palsy: a Bayesian approach.

Leen Van Gestel¹, Mike Schwartz²,³, Tinne De Laet¹, Herman Bruyninckx², Guy Molenbergs²,³, Anja Van Campenhout²,³, Enrico Di Lello², Erwin Aertebiën², Hans Wambacq², Paul De Cock², Kaat Desloovere²,³

¹Dep. of Rehabilitation Sciences, Fac. of Kinesiology and Rehabilitation Sciences, KULeuven, Belgium; ²Gillette Children’s Specialty Healthcare, James R. Gage Center for Gait and Motion Analysis, St. Paul, MN, USA; ³Univ. of Minnesota, Dep. of Orthopaedic Surgery and Biomedical Engineering, Minneapolis, MN, USA; ⁴Div. of Production Engineering, Machine Design and Automation, Dep. of Mechanical Engineering, KU Leuven, Belgium; ⁵Clinical Motion Analysis Laboratory, CERM, University Hospital Leuven, KULeuven, Belgium; ⁶Dep. of Paediatric Orthopaedics, University Hospital Pellenberg, Belgium; ⁷Musculoskeletal Sciences, KULeuven, Belgium; ⁸Department of Pediatrics, University Hospital Leuven, Belgium

SUMMARY
Three dimensional gait analysis (3DGA) is a vital element in the multidisciplinary treatment of children with cerebral palsy. As 3DGA generates an extensive amount of data, preprocessing steps are needed to facilitate the clinical interpretation and decision-making process. Pre-processing often includes classification of gait into distinct gait patterns, using either qualitative or quantitative approaches. While the former can be subjective and have limited repeatability, the latter often end up with clinically irrelevant gait patterns [1]. Therefore, new strategies for efficient data-analysis need to be explored. This study introduces Bayesian networks (BN) for preprocessing 3DGA data. The developed BN integrated prior knowledge (clinical expertise) to automatically classify the movement patterns of the ankle and knee in the sagittal plane. A group of 139 patients with CP were recruited and divided into a training set (n= 80% of all patients) and a validation set (n= 20% of all patients). The proposed BN reached accuracies between 73% - 100% for automatic gait classification and was found to be a successful tool for the classification of ankle and knee patterns.

INTRODUCTION
A central element in the treatment of children with Cerebral Palsy (CP) is the detailed three dimensional gait analysis (3DGA). Because 3DGA generates a vast and complex dataset, an efficient data reduction scheme is imperative to facilitate clinical interpretation and decision making. Classification of gait into distinct gait patterns proves to be an efficient method for preprocessing 3DGA data. As reviewed by Dobson, qualitative and quantitative strategies for classification construction can be distinguished[1]. Qualitative strategies distinguish clinically relevant groups[1] but often are limited by their subjective nature[1] and limited repeatability[2]. On the other hand, quantitative strategies provide a systematic and structured approach to analyze gait data[1]. Nevertheless, such approaches often generate clinically irrelevant groups and have not proven to be useful additions to the clinical interpretation process[1]. Furthermore, as CP gait is often a mix of several gait patterns and both qualitative and quantitative methods tend to force all gait into one single gait pattern category, finding the adequate method for the classification of 3DGA in patients with CP remains a challenge. Recent research indicates that Bayesian Networks (BN) are a promising tool for the analysis of the 3DGA datasets. A BN identifies and trains relationships within complex heterogeneous datasets in the face of incomplete and/or imperfect data[3]. Automatic classification of 3D movement patterns includes modeling the complex relations between the 3DGA curves and variables, and the observed movement patterns in the different joints. BNs have many advantages when compared to other quantitative techniques, one of them being the incorporation of prior knowledge such as clinical expert knowledge in the identification and training of relations. Furthermore, integration of clinical knowledge into a BN for gait classification ensures the classification of movement patterns into clinically relevant patterns. Finally, BNs do not classify gait into just one movement category but assign probability scores for belonging to each category (movement pattern). Because of these strengths, BNs may be a promising tool for analysis of the 3DGA dataset.

METHODS
This study developed a BN for the automatic classification of CP gait into clinically relevant movement patterns. (Fig. 1) The movement patterns, and their determining gait features, were defined by a multidisciplinary team of experts in gait analysis (orthopaedic surgeons, kinesiologists and physical therapists). A gait feature is a clinically relevant (kinematic or kinetic) parameter extracted from the gait cycle, such as the maximal or minimal angle of a joint. Several iterations were required to define the final set of movement patterns and their corresponding gait features. These movement patterns are described through qualitative and quantitative rules. Qualitative rules define which gait features determine a certain gait pattern and their required feature state(s) (e.g. feature state: ‘decreased knee flexion in swing’ was required for pattern: ‘stiff knee’). The corresponding quantitative rules determine the threshold values of the different states of a feature e.g. ‘normal’ [57.7° -> 67.6°] and ‘decreased’ [5.0° -> 57.7°] knee flexion angle in swing. The developed BN focused on the automatic classification of the movement patterns of the knee and ankle in the sagittal plane during stance and swing phase. It consisted of nodes (gait variables), arcs (which link variables to each other) and underlying Condition Probability Tables (CPT) (describing how variables are related). The
database of the Pellenberg Movement Analysis Laboratory (>4000 patients) was searched for gait analyses to train and test the BN. Inclusion criteria were: (a) predominantly spastic type of CP, (b) age between 5 and 12 years, (c) no history of lower limb surgery, (d) no lower limb Botulinum Toxin A treatment within 6 months prior to the 3DGA, (e) ability to walk independently, (f) 3DGA data acquired with accurate marker placement and definition of the knee flexion/extension alignment by the knee alignment device[4] (Vicon, Oxford Metrics, Oxford, UK), and (g) at least two gait trials with full kinematics and kinetics available. One member of the multidisciplinary team classified the ankle and knee movement patterns of all patients based on the qualitative kinematics and kinetics available. One member of the multidisciplinary team classified the ankle and knee movement patterns of all patients based on the qualitative rules. Similar to existing classifications, this expert was instructed to classify each patient (n= 139; n= 814 trials) into a single movement pattern. This classification was used as a reference to evaluate the performance of the developed BN. The required gait features were automatically extracted from the continuous gait waveforms.

RESULTS AND DISCUSSION
There were 139 children who met the inclusion criteria (n=58 hemiplegia, n=80 diplegia, n=1 triplegia). Of the selected patients, 103 were classified as a GMFCS level I and 35 as GMFCS level II (1 missing data point). The average age was 8 yrs 10 months ± 2 yrs 2 months. Netica® was used as a software tool to create the proposed BN[5]. Figure 1 shows the BN that was developed for this study based on prior knowledge. The top layer of nodes represent the clinically relevant movement patterns for the ankle and knee in the stance and swing phase of gait. The nodes below represent the determining gait features for these movement patterns. Arcs connect each movement pattern node (e.g.: ‘Ankle pattern in swing’) to its own set of gait features (e.g. ‘maximal dorsiflexion angle in swing’ and the ‘ankle angle at initial contact’). To train the parameters of the BN, patient cases (3DGA) containing data for every node in the BN were needed. Of all included patients, 80% (n=111 patients or 651 trials) were randomly selected for the training phase. All gait features and classification data of these patients were used to train the BN. The gait data of the remaining 28 patients (163 trials) was used as a validation dataset for testing the accuracy of the BN. Only the extracted gait features of the validation patients were used as an input for the BN. The network then used probabilistic inference to classify the gait patterns. For every case, the predicted movement pattern was compared to the reference classification (done by a member of the multidisciplinary team) to determine the network’s accuracy expressed as a performance rate per subset of movement patterns. As mentioned before, one of the advantages of a BN is that each gait trial is not forced to fit into just one movement class. For each gait trial, a probability for belonging to each movement class is calculated, enabling the classification of mixed movement patterns. In a second phase, a performance rate was calculated separately for trials with a distinct and a mixed movement pattern. Every case (3DGA trial) was checked to see whether the predicted probabilities were spread out evenly over two or more movement patterns (‘mixed’ movement pattern) or if there was one dominant predicted movement pattern (probability >65%, ‘distinct’ movement pattern). To evaluate the performance for the mixed gait trials, the highest and second highest probable movement patterns were compared to the reference classification. The highest performance rate for the distinct movement patterns was achieved for the classification of the ankle pattern in stance with 91.3% of all patients classified correctly. The movement patterns for the ankle in swing and the knee in stance and swing were classified with an accuracy of 86.7%, 81.9%, 89.5% respectively. For the mixed gait patterns, the reference movement classification was almost always reflected in one of the two most probable movement patterns. The mixed ankle and knee movement patterns in swing achieved a classification accuracy of 100% while their classification in stance achieved an accuracy of 92.9% and 86% respectively.

CONCLUSIONS
The focus of this study was the introduction of a new tool for automatic gait classification. Bayesian networks were selected because of their ability to (1) handle complex decision-making in the face of incomplete and/or imperfect data, (2) include clinical expertise when processing 3DGA and (3) classify mixed gait patterns such as CP gait. The proposed BN achieved promising accuracy rates as 73% - 100% of all patients were classified correctly by the network.

ACKNOWLEDGEMENTS
LVG received a PhD fellowship of the Research Foundation Flanders (FWO), Belgium.

REFERENCES
4. Schwartz et al., Gait Posture 20: 196-203, 2004
5. Netica 4.08, Norsys Software corp., Vancouver, Canada

Figure 1: Bayesian network for the automatic classification of gait patterns observed in the ankle and knee in the sagittal plane, for the stance and swing phase separately.