RECOGNITION OF GAIT DISTURBANCES IN PATIENTS WITH NORMAL PRESSURE HYDROCEPHALUS USING A COMPUTER DYNOGRAPHY SYSTEM

Gait and body balance disturbances are important clinical problems in patients with normal pressure hydrocephalus (NPH). They affect patients' locomotion and lead to a higher risk of falls. The gait pattern may be described as durations of the single and double support and of a stance phase. The aim of the present study was to apply the pattern recognition methods for the evaluation of gait disturbances in patients with NPH before and after neurosurgical treatment (shunt implantation). The results indicate that the parameters measured with a Computer DynoGraphy (CDG) system may effectively differentiate changes of gait in patients with NPH.

Key words: computer dynography system, gait analysis, k-NN classifier, normal pressure hydrocephalus, pattern recognition

INTRODUCTION

Normal pressure hydrocephalus (NPH) is known as a triad of progressive mental disturbances, urinary incontinence, and gait disturbance (1-3). Gait disturbances are usually the earliest symptoms in normal pressure hydrocephalus (NPH) and are described as apraxia (4-6). The NPH patients have difficulties in
turning, starting and stopping to walk. The NPH gait can finally lead to a total loss of ability to move. The NPH in its several forms and differential complications is a major problem in neurosurgery practice (7, 8).

The gait analysis and a study of movement patterns, has changed very fast over the last few decades (9, 10). Advances in computer technology (11, 12) and data analysis techniques (13, 14) have influenced the rapid development and progress of this new field. Computerized gait analysis has become a valuable tool in clinical practice (7, 15). On the other hand, the evaluation of gait disturbances in patients with NPH is very difficult in neurological categories and requires great clinical experience (6, 7, 16). Recently there have appeared new pointing to the use of studies the Computer DynoGraphy (CDG) system for measuring the ground reaction force, enabling a quantitative analysis in several gait disorders (17-24). It could also be helpful in the evaluation of effects of treatment and rehabilitation. The CDG system has already been successfully used: (i) in patients with subcortical vascular encephalopathy (17), (ii) in hemiplegic patients (18), (iii) in patients with upper and lower leg amputations (19), (iv) in rheumatologic patients before and after knee surgery (20), (v) in cases of patellofemoral pain syndrome (21), or (vi) in post-stroke patients (22). Our Department of Neurosurgery provides studies on gait by means the CDG system in patients with NPH before and after neurosurgical treatment, i.e. a shunt implantation (23, 24).

The aim of the present study was to apply pattern recognition algorithms for the evaluation of gait disturbances in patients with NPH.

MATERIAL AND METHODS

Subjects and measurements

Three groups were studied (age range 50-65 years): (i) 21 healthy subjects as a control group (11 males and 10 females), (ii) 21 patients with the NPH (11 males and 10 females), and (iii) 21 patients with the NPH after the shunt implantation (11 males and 10 females). The gait pattern was measured with the Infotronic Ultraflex (version 2.70), with the use of the CDG system of calculating the distribution of forces of ground reaction during walking. The system has already been described (23). The gait cycle consists of a single support, a double support and a stance phase. Hence, a set of examined parameters of the gait pattern includes: the duration of the single and double support and the duration of a stance phase, determined for each foot (leg) separately. Table 1 presents these variables defined as features and studied groups, named classes.

Statistical analysis with application of pattern recognition algorithms

The features presented in Table 1 were used for the classifier construction based on the k-nearest neighbors (k-NN) rule, known from the theory of pattern recognition (25, 26). A classifier is an algorithm for recognizing the class of an object if values of its features are known. The k-NN classifier requires a set of objects with known class membership. Such a set is called a reference set (X). The new object x, from outside the reference set, is assigned to the class most frequently represented among its k closest neighbors, which are searched in the reference set. The value of k
can be established experimentally by the leave-one-out method (25) which is a method for
eperimental estimation of the misclassification rate. It consists in classifying each object \( x \) in \( X \) by
\( k \)-NN classifier operating with the reference set decreased by a currently classified object \( x \), \( i.e. \) the
\( k \) nearest neighbors are searched for in the set \( X-\{x\} \). Some of these objects may be misclassified.
The ratio \( E_r=r/m \) (error rate) estimates the expected percentage of misclassifications where \( r \) is a
number of misclassified objects and \( m \) is the numerical force of the reference set \( X \). To determine
the optimal value of \( k \), the error rates \( E_r \) are calculated for \( k=1, 2, … , m \). We then choose the proper
\( k \) which corresponds with the lowest \( E_r \). The reference set can contain features not related to the
considered classes. The presence of such features usually increases the misclassification rate. This
is why feature selection is essential. If the number of features is sufficiently small, which is our case,
the error rates, corresponding with the optimal value of \( k \), are calculated for all feature
combinations. The feature combination that offers the lowest error rate \( E_r \) forms a set of selected
features. To classify new objects, \( i.e. \) objects from outside the reference set, the \( k \)-NN rule is used
with optimal \( k \) and selected features.

The standard \( k \)-NN rule, described above, promises satisfactory performance, but better results
can be obtained by pair-wise \( k \)-NN classifier. The \( k \)-NN classifiers are constructed separately for
each class pair, including feature selection and determination of the optimal \( k \). The new object
(\( \text{which does not belong to} \ X \)) is classified by voting of these class pair classifiers, \( i.e. \) \( k \)-NN classifier
can decide between two classes only.

The \( k \)-NN rule was used to establish the dependence between features (\( i.e. \), parameters of the
gait pattern) and classes (studied groups). If the dependence between classes and features is strong
then the percentage of misclassifications is low and \( \text{vice versa} \). To verify whether the dependence
between classes and features is statistically significant the Chi-square test was used.

## RESULTS AND DISCUSSION

We performed an analysis of gait parameters with the use of the \( k \)-NN
classifier in three groups of subjects (\textit{Table 1}). The results of \( k \)-NN classification
for the groups studied in various combinations are presented in \textit{Table 2}.

Each single feature, then a set of features of the left leg \{F1, F2, F3\} and the
right leg \{F4, F5, F6\}, and also both legs (all features together) were analyzed.
Error rates for all classes treated jointly and for the two-class comparison, patients
before and after surgery (II vs. III class), are large, \( i.e. \), from 28.6 to 42.9%.
Differentiation between all three classes and between the two classes of patients
is difficult (error rates from 25.4 to 31.0%), no matter form which leg the features
were measured. The use of features measured from both legs jointly slightly
improves differentiation between all classes (\( E_r=20.6\% \)) and between the groups
of patients (\( E_r=23.8\% \)). Further improvements were achieved by feature selection

### Table 1. Groups and variables.

<table>
<thead>
<tr>
<th>GROUPS (Classes)</th>
<th>VARIABLES (features) – durations of gait phases</th>
<th>Left leg</th>
<th>Right leg</th>
</tr>
</thead>
<tbody>
<tr>
<td>I – Control (healthy subjects)</td>
<td>F1 – single support</td>
<td>F4 – single support</td>
<td></td>
</tr>
<tr>
<td>II – NPH before the shunt implantation</td>
<td>F2 – double support</td>
<td>F5 – double support</td>
<td></td>
</tr>
<tr>
<td>III – NPH after the shunt implantation</td>
<td>F3 – stance phase</td>
<td>F6 – stance phase</td>
<td></td>
</tr>
</tbody>
</table>
The misclassification rates were remarkably lower for differentiation between the control subjects and the patient groups (class I vs. class II and class I vs. class III). The smallest $E_r$ was achieved for "stance" features F3 (9.5% and 11.9% respectively) and F6 (4.8% and 14.3% respectively). When all features of the right leg were used, the error rate for recognition between the classes I and II slightly decreased (from 9.5 to 7.1%) and for the differentiation between the classes I and III remained unchanged at 11.9%. Using all features of the left leg for the two above-mentioned comparisons, reduced the error rate to 4.8%. Perfect differentiation of the classes I and II was obtained if the features of both legs were used ($E_r=0.0\%$). For the classes I and III, the result with the use of both leg features was worse (9.5%) as compared with that for features of the left leg only (4.8%). Feature selection reduced the misclassification rates to zero for both comparisons (class I vs. class II and class I vs. class III).

The leave-one-out method enables calculation of the number $r_{ij}$ of objects from class $i$ assigned to class $j$. This number enables computation of probability $p_{ij}$ that an object from class $i$ will be classified to class $j$ and the probability $q_{ij}$ that an object assigned to class $i$ belongs to class $j$: $p_{ij}=r_{ij}/\Sigma r_{ij}$ and $q_{ij}=r_{ij}/\Sigma r_{ji}$. These numbers form three types of confusion matrices $R=\{r_{ij}\}$, $P=\{p_{ij}\}$ and $Q=\{q_{ij}\}$.

Table 3 contains detailed results, i.e., the confusion matrices R, P, and Q which concerned differentiation between all three considered classes jointly, i.e., recognition among healthy subjects and NPH patients, before and after the shunt implantation, using the whole set of features (the parameters of both legs) and the set of selected features. In case of matrix P, the percentages are calculated in relation to the class numerical forces, and the matrix Q contains percentages computed in relation to numbers of objects classified to classes I, II, and III.

<table>
<thead>
<tr>
<th>FEATURES</th>
<th>All Classes $E_r(%)$</th>
<th>Control vs. Patients $E_r(%)$</th>
<th>Patients $E_r(%)$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>I-Control</td>
<td>II-Before SI</td>
<td>III-After SI</td>
</tr>
<tr>
<td>F1</td>
<td>38.1</td>
<td>11.9</td>
<td>19.1</td>
</tr>
<tr>
<td>F2</td>
<td>38.1</td>
<td>14.3</td>
<td>14.3</td>
</tr>
<tr>
<td>F3</td>
<td>34.9</td>
<td>9.5</td>
<td>11.9</td>
</tr>
<tr>
<td>F4</td>
<td>31.8</td>
<td>14.3</td>
<td>14.3</td>
</tr>
<tr>
<td>F5</td>
<td>39.7</td>
<td>9.5</td>
<td>21.4</td>
</tr>
<tr>
<td>F6</td>
<td>34.9</td>
<td>4.8</td>
<td>14.3</td>
</tr>
<tr>
<td>Left leg, {F1,F2,F3}</td>
<td>27.0</td>
<td>4.8</td>
<td>4.8</td>
</tr>
<tr>
<td>Right leg, {F4,F5,F6}</td>
<td>25.4</td>
<td>7.1</td>
<td>11.9</td>
</tr>
<tr>
<td>Both legs, Without feature selection:</td>
<td>20.6</td>
<td>0.0</td>
<td>9.5</td>
</tr>
<tr>
<td>After feature selection:</td>
<td>12.7</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>Set of selected features:</td>
<td>{F1,F2,F4,F5}</td>
<td>{F1,F6}</td>
<td>{F1,F3,F4,F6}</td>
</tr>
</tbody>
</table>

SI - shunt implantation, F1-F6 features of the gain pattern; see description in Material and Methods.

Table 2. Misclassification rates ($E_r$) found by means of the $k$-NN classifier for studied groups of subjects in various combinations.
The level of significance can be calculated taking into account two qualitative variables: the true and the assigned class. The assigned class is the one pointed to by the features. Thus, the relation between the true and the assigned classes corresponds to the relation between features and classes. The level of significance for dependence between classes and features can be calculated on basis of matrix R, treated as the contingency table, for the complete set of 6 features. According to the Chi-square test, this dependence is statistically significant (P<0.001). Calculation of P value would not be correct for selected features (27).

When all features were used, the subjects from class I were correctly classified (95.2%) more frequently than the object from class II (76.2%) and class III (66.7%). Among patients classified to class I (control healthy group) 9.1% came in fact from class III and 90.9% were correctly recognized. The least dependable was classification to class III (70.0%); 25% of subjects assigned to class III came really from class II and 5% from class I. After the feature selection, all subjects from class I were correctly classified and all subjects qualified to class I came really from this class. For classes II and III, 81% of subjects were properly classified and exactly the same percentage of subjects assigned to class II or III came from these classes (19% were in fact from class III or II respectively).

In summary (i) The choice of parameters of gait pattern was appropriate. It consisted of single support, double support, and stance phase durations; (ii) The set was very effective in differentiation of gait changes between the patients with NHP and control subjects; (iii) The set was effective in differentiation of gait changes in the NPH patients with and without shunt implantation. The Computer DynoGraphy system can be a useful tool in the diagnosis and assessment of treatment in NPH patients. The present study demonstrates a novel approach to

**Table 3. Confusion matrices for 3 classes and features of both legs: (a) without feature selection and (b) after feature selection**

<table>
<thead>
<tr>
<th>Number of cases from class $i$ (row) assigned to class $j$ (column), matrix R</th>
<th>Probability (%) that a case from class $i$ (row) will be assigned to class $j$ (column), matrix P</th>
<th>Probability (%) that a case assigned to class $i$ (row) belongs to class $j$ (column), matrix Q</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Assigned class</strong></td>
<td><strong>True Class</strong></td>
<td>I</td>
</tr>
<tr>
<td>I</td>
<td>20</td>
<td>0</td>
</tr>
<tr>
<td>II</td>
<td>0</td>
<td>16</td>
</tr>
<tr>
<td>III</td>
<td>2</td>
<td>5</td>
</tr>
</tbody>
</table>

21 cases in control group, NPH without shunt & NPH with shunt implantation, respectively, jointly 63 subjects
Classification rule: 1-NN rule, error rate 20.6%, P<0.001 (Chi-square test), features: {F1, F2, F3, F4, F5, F6}

<table>
<thead>
<tr>
<th>Assigned class</th>
<th><strong>True Class</strong></th>
<th>I</th>
<th>II</th>
<th>III</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>21</td>
<td>0</td>
<td>0</td>
<td>I</td>
</tr>
<tr>
<td>II</td>
<td>0</td>
<td>17</td>
<td>4</td>
<td>II</td>
</tr>
<tr>
<td>III</td>
<td>0</td>
<td>4</td>
<td>17</td>
<td>III</td>
</tr>
</tbody>
</table>

21 cases in the control group, NPH without shunt & NPH with shunt implantation, respectively, jointly 63 subjects
Classification rule: 1-NN rule, error rate 12.7%, features: {F1, F2, F4, F5}
the analysis of NPH gait using the algorithms of the theory of statistical pattern recognition.

Acknowledgments: This study was partly supported by grant from the Polish Ministry of Education and Science (Grant No. R13 041 02).

Conflicts of interest: The authors had no conflicts of interest to declared in relation to this article.

REFERENCES


Received: August 9, 2008
Accepted: August 22, 2008

Author’s address: L. Czerwosz, Medical Research Center, Polish Academy of Sciences, Pawinskiego 5 St., 02-106 Warsaw, Poland; phone: +48 22 608-64-64; e-mail: czerwosz@cmdik.pan.pl