

Multi-parameter-based Radiological Diagnosis of Chiari Malformation using Machine Learning Technology

Bora Tetik¹, Güleç Mert Doğan², Ramazan Paşahan¹, Mehmet Akif Durak¹, Emek Güldoğan¹, Kaya Saraç¹, Çağatay Önal¹, and Ismail Okan Yıldırım¹

¹İnönü Üniversitesi

²Malatya Eğitim ve Araştırma Hastanesi

April 19, 2021

Abstract

Background: The known primary radiological diagnosis of Chiari Malformation-I (CM-I) is based on the degree of tonsillar herniation (TH) below the Foramen Magnum (FM). However, recent data also shows the association of such malformation with smaller posterior cranial fossa (PCF) volume and the anatomical issues regarding the Odontoid. This study presents the achieved result regarding some detected potential radiological findings that may aid CM-I diagnosis using several machine learning (ML) algorithms. **Materials and Methods:** Between 2011 and 2020, radiological examinations of 100 clinically/radiologically proved symptomatic CM-I cases and 100 control were evaluated by matching age and gender. A team of Neuroradiologists had reviewed the MR images of the study population. A total of 11 different radiological parameters were assessed for CM-I diagnosis. The parameters were defined and examined in 5 designed different ML algorithms. Statistical analysis was conducted for data analysis. **Results:** The mean age of patients was 29.92 ± 15.03 years. The primary presenting symptoms were headaches (62%). Syringomyelia and retrocurved-odontoid were detected in 34% and 8% of patients, respectively. All of the morphometric measures were significantly different between the groups, except for the distance from the dens axis to the posterior margin of FM. The Radom Forest model is found to have the best 1.0 (14 of 14) ratio of accuracy in regard to 14 different combinations of morphometric features. **Conclusion:** This study indicates the potential usefulness of ML-guided PCF measurements, other than TH, that may be used to predict and diagnose CM-I accurately. Our results support the view of TH as a single radiological parameter may fail during the diagnosis of CM-I. Combining two or three preferable osseous structure-based parameters may increase the accuracy of radiological diagnosis of CM-I.

Multi-parameter-based Radiological Diagnosis of Chiari Malformation using Machine Learning Technology

Abstract

Background: The known primary radiological diagnosis of Chiari Malformation-I (CM-I) is based on the degree of tonsillar herniation (TH) below the Foramen Magnum (FM). However, recent data also shows the association of such malformation with smaller posterior cranial fossa (PCF) volume and the anatomical issues regarding the Odontoid. This study presents the achieved result regarding some detected potential radiological findings that may aid CM-I diagnosis using several machine learning (ML) algorithms.

Materials and Methods: Between 2011 and 2020, radiological examinations of 100 clinically/radiologically proved symptomatic CM-I cases and 100 control were evaluated by matching age and gender. A team of Neuroradiologists had reviewed the MR images of the study population. A total of 11 different radiological parameters were assessed for CM-I diagnosis. The parameters were defined and examined in 5 designed different ML algorithms. Statistical analysis was conducted for data analysis.

Results: The mean age of patients was 29.92 ± 15.03 years. The primary presenting symptoms were headaches (62%). Syringomyelia and retrocurved-odontoid were detected in 34% and 8% of patients, respectively. All of the morphometric measures were significantly different between the groups, except for the distance from the dens axis to the posterior margin of FM. The Radom Forest model is found to have the best 1.0 (14 of 14) ratio of accuracy in regard to 14 different combinations of morphometric features.

Conclusion: This study indicates the potential usefulness of ML-guided PCF measurements, other than TH, that may be used to predict and diagnose CM-I accurately. Our results support the view of TH as a single radiological parameter may fail during the diagnosis of CM-I. Combining two or three preferable osseous structure-based parameters may increase the accuracy of radiological diagnosis of CM-I.

Key Words: Chiari Malformation; Machine Learning; posterior fossa; tonsillar herniation.

What is already known about this topic?

Chiari malformation (CM) is known as the malformation of varying degrees of the descent of the brain stem and cerebellum below the foramen magnum (FM). The most common form of CM is type I, the elongation of the tonsils of the cerebellum into the upper cervical canal. Although this situation is usually presented as an incidental radiological finding, many symptoms may also guide the practitioner to such a diagnosis. The known primary radiological diagnosis of CM-I is relied on the degree of tonsillar herniation (TH) below the FM. However, recent data also shows the association of such malformation with smaller posterior cranial fossa (PCF) volume and the anatomical issues regarding the Odontoid. Recent studies suggest more radiological criteria that may aid to the diagnosis of CM.

What does this article add?

Machine learning (ML) is one of the traditional methods of computer systems that help to recognize and detect patterns in organizational and commercial decisions (e.g., underlying dimensions/subgroups, nonlinear associations/patterns, and so on). This situation can be used in differentiation, classifying, and organizing many health problems. ML offers an opportunity to overcome obstacles using conventional statistical methods. This study presents the achieved results regarding some detected potential radiological findings that may aid CM-I diagnosis using several ML algorithms.

Introduction

Hans Chiari was the first who defined the term Chiari malformation (CM) in 1891, specifying the malformation of varying degrees of the descent of the brain stem and cerebellum below the foramen magnum (FM) (1). The most common form of CM is type I, the elongation of the tonsils of the cerebellum into the upper cervical canal. Although this situation is usually presented as an incidental radiological finding, many symptoms may also guide the practitioner to such a diagnosis (2).

The prevalent belief is that CM-I could result from an intrinsically smaller overcrowded posterior fossa that predisposes the upper cervical spine to eventual herniation of the cerebellar tonsils, resulting in altered dynamics of Cerebrospinal fluid (CSF) flow (3). The known primary radiological diagnosis of CM-I is relied on the degree of tonsillar herniation (TH) below the FM. However, recent data also shows the association of such malformation with smaller posterior cranial fossa (PCF) volume and the anatomical issues regarding the Odontoid (4-7). Recent studies suggest more radiological criteria that may aid to the diagnosis of CM (5, 6, 7, 16, 18, 21).

Machine learning (ML) is one of the traditional methods of computer systems that help to recognize and detect patterns in organizational and commercial decisions (e.g., underlying dimensions/subgroups, nonlinear associations/patterns, and so on) (8). This situation can be used in differentiation, classifying, and organizing many health problems. ML offers an opportunity to overcome obstacles using conventional statistical methods.

This study presents the achieved results regarding some detected potential radiological findings that may aid CM-I diagnosis using several ML algorithms.

Materials and Methods

Patient and Study Characteristics

Between 2011 and 2020, the Neurosurgery Department database in Malatya Inonu University, School of Medicine was searched to evaluate 100 patients with clinically and radiologically proved symptomatic CM-I diagnosis. Also sex and age were matched to patients with normal MRI reports to form the control group. Patients with missing clinical data or with a history of prior cranial surgery were excluded from the study.

The study was approved by the Ethical Committee of Malatya Clinical Researches in Malatya, Turkey (Protocol number: 2021/104). Since the patients were retrospectively chosen and the MR images examined were fully de-identified, the analysis did not require any informed consent.

Magnetic Resonance Imaging (MRI) Protocol

Three experienced Neuroradiologists who are blinded to patients' data reviewed the MR images (1.5 T; Siemens, Magnetom-Avantou, Germany) of all patients in the study, and 10 measurements other than TH were obtained from this study. On repeated tests, high intraobserver reliability was demonstrated (with 95 percent confidence intervals) in 20 randomly selected situations. Both the CM-I (n=100) and control (n=100) groups were evaluated regarding the axial, sagittal and coronal T1 and T2 weighted MRIs. The measurements were obtained manually from the mid-sagittal images. These measurements consisted of; F1: Tonsillar herniation; F2: Distance from the pons to FM; F3: Distance from corpus callosum to FM; F4: Distance from Fastigium to FM; F5: Distance from the pons to dens axis; F6: Distance from corpus callosum to dens axis; F7: Distance from Fastigium to dens axis; F8: Distance from dorsum sella to dens axis; F9: Clivus length; F10: Distance from dens axis to posterior margin of FM; F11: Clivus canal angle (Weckenheim line- posterior vertebral body line) (Figure 1).

Supervised Machine Learning Models

In the current study, each data point in the training set includes a vector (X, Y), where X denotes the feature vector of (x_1, x_2, \dots, x_n) and Y expresses the output feature that takes values 0 (absence of CM) and 1 (presence of CM). The following five supervised ML models were constructed on the feature vectors of several sizes (1, 2, ..., 11). Single and the combined features achieved from the best performed classifiers were constructed for each measurement.

XGBoost (Extreme Gradient Boosting) algorithm: Xgboost is an algorithm that employs a gradient boosting decision tree and can compute boosted trees efficiently in parallel. XgBoost divides its tree models into regression and classification trees (9). The biggest advantage of XGBoost is its scalability in all situations. The method for measuring timing helps to simulate distributed real-time applications with millions of simultaneous connections. The efficacy of XGBoost is owing to various well-differentiated systems and algorithmic optimizations. A novel tree learning method is used to handle sparse data, and a technically justified weighted quantile sketch technique is implemented in approximate tree learning to handle instance weights (10). The details of the XGBoost can be achieved from the relevant study (10).

Stochastic Gradient Boosting (SGB): Gradient boosting is a supervised machine learning approach for regression and classification issues in the SGB algorithm, which creates a predictive model using a collection of weak learners/classifiers, usually decision trees. SGB builds the model in a stage-wise manner and constructs additive regression models by fitting a simple parameterized function (base learner) to the current pseudo-residuals by at least squares at each iteration. The pseudo-residuals are the minimization of the functional loss gradient. It's worth noting that integrating randomization into the technique improves gradient boosting's approximation accuracy and execution speed. For prediction models, the SGB parameters (depth of interaction, tension number, and shrinkage) were optimized as soon as possible (11). The entire information on the SGB was reported by the related study (12).

Bagged classification and regression trees (Bagged CART): CART has been widely utilized for ML modeling with adequate results, including potential prediction of the diseases in health field. As the CART is regarded

as an unstable model, the bagging method can significantly increase its precision. The bagged CART effectively eliminates the variance in estimation and greatly improves accuracy and over-fitting of classification. Thus, it is predicted that promising results can be obtained by employing the bagged CART in the new application of potential classification problems (13). A comprehensive explanation for the bagged CART was given in the concerned reference (14).

Random Forest (RF): An ensemble method, which constructs several decision trees that will be used by the majority vote to classify a new case, is Random Forest. A subset of attributes selected randomly from the whole original set of attributes is used by each decision tree node. Additionally, in the same way as bagging, each tree uses a different bootstrap sample data. Besides, Random Forest is a technique that is computer-efficient and can run rapidly over large datasets. It has been used in diverse domains in many recent research projects and real-world applications (14). The details of the Random Forest was introduced with the relevant paper (15).

Logistic Regression (LR): Logistic regression analysis is used to obtain an "odds ratio" that incorporates two or more covariates/features. The process is similar to multiple linear regression, except that the response attribute is predicted from a binomial distribution. The data and the results of each variable indicate the strength of each predictor on the observed outcome of interest. This method can be used to estimate the probability $P(Y = 1 | X)$ for an observation point of X and considers a different transformation of a linear combination for the input features. Here, $(Y = 1 | X) = \frac{1}{1 + e^{-(b_0 + b_1x_1 + \dots + b_nx_n)}}$, where b_0, b_1, \dots, b_n are parameters of the model, which are calculated from the training stage with maximum likelihood technique (16,17).

Validation and performance evaluation

Validation of the algorithms were performed using five-fold cross-validation method. Hyperparameters of the interested models were tuned by random search optimization in the current study. We computed the following performance metrics to measure the quality of a classifier using five -fold cross-validation: Accuracy, sensitivity, specificity, positive/negative predictive values, F-measure, and G-mean. All performance metrics were calculated through DTROC software. The supervised ML algorithms mentioned earlier were implemented using R Studio of R programming language.

Statistical Analysis

IBM SPSS Statistics 26.0 software was utilized for the analysis. Quantitative data were given as mean, standard deviation, %95 confidence intervals (CI), and qualitative data were given as numbers (percentages). Conformity to normal distribution was made using the Shapiro-Wilk test. Mann-Whitney U test, Independent Sample t-Test, and Spearman Chi-square test were used in statistical analyzes as appropriate. All results were reported in line with APA style. A p-value of <0.05 was regarded as statistically significant.

Results

A total of 200 radiological examinations was evaluated in this work. Table 1 represents the baseline characteristics of the CM-I group. The mean age was 29.92 +- 15.03 years, and the study population consisted of 55 females and 45 males. The primary presenting symptoms were headache (62%), neck pain (30%), sensorial disturbance (18%), and paraparesis (15%). Syringomyelia and retro-odontoid were detected in 34% and 8% of patients, respectively.

The measured parameters were identified to ML algorithms, and all of the morphometric measures were significantly different between the groups, except for the distance from dens axis to posterior margin of FM (F10). Table 2 represents the statistics regarding the measured parameters.

For the first parameter (F1) XGBoost, SGB, Bagged CART, Random Forest, and Logistic Regression values were found to be 100% as expected. For the other parameters (F2-11), the most accurate values were 0.88 for F6 and F10 (Random Forest model), 0.87 for F5 (Random Forest model), and 0.86 for F2 (Random

Forest model). The accuracy of the Logistic Regression model for F2-11 was calculated as 0.89, whereas it was 100% for other models. Detailed data are presented in Table 3.

The results indicate that the Random Forest model has produced the best 1.0 (14 of 14) ratio of accuracy regarding 14 different combinations of morphometric features. This ratio is found to be 0.50, 0.29, 0.21, and 0.07 in the Bagged CART, Stochastic Gradient Boosting, XGBoost, and Logistic Regression models, respectively.

Discussion

Recent clinical and experimental studies have underlined the TH in CM-I due to embryological developmental anomalies in the occipital bone and a hypoplastic PCF (5,6,7,18,19). However, tonsillar herniation as a single radiological diagnostic criterion for CM-I is insufficient and more parameters are needed to prevent under or misdiagnosis. A reported study by Alperin et al. identified four measures that distinguished CM-I patients from a healthy cohort with 97 percent sensitivity and 100 percent specificity (4). Another study produced a probability predictor based on a logistic regression (LR) model that took into account seven PCF measures and had a sensitivity of 93% and specificity of 92% in distinguishing patients with classic CM-I from those with a standard PCF (16).

Many studies have searched for alternatives or adjunct diagnostic radiological parameters for CM-I (4,5,16,20-23). In our study, a large CM-I population was compared with a sex-age matched control group using 5 different supervised ML programs. The primary aim was to detect the test with the most accurate result and minimal parameters.

Machine learning is a potential diagnostic tool that may be used in many disorders. The quality of each classifier was tested using three-fold cross-validation. Also, calculation of the variable importance for the features can be divided into model-based and model-independent approaches. The benefit of using a model-based method is that it is more closely related to the performance of the model and that the association structure between the predictors can be integrated into the estimation of significance. Based on the advantages of model-based methods, we used this approach for calculating the variable importance in the current study.

It has been shown that cross-validation techniques significantly improve the outcome of the models by helping to choose the best coefficients taken into consideration for the model (24). To boost the efficiency of the models, k-fold cross-validation is employed in the training stage. In the present study, three-fold cross-validation was utilized for assessing the predictive power of the supervised ML models. The random search optimization approach deals with techniques for a given problem to achieve the 'most favorable' solution and selects the most relevant features with the output variable.

Bagged CART (F4&F11) and Random Forest (F2&F11) methods have revealed a success rate of 98% using two different parameters. This ratio increased to 100% when a third parameter was included in the analysis. Also, the diagnostic ratio of the XGBoost method was 99% using three parameters (F2, F4, and F6) but increased to 100% when a fourth parameter was included. These results reflect the usefulness and accuracy of such ML systems. The Random Forest method is found to be the most accurate test using the parameters of distance from the pons to FM (F2); fastigium to FM (F4), and Clivus canal angle (F11). This outcome supports the latest report by Urbizu et al. but with augmented accuracy (16). The high diagnostic ratio of our methods seems to be due to the matched control group and the use of advanced ML algorithms.

It is known that the posterior basicranium undergoes retroflexion during the fetal phase, pushing the basion upward and dorsally and decreasing the PCF ventral depth (25). Any developmental problem during this phase, as the spheno-occipital synchondrosis closes, may impede the normal growth of the basilar portion of the occipital bone until the 2nd decade of life (25). In our study, the distance from Fastigium to FM (F4) was the standard parameter that was used and obtained higher accuracy. This result may support the suggested role of the decreased ventral depth of PCF and hypoplastic PCF in CM-I etiology (2,5,6,7,18,21,25).

It was suggested that the basion and the dens might grow in the embryo jointly (sclerotome resegmentation), supporting the theory that a shift in the location of the basion could result in a concomitant change in the

dens position, leading to more craniocervical changes in these patients (26). This relation was concluded as a necessity to include the Odontoid in such measurements (16). In our study, the Odontoid was included as a parameter but provided different values of accuracy that in Random Forest test no change was observed; but in Bagged CART test, the value of accuracy was increased and decreased in different combinations. Also, inclusion of clival and Corpus Callosum measurements was found to increase the classification accuracy. The crowded PCF and CSF dynamics may decrease the reliability of mobile points like pons when compared with osseous structures like Clivus, dorsum sellae and odontoid. These findings should be evaluated in cooperation with literature-based data that support the role of crowded PCF in the etiology of CM-I (5,27,28,29).

The optimal surgical management of CM-I is still controversial. Posterior fossa decompression with or without duroplasty is the most recommended surgery (28, 30-32). However, our results may guide further studies that anteriorly localized structures seem to influence the PCF and etiology of CM-I as much as posterior ones. This may also aid different surgical approaches and management ways.

Regarding the limitation of the study, even statistically matched, the dominance of female cases in the study population might be considered as a minor bias. A larger and more homogenous study population may be more appropriate. Also, TH was calculated using midsagittal MR images, excluding any possible asymmetry regarding parasagittal structures that may influence symptomatology. Larger studies may be designed in order to exclude any possible difference in race and regional factors that may have an effect on morphometric anatomy.

Conclusion

Our study indicates the potential usefulness of ML-guided PCF measurements, other than TH, that may be used to predict and diagnose CM-I accurately. Our results support the view of TH as a single radiological parameter may fail during the diagnosis of CM-I. Combining two or three preferable osseous structure-based parameters may increase the accuracy of radiological diagnosis of CM-I.

References

1. Chiari HJD-DMW. Ueber Veränderungen des Kleinhirns infolge von Hydrocephalie des Grosshirns. *Dtsch Med Wochenschr* 1891; 17: 1172-1175
2. Milhorat TH, Chou MW, Trinidad EM et al. Chiari I malformation redefined: clinical and radiographic findings for 364 symptomatic patients. *Neurosurgery* 1999 May;44(5):1005-17.
3. Siri Sahib S Khalsa, Alan Siu et al. Comparison of posterior fossa volumes and clinical outcomes after decompression of Chiari malformation Type I. *Journal of Neurosurgery Pediatrics* 2017; 19: 511-517.
4. Bagci AM, Lee SH, Nagornaya N et al. Automated posterior cranial fossa volumetry by MRI: applications to Chiari malformation type I." *American Journal of Neuroradiology* 2013 34: 1758-1763.
5. Dagtekin A, Avci E, Kara E, et al. Posterior cranial fossa morphometry in symptomatic adult Chiari I malformation patients: comparative clinical and anatomical study. *Clin Neurol Neurosurg* 2011;113: 399-403.
6. Nishikawa M, Sakamoto H, Hakuba A et al. Pathogenesis of Chiari malformation: a morphometric study of the posterior cranial fossa. *J Neurosurg* 1997; 86: 40-47.
7. Noudel R, Jovenin N, Eap C et al. Incidence of basioccipital hypoplasia in Chiari malformation Type I: comparative morphometric study of the posterior cranial fossa. *J Neurosurg* 2009; 111: 1046-1052.
8. Jiang T, Gradus JL, Rosellini AJ. Supervised Machine Learning: A Brief Primer. *Behavior Therapy*. 2020; 51: 675-87.
9. Zheng H, Yuan J, Chen LJE. Short-term load forecasting using EMD-LSTM neural networks with a Xgboost algorithm for feature importance evaluation. *Energies* 2017; 10: 1168.
10. Chen T, Guestrin C, editors. Xgboost: A scalable tree boosting system. Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining; August 2016. 785-794

11. Xu Q, Xiong Y, Dai H, Kumari KM, Xu Q, Ou H-Y, et al. PDC-SGB: Prediction of effective drug combinations using a stochastic gradient boosting algorithm. *J Theor Biol* 2017; 21;417:1-7.
12. Friedman JH. Stochastic gradient boosting. *Computational Statistics & Data Analysis* 2002; 38: 367-78.
13. Mosavi A, Hosseini FS, Choubin B, Goodarzi M, Dineva AA, Sardooi ERJWRM. Ensemble Boosting and Bagging Based Machine Learning Models for Groundwater Potential Prediction. *Water Resources Management* 2021; 35: 23-37
14. Sutton CD. Classification and regression trees, bagging, and boosting. *Handbook of Statistics* 2005;24: 303-29.
15. Breiman L. Random forests. *Machine Learning*. 2001; 45: 5-32.
16. Urbizu A, Martin BA, Moncho D et al. Machine learning applied to neuroimaging for diagnosis of adult classic Chiari malformation: role of the basion as a key morphometric indicator. *J Neurosurg*. 2018; 129: 779-791.
- 17 Colak C, Colak MC, Orman MN. The comparison of logistic regression model selection methods for the prediction of coronary artery disease. *Anadolu Kardiyol Derg*. 2007; 7: 6-11.
18. S. Aydin, H. Hanimoglu, T. Tanriverdi et al. Chiari type I malformations in adult: a morphometric analysis of the posterior cranial fossa *Surg Neurol* 2005; 64:237-241
19. F. Karagoz, N. Izgi, S. Kapicioglu Sencer. Morphometric measurements of the cranium in patients with Chiari type I malformation and comparison with the normal population *Acta Neurochir (Wien)* 2002; 144: 165-171
20. Halvorson KG, Kellogg RT, Keachie KN et al. Morphometric Analysis of Predictors of Cervical Syrinx Formation in the Setting of Chiari I Malformation. *Pediatr Neurosurg* 2016; 51:137-141
21. Aydin S, Hanimoglu H, Tanriverdi T, et al. Chiari type I malformations in adults: a morphometric analysis of the posterior cranial fossa. *Surg Neurol* 2005; 64: 237- 241.
22. Karagoz F, Izgi N, Kapijicijoglu Sencer S. Morphometric measurements of the cranium in patients with Chiari type I malformation and comparison with the normal population. *Acta Neurochir (Wien)* 2002; 144: 165- 171.
23. Trigylidas T, Baronia B, Vassilyadi M, et al. Posterior fossa dimension and volume estimates in pediatric patients with Chiari I malformations. *Childs Nerv Syst* 2008; 24: 329- 336.
24. Saud S, Jamil B, Upadhyay Y, Irshad KJSET. Performance improvement of empirical models for estimation of global solar radiation in India: A k-fold cross-validation approach. *Sustainable Energy Technologies and Assessments* 2020; 40: 100768.
25. Krishan K, Kanchan T: Evaluation of spheno-occipital synchondrosis: A review of literature and considerations from forensic anthropologic point of view. *J Forensic Dent Sci* 2013; 5:72-76,
26. Pang D, Thompson DN: Embryology and bony malformations of the craniovertebral junction. *Childs Nerv Syst* 2011; 27: 523-564
27. Cesmebasi A, Loukas M, Hogan E et al. The Chiari malformations: a review with emphasis on anatomical traits. *Clin Anat* 2015; 28: 184-194,
28. Chotai S, Medhkour A. Surgical outcomes after posterior fossa decompression with and without duraplasty in Chiari malformation-I. *Clin Neurol Neurosurg* 2014; 125:182-188
29. Isik N, Elmaci I, Kaksi M et al. A new entity: Chiari Zero malformation and its surgical method. *Turk Neurosurg* 2011; 21: 264-268

30. Forander P, Sjavik K, Solheim O et al. The case for duraplasty in adults undergoing posterior fossa decompression for Chiari I malformation: a systematic review and meta-analysis of observational studies. Clin Neurol Neurosurg 2014; 125: 58–64
31. Tubbs RS, Oakes WJ. The Chiari Malformations. J Neurosurg. 2007; 106: 329-330.
32. Zhao JL, Li MH, Wang CL, Meng W: A systematic review of Chiari I malformation: techniques and outcomes. World Neurosurg 2016; 88: 7–14

Table 1. Basic clinical characteristics of the CM-I group

Demographics

Characteristic	Mean ± SD	Mean ± SD
Age at diagnosis	29.92 ± 15.03	29.92 ± 15.03
Characteristic	n	%
Male	55	55.0
Female	45	45.0
Signs and Symptoms	Signs and Symptoms	Signs and Symptoms
Characteristic	Mean ± SD	Mean ± SD
Time elapsed from onset (mo)	111.33± 118.75	111.33± 118.75
Characteristics	n	%
Hydroxyringomyelia	34	34.0
Retrocurved odontoid	8	8.0
Headaches	62	62.0
Occipital headache/cervicalgia	30	30.0
Dizziness	9	9.0
Vertigo	8	8.0
Visual alterations	3	3.0
Kyphoscoliosis	8	8.0
Sensory loss	18	18.0
Motor weakness	12	12.0
Gait disturbances	13	13.0
Paresthesia/pain upper limbs	13	13.0
Paresthesia/pain lower limbs	15	15.0
Therapeutic procedures	Therapeutic procedures	Therapeutic procedures
Characteristic	n	%
Surgical treatment	50	50.0

Table 2. Baseline characteristics of the study sample per the groups

Feature	Group	Group	Group	Group
	Non-CM-I n=100	Non-CM-I n=100	Non-CM-I n=100	Non-CM-I n=100
	Mean	SD	95.0% Lower CL for Mean	95.0% Upper CL for Mean
F ₁	-3.1 _a	2.1	-3.5	-2.7
F ₂	40.5 _a	4.2	39.7	41.4
F ₃	60.8 _a	3.6	60.1	61.5
F ₄	27.0 _a	2.9	26.4	27.6
F ₅	49.2 _a	4.6	48.3	50.1
F ₆	65.8 _a	3.8	65.1	66.6
F ₇	30.0 _a	3.4	29.4	30.7
F ₈	47.5 _a	4.5	46.6	48.4

Feature	Group	Group	Group	Group
F ₉	37.8 _a	3.1	37.2	38.5
F ₁₀	36.6 _a	3.8	35.8	37.3
F ₁₁	158 _a	6	157	159
Age	29.9	15.0	27	33
Gender	n	%		
1	45	45.0		
2	55	55.0		

Values in the same row and subtable not sharing the same subscript are significantly different at $p < .05$ in the two-sided test of equality for column means by APA style.

Table 3: Accuracy values of single and combined morphometric measurements for the best performed classifiers

Feature(s)	XGBoost
F ₁	1
F ₂	0.845
F ₃	0.775
F ₄	0.83
F ₅	0.78
F ₆	0.805
F ₇	0.77
F ₈	0.785
F ₉	0.79
F ₁₀	0.84
F ₁₁	0.81
Accuracy of the combined features achieved from the best performed classifiers	Accuracy of the combined features achieved from the best performed classifiers
2	0.99
3	0.955
4	1
5	0.995
6	0.995
7	1
Features combination for the best performed classifiers	Features combination for the best performed classifiers
2	4,2
3	4,2,6
4	4,2,6,5
5	4,2,6,5,9
6	4,2,6,5,9,7
7	4,2,6,5,9,7,8

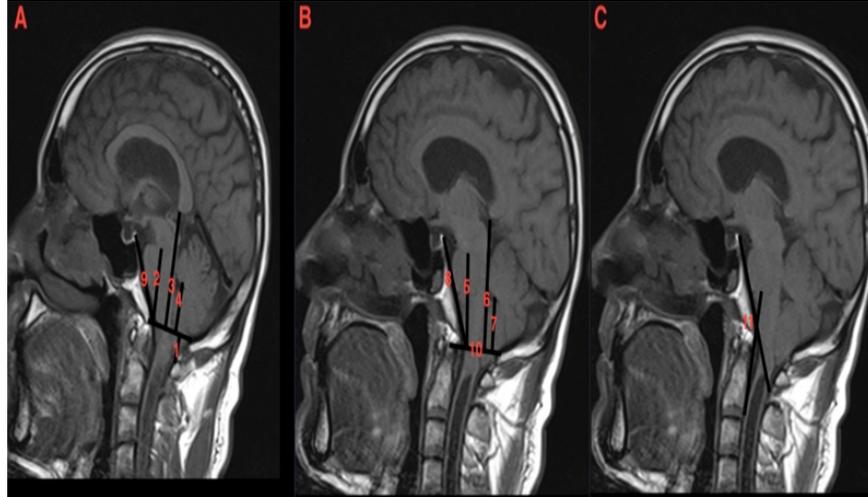
Figure Legends

Figure 1. Sagittal T1-weighted MR images show the measurements of structures related to posterior cranial fossa morphology. The numbers represent the measured distance: 1. Tonsillar herniation; 2. Distance from pons to FM; 3. Distance from corpus callosum to FM; 4. Distance from Fastigium to FM; 5. Distance from pons to dens axis; 6. Distance from corpus callosum to dens axis; 7. Distance from Fastigium to dens axis; 8. Distance from dorsum sellae to dens axis; 9. Clivus length; 10. Distance from dens axis to posterior margin of FM; 11. Clivus canal angle (Weckenheim line- posterior vertebral body line).

Table 1. Basic clinical characteristics of the CM-I group

Table 2. Baseline characteristics of the study sample per the groups

Table 3: Accuracy values of single and combined morphometric measurements for the best performed classifiers



Hosted file

Table 1.pdf available at <https://authorea.com/users/408787/articles/518653-multi-parameter-based-radiological-diagnosis-of-chiari-malformation-using-machine-learning-technology>

Hosted file

Table 3.pdf available at <https://authorea.com/users/408787/articles/518653-multi-parameter-based-radiological-diagnosis-of-chiari-malformation-using-machine-learning-technology>