PredictMed: A logistic regression-based model to predict health conditions in cerebral palsy

Carlo M Bertoncelli
Florida International University, USA; Children Hospital E.E.A.P. H. Germain, France

Paola Altamura
D’Annunzio University of Chieti–Pescara, Italy

Edgar Ramos Vieira
Sundaraja Sitharama Iyengar
Florida International University, USA

Federico Solla
Florida International University, USA; Lenval University Children’s Hospital of Nice, France

Domenico Bertoncelli
University of L’Aquila, Italy

Abstract
Logistic regression–based predictive models are widely used in the healthcare field but just recently are used to predict comorbidities in children with cerebral palsy. This article presents a logistic regression approach to predict health conditions in children with cerebral palsy and a few examples from recent research. The model named PredictMed was trained, tested, and validated for predicting the development of scoliosis, intellectual disabilities, autistic features, and in the present study, feeding disorders needing gastrostomy. This was a multinational, cross-sectional descriptive study. Data of 130 children (aged 12–18 years) with cerebral palsy were collected between June 2005 and June 2015. The logistic regression–based model uses an algorithm implemented in R programming language. After splitting the patients in training and testing sets, logistic regressions are performed on every possible subset (tuple) of independent variables. The tuple that

Corresponding author:
Carlo M Bertoncelli, Hôpital pour Enfants, E.E.A.P. H. Germain, 337 Chemin Saint Antoine de Ginestiere, 06200 Nice, France.
Email: bertoncelli@unice.fr
shows the best predictive performance in terms of accuracy, sensitivity, and specificity is chosen as a set of independent variables in another logistic regression to calculate the probability to develop the specific health condition (e.g. the need for gastrostomy). The average of accuracy, sensitivity, and specificity score was 90%. Our model represents a novelty in the field of some cerebral palsy–related health outcomes treatment, and it should significantly help doctors’ decision-making process regarding patient prognosis.

**Keywords**
clinical decision-making, databases, data mining, decision-support systems, machine learning

**Introduction**

Cerebral palsy (CP) is the term used for a group of nonprogressive disorders of motor control and posture that occur secondary to damage to the developing brain during the early stages of its development. Predictive models using logistic regression (LR) are widely used in the healthcare, but only recently, they started to be used to predict health-related outcomes in subjects with CP. This article discusses the potential of using a LR-based predictive model (hereby named PredictMed) for health outcomes and presents examples of the application of the model. PredictMed is an algorithm based on LR s implemented in R programming language. It has been developed to predict scoliosis development and to identify factors associated with autistic features, intellectual disabilities, and impaired adaptive functioning in subjects with CP.

The purpose of this article was to explain the overall procedures used and highlight the versatility of the PredictMed model using gastrostomy as an additional example.

**Methods**

**Ethical statements**

All procedures were performed in accordance with the ethical standards of the institutional research committee and with the Declaration of Helsinki 1975, revised Hong Kong 1989 and its later amendments. All participants and/or parents consented to participate. The data were anonymized and analyzed according to the requirements of French Reference Method 003. Ethics committee approval and informed consent was registered with number “2017728 v 0-MR003 (Reference Method 003).”

**Design**

Linear regression and LR are among the first machine learning (ML) algorithms used to implement predictive models in healthcare. The number of parameters (hereby named features or independent variables) should not be too large but should contain enough information to predict the output. We deal with the Feature Reduction problem in two steps: first, by using Fisher’s exact test to identify parameters (i.e. the LR independent variables) associated with the binary dependent variable (e.g. need/not need for gastrostomy) and second, by testing and ranking the predictive performance of each of all possible subsets (hereby named tuples) of the parameters found in the first step. The PredictMed algorithm takes as input a data matrix having patients as rows and parameters (variables) as columns. The first column contains the values of the binary dependent variable while other columns those of independent variables (Table 1).

If we have N independent variables, we can form $2^{N-1}$ different tuples of independent variables. For example, if the independent variables were A, B, and C, we would have $2^{3-1} = 7$ tuples: ABC,
AB, AC, BC, A, B, C. After splitting the data in a training and test set, PredictMed provides the following outputs for each tuple of the $2^{N-1}$:

1. A classification using LR, predicting the outcome of the dependent variable based on the probability of this outcome (if the probability of positive outcome is above a threshold, then we predict the outcome as positive: we try different thresholds in a range from 0.1 to 0.9).
2. Accuracy, sensitivity, and specificity of the prediction (calculated on the test set)
3. Ranking of the odds ratio and p-value significance of each independent variable of the tuple in influencing the outcome of the dependent variable.

Based on steps 1 and 2, we can find out the tuple with the best predictive performance in predicting the outcome of the dependent variable. Since this tuple is a subset of the initial set of all the independent variables (features), in doing this, we perform a Feature Reduction.

**Data sources and study population**

Data were collected between June 2005 and June 2015 for a multinational, double-blinded, cross-sectional descriptive study on the need of gastrostomy placement in children with CP. Narrative notes were coded and entered into the electronic database PredictMed. Assessments and data collection for the implementation of the model were conducted in the last 6 months of 2014, and data analysis began in December 2015 and lasted 12 months. The fields included patient demographics, the results of diagnostic procedures, and functional and neurological assessments following the “European Society for Pediatric Gastroenterology, Hepatology and Nutrition Guidelines for the Evaluation and Treatment of Gastrointestinal and Nutritional Complications in Children with Neurological Impairment.” The study followed the guidelines of the transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD) statement.

Nine hundred twenty-six potentially eligible patients with severe developmental disabilities and CP were screened (Figure 1). Patients were included in the analysis if they were between 12 and 18 years old; had CP diagnosed by a pediatric neurologist; had spastic, dystonic, mixed spastic/dystonic, or hypotonic CP evaluated and classified using the “Surveillance of Cerebral Palsy in..."
Europe" system;\textsuperscript{11} and were followed for at least 3 years after initial diagnosis. Patients with pro-
gressive encephalopathy or spinal neuropathology, well-developed gross motor functional capaci-
ties (defined by Gross Motor Function Measure-88 > 70),\textsuperscript{12} and good manual abilities (Manual 
Ability Classification System (MACS) < 2)\textsuperscript{12} were not included. Our study included a total of 130 
adolescents (72 male, 58 female). The mean age was 16 years 6 months (standard deviation (SD) 
1.8). Mean follow-up time was 5 years (range: 3–12 years) (Table 2).

Patients’ nutritional and respiratory status. Patients’ nutritional status was assessed by pediatricians 
based on weight, height, body mass index (BMI in kg/m\textsuperscript{2}), and skinfold thickness. Patients were 
classified as undernourished if they had a falling weight centile, weight loss, a discrepancy between 
weight and measured or calculated height centile BMI under the 3rd centile, or BMI < 18 kg/m\textsuperscript{2}.

Dysphagia was defined as difficulty in swallowing manifested as coughing, spluttering, or choking 
during eating⁄drinking; aversion to eating; or excessive time to eat a meal (longer than 30 min).\textsuperscript{9}

Gastroesophageal reflux (GER) was defined as the passage of gastric contents into the esopha-
gus with or without regurgitation and vomiting. Gastroesophageal reflux disease (GERD) was 
defined as a complication of GER, like failure to thrive, Sandifer syndrome, or frequent nighttime 
waking. Respiratory disease was defined as two or more episodes of respiratory infection requiring 
antibiotic therapy and/or admission to hospital for treatment of pneumonia or recurrent wheezing 
requiring bronchodilators.\textsuperscript{9}

Feeding abilities. Assessment of eating and drinking performance was made using the Eating and 
Drinking Ability Classification System for Individuals with Cerebral Palsy (EDACS)\textsuperscript{13}(Table 3).

Feeding autonomy was classified by a pediatrician as children who had “some self-feeding 
skills” (mild or no feeding disorders) and those who were “fed by others,” or who needed gastro-
tomy/laparoscopic antireflux surgery (LARS) (with severe feeding disorders). The indications for 
gastrostomy placement included malnutrition or failure to thrive, as well as oropharyngeal dyspha-
gia.\textsuperscript{14} Types of gastrostomy included surgical gastrostomy (SG) using the Stamm technique,

Figure 1. Flow diagram of study participants for analysis.
percutaneous endoscopic gastrostomy (PEG),\textsuperscript{15} and laparoscopic-assisted PEG. LARS, like Nissen fundoplication, was the treatment option for children with GERD when medical therapy failed and GER was confirmed by esophageal pH Monitoring.\textsuperscript{16}

**Etiology.** Developmental disability (DD) and CP etiology were classified as pre-natal (genetic, cerebrovascular malformation, infection, or vascular), peri-natal (anoxic-ischemic or infectious) or post-natal (cranial trauma, infectious, epilepsy, or post-natal anoxic/ischemic injury).

**Motor status.** All patients were assessed for gross motor function using the gross motor function classification system (GMFCS),\textsuperscript{17} manual abilities by the MACS\textsuperscript{18}(Table 3) following the recommendation of the “Surveillance of Cerebral Palsy in Europe” proposal.\textsuperscript{19} They use a 5-level classification system (from I to V) and higher scores indicate lower levels of motor functioning.

**Epilepsy.** Presence of epilepsy was determined by a pediatric neurologist and identified as “well controlled” or “intractable.” We followed the definition of the International League against Epilepsy, which defines intractable epilepsy as patients who follow treatment with at least 2 appropriate antiepileptic drugs treated at a comprehensive epilepsy center during long-term follow-up.\textsuperscript{20}

**Neurological status.** Neurological status was classified according to the anatomy of spastic disorder (hemiplegia, diplegia, and tri/quadriplegia) and the presence of dystonia. Spasticity was quantified using the Bohannon and Smith modified Ashworth Scale and/or the Modified Tardieu Scale.\textsuperscript{21} Trunk muscles’ tone was also assessed as hypotonic, spastic, or normal. The presence of scoliosis was determined by the presence of a spinal radiograph with a Cobb angle >10°, and it was labeled as “severe” if the Cobb angle was >40°.\textsuperscript{22,23}

---

### Table 2. Clinical presentation according to the presence and absence of self-feeding skills.

| Patients profile | Pediatric hospital A | Children hospital B | Multicenter | Total (%) |
|------------------|----------------------|---------------------|-------------|-----------|
|                  | No self-feeding skills\textsuperscript{a} | Some self-feeding skills | Total | No self-feeding skills\textsuperscript{a} | Some self-feeding skills | Total | Overall |
| Number of patients, n (%) | 442 (52) | 338 (48) | 880 (100) | 336 (72) | 114 (28) | 550 (100) | 1130 (100) |
| Male, n (%) | 24 (53) | 21 (47) | 45 (100) | 20 (74) | 7 (26) | 27 (100) | 72 (100) |
| Female, n (%) | 18 (51) | 17 (49) | 35 (100) | 16 (69) | 7 (31) | 23 (100) | 58 (100) |
| Mean age (SD) | 16.2 (1.6) | 16.6 (1.7) | 16.4 (1.7) | 15.1 (1.8) | 15.3 (1.5) | 15.1 (1.8) | 15.7 (1.8) |
| Gastrostomy, n (%) | 19 (100) | 0 (0) | 19 (100) | 6 (100) | 0 (0) | 6 (100) | 25 (100) |
| Spasticity, n (%) | 40 (63) | 23 (37) | 63 (100) | 30 (68) | 14 (32) | 44 (100) | 107 (100) |
| Dystonia, n (%) | 7 (70) | 3 (30) | 10 (100) | 11 (78) | 3 (22) | 14 (100) | 24 (100) |
| No epilepsy, n (%) | 4 (21) | 15 (79) | 19 (100) | 6 (35) | 11 (65) | 17 (100) | 36 (100) |
| Well-controlled epilepsy | 26 (55) | 21 (45) | 47 (100) | 17 (74) | 6 (26) | 23 (100) | 70 (100) |
| Intractable epilepsy, n (%) | 13 (93) | 1 (7) | 14 (100) | 10 (100) | 0 (0) | 10 (100) | 24 (100) |
| Etiology causes, n (%) | 42 (52) | 38 (48) | 80 (100) | 36 (72) | 14 (28) | 50 (100) | 130 (100) |
| Ante-natal | 21 (46) | 25 (54) | 46 (100) | 16 (69) | 7 (31) | 23 (100) | 69 (100) |
| Peri-natal | 14 (61) | 9 (39) | 23 (100) | 17 (71) | 7 (29) | 24 (100) | 47 (100) |
| Post-natal | 7 (64) | 4 (36) | 11 (100) | 3 (100) | 0 (0) | 3 (100) | 14 (100) |
| Neuromuscular scoliosis | 24 (77) | 7 (23) | 31 (100) | 23 (82) | 5 (18) | 28 (100) | 59 (100) |

SD: standard deviation.

\textsuperscript{a}No self-feeding skills include oral feeding by others and gastrostomy.
Prediction model

The algorithm consists of a 5-step-process (Figure 2). Steps from 1 to 4 involve Feature Selection based on Fisher’s test and a brute-force algorithm using LR. Step 5 involves the calculation of the outcome probability for new patients using the selected features as independent variables.

First step: finding potential predictors. The first statistical analysis was performed with “OpenEpi,” a web-based epidemiological calculator,24 by creating contingency tables and using Fisher exact test to identify factors associated with feeding disorders causing the need for a gastrostomy. Confidence intervals and distribution frequencies were calculated for the binary dependent variable “need for gastrostomy (yes/no)” versus each independent variable. Statistical significance was set at 95% confidence interval. A total of 10 independent variables (potential predictors hereby named features) were associated with feeding disorders: etiology (ET), spasticity (SP), dystonia (D), epilepsy (E), gender (SE), neuromuscular scoliosis (NS), truncal tone disorders (TT), GMFCS score (GMFCS), MACS score (M), and EDACS score (ED).

Second step: generating all possible subsets (tuples) of the potential predictors. All the possible subsets (hereby named tuples) of the 10 parameters (features) found in step 1 were studied. For a set of n elements, the number of all the possible subsets of m (with m < n) elements of the set is24
The total number of tuples that can be generated with \( n \) elements and every value of \( m \) between 1 and \( n \) is \( C(n) = 2^{n-1} \).

For \( n = 10 \) parameters, we have a total of \( C(10) = 2^{10-1} = 1023 \) tuples (Table 4). Each parameter in a tuple is an independent variable for the LR to predict the outcome of the dependent binary variable (need/not need for gastrostomy).

**Third step.** For each tuple, perform a LR and assess the tuple performance in predicting the outcome. For each tuple of the 1023, an LR was run to predict the outcome (need/not need for gastrostomy). The patients were randomly split in a “training set” to train the LR model and a “test set” to test the performance of the model. Model performance on each tuple was evaluated based on the accuracy, sensitivity, and specificity\(^{25} \) of the predictions. The open source software “R,”\(^{2–4} \) implementing the general linear model, “glm()” was used to predict the probability that each subject would have the outcome (need/not need for gastrostomy). If the probability was above a threshold, the subject was classified as “will need gastrostomy placement.”

Sensitivity, specificity, and accuracy were described in terms of true positives (TPs), true negatives (TNs), false negatives (FNs), and false positives (FPs).\(^{25} \)

Sensitivity was defined as the proportion of actual positives and identified as such (TP/(TP + FN)). Specificity was defined as the proportion of actual negatives and identified as such (TN/(TN + FP)). Accuracy was defined as the proportion of TP and TN in all assessments and identified as (TP + TN)/(TP + TN + FP + FN).

To minimize dependency from the composition of training and test sets, we used cross validation—a technique for assessing how the results of a statistical analysis will generalize to an independent data set. By splitting the original data set, we randomly generated 20 different couples of training and test sets, and for each couple, we calculated the accuracy, sensitivity, and specificity.
of the predictions with different thresholds, and finally, we averaged (on the 20 rounds), so getting (for each threshold) the average accuracy, sensitivity, and specificity.

**Fourth step.** Finding which tuple best predicts the outcome. The 1023 tuples were so tested to find the one with the best predicting performance in terms of accuracy, sensitivity, and specificity. When predicting the need for gastrostomy, the best tuple (with a probability threshold of 0.5) was composed of 7 parameters: NS, TT, SP, DY, EP, GMFCS, and ED.

Table 5 shows that the probability threshold of 0.5 was the best to be used in conjunction with this 7-factor tuple to achieve the highest accuracy, sensitivity, and specificity (each one calculated as the average of the results on 20 different test sets).

Table 6 shows as among the 7-elements tuples (120 out of 1023), the tuple with threshold=0.5 showed the best overall performance in terms of accuracy, specificity, and sensitivity.

Moreover, we make sure that this 7-element tuple performs better even if compared with all other 1023 tuples of 1, 2, 3, . . .10 elements. The choice of the best tuple to be used in LR is not just based on the best predictive performance. The tuple itself must also make sense from a clinical point of view in accordance with the aim of the research.
Table 6. Overall performances in terms of accuracy, specificity, and sensitivity.

| Tuples | Threshold Average | Accuracy | Sensitivity | Specificity |
|--------|-------------------|----------|-------------|-------------|
|        |                   |          |             |             |
| 1      | NS TT ET SP DY EP SU | 0.5      | 0.8044615   | 0.1231151   | 0.9638505   |
| 2      | NS TT ET SP DY EP G  | 0.5      | 0.7754615   | 0.1750595   | 0.9197582   |
| 3      | NS TT ET SP DY EP ED | 0.5      | 0.9401538   | 0.7525198   | 0.9760981   |
| 4      | NS TT ET SP DY EP S  | 0.5      | 0.8238462   | 0.2236706   | 0.9601501   |
| 5      | NS TT ET SP DY SU G  | 0.5      | 0.7501538   | 0.0083333   | 0.923214    |
| 6      | NS TT ET SP DY SU ED | 0.5      | 0.9398462   | 0.7543849   | 0.970618    |
| 7      | NS TT ET SP DY EP G ED| 0.5      | 0.948      | 0.7664087   | 0.9837132   |
| 8      | NS TT ET SP DY SU S  | 0.5      | 0.787      | 0.0083333   | 0.966243    |
| 9      | ET SP DY EP SU G S   | 0.5      | 0.7984615   | 0.2474802   | 0.9289542   |
| 10     | ET SP DY EP SU ED S  | 0.5      | 0.9343846   | 0.7484325   | 0.9676055   |
| 11     | ET SP DY EP G ED S   | 0.5      | 0.9460769   | 0.7589683   | 0.9807556   |
| 12     | ET SP DY SU G ED S   | 0.5      | 0.9594615   | 0.8381944   | 0.9837366   |
| 13     | ET SP EP SU G ED S   | 0.5      | 0.9403846   | 0.7601587   | 0.9734703   |
| 14     | ET DY EP SU G ED S   | 0.5      | 0.9498462   | 0.7798016   | 0.9797717   |
| 15     | SP DY EP SU G ED S   | 0.5      | 0.9422308   | 0.7571825   | 0.978124    |

NS: neuromuscular scoliosis; TT: truncal tone disorders; ET: etiology; SP: spasticity; DY: dystonia; EP: epilepsy; ED: EDACS score; EDACS: Eating and Drinking Ability Classification System; SU: previous surgeries.

Table 7. List of the predictive logistic regression coefficients (independent variables) for the need of gastrostomy.

Logistic regression for gastrostomy need

| Independent variables | Odds estimate | Standard error | Z ratio | Prob(>|z|) | p value |
|-----------------------|---------------|----------------|---------|-----------|---------|
| Intercept             | –13.1294      | 4.6271         | –2.838  | 0.004547  |         |
| Neuromuscular scoliosis| 0.8034        | 1.1997         | 0.670   | 0.503049  |         |
| Truncal tone disorders (TT) | –0.1718       | 0.8340         | –0.206  | 0.836773  |         |
| Spasticity(SP)        | –0.2692       | 0.5180         | –0.520  | 0.603232  |         |
| Dystonia (D)          | –1.6755       | 1.2376         | –1.354  | 0.175786  |         |
| Epilepsy (EP)         | 2.2939        | 1.0756         | 2.133   | 0.032954  |         |
| GMFCS score (G)       | 3.3473        | 1.6499         | –2.029  | 0.042484  |         |
| EDACS score (ED)      | 6.0862        | 1.6487         | 3.692   | 0.000223  |         |
| Performance prediction| 0.948         | 0.7664087      |         | 0.9837132 |         |

GMFCS: Gross Motor Function Classification System; EDACS: Eating and Drinking Ability Classification System. Logistic regression: the increasing of EP, G, and ED (positive values) favors the need of gastrostomy ("Estimate" column). For every unit increase in EP, the log odds = ln(p/1 − p) increases 2.2939 times (where p = probability to have gastrostomy). In the same way, for every unit increase in SP, the log odds = ln(p/1 − p) decreases –0.2692 times. The "Pr(>|z|)" column indicates the significance strength of the respective parameter in terms of p-value as gastrostomy need predictor. The significance of EP, G, and ED in predicting gastrostomy need is very probable, with a p-value < 0.05.

The best tuple in predicting the outcome of our binary dependent variable (need for gastrostomy: yes/no) had the following 7 parameters: NS, TT, SP, D, EP, GMFCS, and ED (Table 7).

Multiple LR is performed with open source software “R”,2–4 using the general linear model, “glm(),” to predict the probability that each subject would need gastrostomy. If the probability is
above the threshold of 0.5 (found in step 4), we classify the subject needing gastrostomy placement.

**Fifth step.** Use the best tuple to predict the binary outcome for new incoming patients. We use again LR to predict the outcome of our binary dependent variable (need for gastrostomy: yes/no) having as independent variables the 7 parameters found in step 4: NS, TT, SP, D, EP, GMFCS, and ED.

**Model highlights**

**Feature reduction**

PredictMed performs a Feature Reduction, using a brute-force algorithm that ranks the predictive performance of all the possible tuples of features. This allows us to find the features (e.g. independent variables) that most influence the outcome of the dependent variable. In the end, this Feature Reduction discovers the best predictors of the outcome of the dependent variable.

**Cross correlation**

As explained above, for each tuple, PredictMed cross-correlates our regression on multiple training and test sets and then averages accuracy, sensitivity, and specificity.

This means that for each tuple, we randomly re-shuffle the composition of training and test. On a total of 130 patients, 80% of them (104) are in the training set and 20% (26) in the test set. Composition of the training a test set was randomly determined through the R function createDataPartition(). At each re-shuffle, we perform regression, prediction, and computation of accuracy, sensitivity, and specificity. Finally, we compute the average on all re-shuffles (about 20) of accuracy, sensitivity, and specificity. The average values we find are reliable since they do not depend from a composition of training and test set.

**Threshold**

PredictMed fixes a probability threshold. After predicting the outcome of the dependent variable if the probability of this outcome is above the threshold, we classify the outcome as positive (or vice versa). Our algorithm tries different values of thresholds, finding the values that lead to the best predictive performance in terms of accuracy, sensitivity, and specificity. Scholars can then decide which threshold to choose to maximize accuracy or specificity or sensitivity depending on the disorders we want to predict (see Discussion section).

**Results**

A breakdown of the clinical presentation according to the presence or absence of feeding disorders and gastrostomy placement is shown on Table 2. Factors associated with gastrostomy placement in multivariate analysis were epilepsy (p .03), poor motor function (p .04), and male gender (p .04). The average score of accuracy, sensitivity, and specificity was 90% (Table 7).

**Model development path and its adaptability to multiform neurological diseases in CP**

The model has been developed and validated through the previous studies1–4 and first in literature showed the following results:
On 2017, PredictMed highlighted the presence of hip surgery (p .002), intractable epilepsy (p .04), and female gender (p .07) as risks factors of scoliosis.\textsuperscript{1,2} Average of accuracy, sensitivity, and specificity of the predictive model was 74%.

On 2018, PredictMed identified that factors associated with intellectual disabilities and impaired adaptive functioning\textsuperscript{3} were poor manual abilities (p $\leq$ .001), gross motor function (p $\leq$ .001), and type of epilepsy (intractable: p .04; well controlled: p .01). Best threshold performed an accuracy of 74%, sensitivity of 77%, and specificity of 70%, showing an average of 74%.

On 2019, PredictMed defined type of spasticity (hemiplegia $>$ diplegia $>$ tri/quadriplegia; p .04), communication disorders (p $<$ .001), intellectual disability (p .05), feeding abilities (p .002), and motor function (p.01) as factors associated with autism spectrum disorders.\textsuperscript{4} The best prediction model score performed an accuracy of 73%, sensitivity of 79%, and specificity of 72%, showing an average of 75%.

**Discussion**

*Should we choose a tuple based on the higher sensitivity or specificity?*

The choice depends on the aim of the research and its consequences.\textsuperscript{26} For studies with purpose to screen frequent diseases (e.g. breast cancer or HIV infection), higher sensitivity than specificity is required to limit FN and detect the biggest proportion of potentially sick patients. Conversely, higher specificity would be more useful if we were to carry out a “preventive” treatment without precisely knowing whether the subject actually had the comorbidity or the disease (e.g. GERD) or if we had to perform invasive or expensive investigations to ascertain the disease, for example, esophageal pH monitoring.

For this study, we opted to choose the highest specificity (98%)/accuracy (95%) and a good sensitivity (77%) (Tav.5) to avoid overestimation of likelihood of reflux.

Knowing that gastrostomy will likely be needed can inform patient management and family preparation for the procedure. Knowing about the likely need for the surgery allows for the setup of early dietary and food follow-up to determine the ideal moment for the gastrostomy to take place. The evaluation period of enteral nutrition tolerance with a nasogastric tube may be shortened; therefore, iatrogenic complications could be limited.

Previous cited researches\textsuperscript{1–4} about comorbidities in patients with CP aimed to identify subjects with higher risk of presenting or developing a specific comorbidity (scoliosis, cognitive impairments, and autistic features). For these patients, we should inform the family and use specific non-invasive diagnostic tools to confirm if the comorbidity is present or not. Therefore, the choice of a specificity–sensitivity–accuracy balance (74% / 75%) seems appropriate to limit both overestimation and underestimation of the risks.

**Future developments**

LR-based predictive models are widely used in the healthcare field, but just recently, it started to be used to predict health conditions such as comorbidities in children with CP. These models can provide faster and more objective predictions, potentially improving care of patients with CP and reducing healthcare costs and suffering. In our research, we now focus on the study of improvements of our prediction algorithm. At present, the LR algorithm can deal only with binary dependent variables.

Our next step is to modify it to perform also multi-class classification to work with dependent variables having more than two values as multiform types of epilepsies (in development). We also plan to investigate different methods for features reduction.\textsuperscript{27–31} In fact, with a higher number of
independent variables, the computation time increases accordingly (since we have to compute LR on all the possible tuples of independent variables). Given this, we are investigating other ML techniques used in various fields of healthcare. A new method of feature reduction could use, for example, information theory. Another topic to be studied in the future regards if and how to replace the LR algorithm with other types of supervised classification algorithms like support-vector machines or neural networks.

**Limitations**

A limitation of our prediction model is the small number of patients to study (hundreds), that could lead to a model overfitting. Thus, we first need to study and fine-tune the predictive model when increasing the number of patients (at least thousands) while maintaining the same number of independent variables (<15). Second, we need a further external validation of the model on a larger data set of new incoming patients (especially for step 5).

**Computing time**

Usually, we deal with data matrices having 100–300 rows (e.g. patients) and 5–15 columns (1 dependent variable + N independent variables). As previously discussed, the number of tuples and thus of LRs we have to perform is 2^{N-1}. If having, for example, 100 patients and N = 14, we have 2^{14-1} = 16383 tuples, for example, LRs to perform. Using a notebook with an i7 Intel processor, the computing time was around 20 hours and we can expect it to double for every unit increase of N (or if doubling the number of patients). Supposing to increase the number N of independent variables or the number of patients, a Feature Reduction algorithm to reduce N prior to implement LR will be needed. The outcome used to create and evaluate the validity of the PredictMed was gastrostomy placement independently of the time interval; in future studies, we will control for the time of assessment of the predictors and the time of gastrostomy placement.

**Conclusion**

We developed a LR-based predictive model (PredictMed), and we tested it on 5 different data sets of children with CP, each dataset with a different outcome (a binary dependent variable): (1) feeding disorders needing gastrostomy, yes/no (present study); (2) likelihood of scoliosis, yes/no; (3) presence of profound intellectual disability, yes/no; and (4) presence of autistic features, yes/no. In each case, we evaluated accuracy, sensitivity, and specificity in predicting the outcome of the dependent variable. Even if PredictMed needs to be fine-tuned and validated on larger data sets, it showed on presently available data sets a predictive performance in terms of accuracy, sensitivity, and specificity between 74% and 90%. In our vision, PredictMed should significantly help doctors’ decision-making process regarding patient prognosis, improving care of patients with CP and reducing healthcare costs and suffering.

**Authors’ note**

We certify that this manuscript is not under review by any other journal. All authors have contributed to and read the manuscript and approved the final copy.

**Declaration of conflicting interests**

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.
**Funding**

The author(s) received no financial support for the research, authorship, and/or publication of this article.

**ORCID iD**

Carlo M Bertoncelli https://orcid.org/0000-0002-9689-0209

**References**

1. Bertoncelli CM, Solla F, Loughenbury PR, et al. Risk factors for developing scoliosis in cerebral palsy: a cross sectional descriptive study. *J Child Neurol* 2017; 32(7): 657–662.
2. Bertoncelli CM, Bertoncelli D, Elbaum L, et al. Validation of a clinical prediction model for the development of neuromuscular scoliosis. *Pediatr Neurol* 2018; 79: 14–20.
3. Bertoncelli CM, Altamura P, Vieira ER, et al. Identifying factors associated with severe intellectual disabilities in teenagers with cerebral palsy using a predictive learning model. *J Child Neurol* 2019; 34(4): 221–229.
4. Bertoncelli CM, Altamura P, Vieira ER, et al. Using artificial intelligence to identify factors associated with autism spectrum disorder in adolescents with cerebral palsy. *Neuropediatrics* 2019; 50(3): 178–187.
5. Vapnik V. *The nature of statistical learning theory*. New York: Springer, 2000.
6. Fisher RA. Statistical methods for research workers. In: Kotz S and Johnson NL (eds) *Breakthroughs in statistics*. New York: Springer, 1992, pp. 66–70.
7. Romano C, van Wynckel M, Hulst J, et al. European Society for Paediatric Gastroenterology, Hepatology and Nutrition guidelines for the evaluation and treatment of gastrointestinal and nutritional complications in children with neurological impairment. *J Pediatr Gastroenterol Nutr* 2017; 65(2): 242–264.
8. Collins GS, Reitsma JB, Altman DG, et al. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. *BMJ* 2014; 350: g7594.
9. Somerville H, Tzannes G, Wood J, et al. Gastrointestinal and nutritional problems in severe developmental disability. *Dev Med Child Neurol* 2008; 50(9): 712–716.
10. Evans PM, Evans SJW and Alberman E. Cerebral palsy: why we must plan for survival. *Arch Dis Child* 1990; 65(12): 1329–1333.
11. Gainsborough M, Surman G, Maestri G, et al. Validity and reliability of the guidelines of the surveillance of cerebral palsy in Europe for the classification of cerebral palsy. *Dev Med Child Neurol* 2008; 50(11): 828–831.
12. Monbaliu E, De La Pena MG, Ortibus E, et al. Functional outcomes in children and young people with dyskinetic cerebral palsy. *Dev Med Child Neurol* 2017; 59(6): 634–640.
13. Goh YR, Choi JY, Kim SA, et al. Comparisons of severity classification systems for oropharyngeal dysfunction in children with cerebral palsy: relations with other functional profiles. *Res Dev Disabil* 2018; 72: 248–256.
14. Hoshino M, Omura N, Yano F, et al. Effects of the body mass index (BMI) on the surgical outcomes of laparoscopic fundoplication for gastro-esophageal reflux disease: a propensity score-matched analysis. *Surg Today* 2017.
15. Grant JP. Comparison of percutaneous endoscopic gastrostomy with Stamm gastrostomy. *Ann Surg* 1988; 207(5): 598–603.
16. Vandenplas Y, Rudolph CD, Di Lorenzo C, et al. Pediatric gastroesophageal reflux clinical practice guidelines: joint recommendations of the North American Society for Pediatric Gastroenterology, Hepatology, and Nutrition (NASPGHAN) and the European Society for Pediatric Gastroenterology, Hepatology, and Nutrition (ESPGHAN). *J Pediatr Gastroenterol Nutr* 2009; 49(4): 498–547.
17. Russell DJ, Avery LM, Rosenbaum PL, et al. Improved scaling of the gross motor function measure for children with cerebral palsy: evidence of reliability and validity. *Phys Ther* 2000; 80(9): 873–885.
18. Baranello G, Rossi Sebastian D, Pagliano E, et al. Hand function assessment in the first years of life in unilateral cerebral palsy: correlation with neuroimaging and cortico-spinal reorganization. *Eur J Paediatr Neurol* 2016; 20(1): 114–124.
19. Cans C. Surveillance of cerebral palsy in Europe: a collaboration of cerebral palsy surveys and registers. *Dev Med Child Neurol* 2000; 42(12): 816–824.
20. Sinha S and Siddiqui KA. Definition of intractable epilepsy. *Neurosciences* 2011; 16: 3–9.
21. Hareb F, Rampal V, Bertoncelli CM, et al. Botulinum toxin in children with cerebral palsy: an update. *Clin Cases Miner Bone Metab*. Epub ahead of print 4 September 2019. DOI: 10.1055/s-0039-1694988.
22. McCarthy JJ, D’Andrea LP, Betz RR, et al. Scoliosis in the child with cerebral palsy. *J Am Acad Orthop Surg* 2006; 14(6): 367–375.
23. Solla F, Clement JL, Doria C, et al. Adolescent idiopathic scoliosis exceeding 70°: a single unit surgical experience. *Minerva Ortop Trauma* 2018; 69(3): 69–77.
24. Stanley RP. What is enumerative combinatorics? In: Stanley RP (ed.) *Enumerative combinatorics*. Boston, MA: Springer, 1986, pp. 1–63.
25. Sullivan K, Dean A and Soe MM. OpenEpi: a web-based epidemiologic and statistical calculator for public health. *Public Health Rep* 2009; 124(3): 471–474.
26. Solla F, Tran A, Bertoncelli D, et al. Why a p-value is not enough. *Clin Spine Surg* 2018; 31(9): 385–388.
27. Blum AL and Langley P. Selection of relevant features and examples in machine learning. *Artif Intell* 1997; 97(1): 245–271.
28. Chen Y-W and Lin C-J. Combining SVMs with various feature selection strategies. In: Guyon I, Nikravesh M, Gunn S, et al. (eds) *Feature extraction* (Bibliography 88). Berlin: Springer, 2006, pp. 315–324.
29. Liu H and Motoda H. *Computational methods of feature selection*. Boca Raton, FL: Chapman & Hall/ CRC, 2007.
30. Tohka J, Moradi E and Huttenen H. Comparison of feature selection techniques in machine learning for anatomical brain MRI in dementia. *Neuroinformatics* 2016; 14(3): 279–296.
31. Mwangi B, Tian TS and Soares JC. A review of feature reduction techniques in neuroimaging. *Neuroinformatics* 2014; 12(2): 229–244.
32. He GW, Wang TY, Chiang AS, et al. Soma detection in 3D images of neurons using machine learning technique. *Neuroinformatics* 2018; 16(1): 31–41.
33. Reza FM. *An introduction to information theory*. New York: Dover Publications, 1961.
34. Peng H, Long F and Ding C. Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy. *IEEE Trans Pattern Anal Mach Intell* 2005; 27(8): 1226–1238.
35. Liu H, Liu L and Zhang H. Feature selection using mutual information: an experimental study. In: *Proceedings of the 10th Pacific Rim international conference on artificial intelligence: trends in artificial intelligence, PRICAI ’08*, Hanoi, Vietnam, 15–19 December 2008, pp. 235–246. Berlin: Springer-Verlag.
36. Chen P, Lin C and Scholkopf B. A tutorial on v-support vector machines. *Appl Stoch Model Bus Ind* 2005; 21(2): 111–136.
37. Buscema M, Grossi E, Snowdon D, et al. Artificial neural networks and artificial organisms can predict Alzheimer pathology in individual patients only on the basis of cognitive and functional status. *Neuroinformatics* 2004; 2(4): 399–416.