# ABSTRACT

Background:

Ommaya reservoir can be used to treat posthemorrhagic hydrocephalus secondary to intraventricular hemorrhage of prematurity until an acceptable weight can be obtained to place a permanent shunt. Identifying newborns at higher risk of developing shunt conversion may improve the management of these patients.

Objective:

Develop a predictive algorithm for the conversion of an Ommaya reservoir to a permanent shunt using artificial intelligence techniques and “classical” statistics.

Methods:

Database of 43 preterm patients weighing ≤ 1500g with posthemorrhagic hydrocephalus (Papile Grades III and IV with Levene index > 4 mm above the 97th percentile) managed with Ommaya reservoir at our institution between 2002 and 2017 was used to train a KNN algorithm. Validation of results with cross validation technique. Three scenarios were calculated. 1: considering all features regardless whether or not they are correlated with the output variable. 2: consider the features as predictors if they have a correlation greater than a 30% with the output variable. 3: consider the output of the previous analysis.

Results:

Results show that when considering the outputs of a previous multivariate analysis the algorithm reaches an 86% of cross validation accuracy.

Conclusion:

The use of machine learning-based algorithms can help in the early identification of patients with permanent need of shunt. We present the development of a predictive algorithm for permanent shunt with an accuracy of 86%, accuracy of the algorithm can be improved with larger volume of data and previous analysis.

# SHORT TITLE:

ML algorithm for preterm hydrocephalus.

# KEY WORDS

Machine Learning

Preterm hydrocephalus

Cephalospinal fluid shunt

Intraventricular Hemorraghe

Ommaya

Algorithm

Neurosurgery

# Development of machine learning-based predictor algorithm for conversion of an Ommaya reservoir to a permanent cerebrospinal fluid shunt in preterm posthemorrhagic hydrocephalus.

# INTRODUCTION

Intraventricular hemorrhage (IVH) is a serious and common pathology in premature infants. Bleeding begins in the fragile subependymal capillary network of the germinal matrix, a richly vascularized collection of neuronal-glial precursor cells in the developing brain1. IVH is the most frequent cause of acquired hydrocephalus2. This situation can lead to various delays in motor, language, and cognition development3.

Permanent shunt can cause severe complications in patients weighting ≤ 1500g4,5,6. Therefore different options have been described to treat posthemorrhagic hydrocephalus until an acceptable weight is reached. Repeated lumbar taps, ventricular taps and placement of an external ventricular drain are associated with high rates of shunt infection and have been shown less effective in reducing the need for permanent shunts than ventricular reservoirs (Ommaya) and ventriculosubgaleal shunting7,8,9. With no evidence in the superiority between the two techniques, presenting shunt conversion rates of 65%–95%11,12.

Permanent shunting at long-term can lead to a greater degree of neurological abnormalities, a smaller cranial circumference, not mentioning the important rate of shunt revision 4,8,11,13. Selection of patients requiring permanent shunting can lead to an earlier shunt implantation and avoid unnecessary taps that result in complications14,15.

Algorithms based on machine learning (ML) are set to occupy an important place in medicine, changing certain jobs for the better16. This is a reality in many companies that use algorithms for their business strategies. In medicine, the sensitivity of the data, the exquisite precision that algorithms must have and the mistrust or lack of knowledge of healthcare professionals have slowed down their implementation17,18.

The objective of this study is to develop a predictive algorithm for the conversion of an Ommaya reservoir to a permanent shunt using the power of artificial intelligence (AI) through ML. In a second phase, an external centre has been contacted to provide more data for the external validation of the algorithm.

# METHODS

This study was carried out following the transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD) guidelines. Informed consent was obtained for each patient. Ethical Committee approval from our institution was obtained.

Although it may be complex for a clinician, we present below the methodology, mandated by the TRIPOD guidelines, that we have followed for the development of the algorithm to serve as an example for the development of future algorithms.

Data were obtained from a database of our hospital records (tertiary hospital) of 43 preterm neonatal patients, gestational age ≤ 34 week, with low birth weight, ≤ 1500 grams, with presence of IVH classified as Papile grade III or IV, with progressive hydrocephalus with frontal Levene index > 4 mm above the 97th percentile who required placement of an Ommaya reservoir device between 2002 and 2017. Those with Ommaya-associated infections were excluded.

Treatment algorithm is summarized in Figure 1.

Demographic and clinical variables were gathered, as well as the rate of permanent shunt placement. Clinical worsening was defined as a bulging fontanelle, bradycardia, or arterial hypertension. Poor sonographic response was defined as a ventricular index greater than the preoperative value. Some cases underwent a lumbar tap as a temporizing measure before placement of an Ommaya reservoir.

Availability of labelled patient data makes especially interesting to consider Supervised ML techniques. Patient is classified as 0 when they do not need a permanent shunt and 1 when they do.

The k-Nearest Neighbors (kNN) algorithm was used the output of the algorithm is the class that is the most frequent among the selected neighbors (0 or 1):

* K was initialized. This means selecting the number of neighbors that will be considered.
* For each example in the data:
  + Distance between this point and the rest of points from the dataset was calculated. Storing these distances would result in a vector of size (1,n), being n the number of examples in the data set
  + Vector was sorted in ascending order
  + First k positions were selected
  + Mode was taken from label assigned to the selected examples

A series of steps were followed. First step is to load the dataset and analyze the distribution of the variables. Variables that were not of interest as predictors were discarded, missing values are imputed, and outliers are detected.

For variables with a higher percentage of outliers in their data, a logarithmic transformation was applied to smooth these variables.

Variables were standardized as having variables with such different magnitudes can penalize the performance of ML algorithms. Standard scaler was used for this purpose.

Correlation between the features and the output variable was calculated to find those that were candidates to be predictors for the model by having closer relationship with the output.

After selection of the variables, the dataset was split into a train set and a test set. Training set was used to fit the parameters of the model. Amount of data chosen for this test should be greater than testing data to contemplate as many cases as possible. This is key in ML because in a classification problem the training should contain data from all the groups that may exist, since if a group is not seen, it will never be predicted correctly.

Test dataset is independent of the training dataset, but follows the same probability

distribution as the training dataset as they belong to the same population.

The same data used to build the model cannot be used in order to evaluate it, because the model could simply memorize the data instead of learning, so any data of the training set would be perfectly predicted or classified but it would not have any generalization. For this reason, a part of the data was saved and the training never saw this data. Once the data was trained, the test is to predict the target variable..

Accuracy will be used to evaluate the performance of the model, to evaluate the accuracy , the target variable with the prediction is compared. Accuracy is defined as:

Obtaining good accuracy doesn’t mean that the algorithm has understood the problem since it could be that the distribution of train and test was favourable to obtain a better accuracy.

For validating the model in a more robust way, the cross-validation technique was used. Dataset was dived into k-folds and each of them is the test set at some point. An example is shown in Table 1. for 4 folds D1,D2,D3 and D4 applied to a KNN algorithm:

The accuracy assigned for each number of neighbors considered (k parameter) is the average of the 4 accuracies of the different datasets resulting from cross validation.

Number of neighbors to be considered and number of folds for cross validation were the ones that achieved the best accuracy using a grid search algorithm. Grid search computes the calculations for all the values specified in the grid. If k\_fold grid is [3,4,5] and k\_neighbors = [3,4,5,6,7,8,9,10] the kNN algorithm will be calculated first using 3 folds (2 for training and one for validation) for all the specified k\_neighbors values, then with 4 folds, and so on with 5 folds. All possible combinations of these parameters were tested and a given accuracy was obtained. The values with the highest accuracy were selected.

Regarding sample size, no other previous research was available for the analysis.

To impute the empty values of the dataset for categorical variables, variables distribution were analyzed with respect to the output variable, the value that is most frequent for that output value was assigned. For continuous variables empty data was filled with the mean of the distribution.

The correlation between variables and the output class was calculated to obtain those that could be most closely related. When selecting the predictors for the model, the fact of having more or fewer predictors affected the performance of the model was analyzed and the correlation values obtained previously would be those on which the selection would be based.

Three different scenarios were compared:

Scenario 1: consider all features regardless of whether they are correlated or not with the output variable

Scenario 2: consider the features as predictors if they have a correlation greater than a 30% with the output variable

Scenario 3: consider the output of previous analyses associated with higher likelihood of permanent CSF shunting19; high CSF lactate levels, absence of symptomatic patent ductus arteriosus, and higher CSF extraction requirement.

Cross validation accuracy was evaluated for the three scenarios and the features involved in the scenario with highest accuracy were considered the selected predictors.

# RESULTS

43 patients met the inclusion criteria, 32 underwent permanent shunting and 11 didn’t show poor sonographic response or clinical worsening and therefore Ommaya reservoir was removed.

Dataset contained 27 variables (Table 2). Some variables had to be eliminated. Poor clinical response during the tapings was eliminated because of very high percentage of empty data and no significant correlation with the output variable. Birth weight percentile and ventricular Levene's index percentile were discarded because of higher empty values than the variables without percentile. Subsequently dataset was reduced to (43,23).

Pearson's correlation of the variables respect to the output was calculated to consider them as predictors as explained in methods section. (Table 3).

Of the KNN algorithm, only the hyperparameter k, the number of neighbors considered to establish the output of the algorithm, was tuned. For this purpose, a grid search between different values of parameter k was established. Considering that the dataset has 43 possible neighbors, the grid was established in all those values of k between 3 and 20. The differences between cross validation with 3 folds and 4 folds were also tested.

Algorithm’s prediction is simple to interpret, 1 means that patient is considered to need permanent shunting and zero that not. Accuracy of this prediction was calculated for each scenario and number of folds (Table 4-6).

Best results were always associated to 4 folds. A configuration of 14 neighbors for the third scenario was the best case achieving an 86% of cross-validation accuracy as shown on table 6. Therefore, by knowing the variables that have the greatest influence on the possibility or not of ending up with a shunt, with previous analysis19, and applying a cross-validation method in which data is divided for training and validation, a higher accuracy is obtained.

This means that physicians would simply enter the required patient data in Table 7 and the algorithm, adjusted by these parameters, will predict with an accuracy of 86% whether a child whether it will need a shunt or not.

# DISCUSSION

AI techniques such as ML and its algorithms are being developed in multiple fields of medicine. An example in neurosurgery is the development of a convolutional neural network with an AUC of ,846 for the diagnosis of intracranial hemorrhage20.

The first limitation of this study is that we have a relatively small dataset with missing values, this is a problem in medicine where it is difficult to obtain a large volume of reliable data. Big data refers to datasets or combinations of datasets whose size (volume), complexity (variability) and speed of growth (velocity) make it difficult to capture, manage, process or analyse using conventional technologies and tools. In these situations of large data volume ML can be of great help. However, ML is not exclusive of big data and can also be used with smaller datasets.

Other limitation is that in this article we have only developed the algorithm, therefore is only reliable in patients from our dataset, its accuracy should be validated with external data. This generates a problem of generability. An accuracy of 86% suggests there is not much overfitting, however external validation would be again necessary 21. We have a second phase pending with another center for external validation with another dataset.

There are different types of bridging treatments until the shunt can be placed7,8,9, with conversion rates of 65%–95%10,11,12 High variability in the management of these patients, increases heterogeneity and limits the applicability of the study

Variables with the greatest weight in the Pearson correlation analysis were the preoperative lactatorrhachia, extraction rate and patent ductus arteriosus this is consistent with the findings previously known from the work of Palpán et al19.

Despite limitations we present an algorithm with an accuracy of 86%. Ommaya reservoir is a widely used device for the management of hydrocephalus22. Also the rest of the variables used for the calculation of this algorithm are variables that are easily obtainable and available. KNN algorithm is arguably the simplest ML algorithm after linear regression. These facts improve algorithm’s implementation.

IVH has a strong social23 and economic impact24. Rate of Ommaya reservoir infection can be associated with tap frequency. Identifying patients at increased risk of developing hydrocephalus can help preventing these complications, improving patient function, and reducing the impact on families and healthcare-system19.

Finally, this is the first study of these characteristics in IVH of prematurity and the first algorithm developed by neurosurgeons in the field of paediatric neurosurgery. We strongly believe this is a field of future and that spreading these algorithms helps to improve and refine them. The main strength of this study is not the accuracy, but that it serves as a starting point for the development of future algorithms and research. “Classical" statistical tests cannot be forgotten, our algorithm shows the highest accuracy rate in scenario 3, when the outputs of previous analyses are known in advance. This suggests that both techniques are complementary and should be used to enhance each other rather than compete.

# CONCLUSION

The use of ML-based algorithms can help in the early identification of patients with permanent need of shunt. We present a predictive algorithm for permanent shunt with an accuracy of 86%, the accuracy of such an algorithm can be improved with a larger volume of data and previous analysis. A validation study is needed to probe algorithm’s accuracy.

# REFERENCES:

1. Egesa WI, Odoch S, Odong RJ, Nakalema G, Asiimwe D, Ekuk E, Twesigemukama S, Turyasiima M, Lokengama RK, Waibi WM, Abdirashid S, Kajoba D, Kumbakulu PK. Germinal Matrix-Intraventricular Hemorrhage: A Tale of Preterm Infants. Int J Pediatr. 2021 Mar 16;2021:6622598. doi: 10.1155/2021/6622598. PMID: 33815512; PMCID: PMC7987455.
2. Dorner RA, Burton VJ, Allen MC, Robinson S, Soares BP. Preterm neuroimaging and neurodevelopmental outcome: a focus on intraventricular hemorrhage, post-hemorrhagic hydrocephalus, and associated brain injury. J Perinatol. 2018 Nov;38(11):1431-1443. doi: 10.1038/s41372-018-0209-5. Epub 2018 Aug 30. PMID: 30166622; PMCID: PMC6215507.
3. Valdez Sandoval P, Hernández Rosales P, Quiñones Hernández DG, Chavana Naranjo EA, García Navarro V. Intraventricular hemorrhage and posthemorrhagic hydrocephalus in preterm infants: diagnosis, classification, and treatment options. Childs Nerv Syst. 2019 Jun;35(6):917-927. doi: 10.1007/s00381-019-04127-x. Epub 2019 Apr 5. PMID: 30953157.
4. Christian EA, Melamed EF, Peck E, Krieger MD, McComb JG: Surgical management of hydrocephalus secondary to intraventricular hemorrhage in the preterm infant. J Neurosurg Pediatr 17:278–284, 2016
5. Taylor AG, Peter JC: Advantages of delayed VP shunting in post-haemorrhagic hydrocephalus seen in low-birth-weight infants. Childs Nerv Syst 17:328–333, 2001
6. Vinchon M, Lapeyre F, Duquennoy C, Dhellemmes P: Early treatment of posthemorrhagic hydrocephalus in low-birth-weight infants with valveless ventriculoperitoneal shunts. Pediatr Neurosurg 35:299–304, 2001
7. Anwar M, Kadam S, Hiatt IM, Hegyi T: Serial lumbar punctures in prevention of post-hemorrhagic hydrocephalus in preterm infants. J Pediatr 107:446–450, 1985
8. Behjati S, Emami-Naeini P, Nejat F, El Khashab M: Incidence of hydrocephalus and the need to ventriculoperitoneal shunting in premature infants with intraventricular hemorrhage: risk factors and outcome. Childs Nerv Syst 27:985–989, 2011
9. Mazzola CA, Choudhri AF, Auguste KI, Limbrick DD Jr, Rogido M, Mitchell L, et al. : Pediatric hydrocephalus: systematic literature review and evidence-based guidelines. Part 2: Management of posthemorrhagic hydrocephalus in premature infants. J Neurosurg Pediatr 14 (Suppl 1):8–23, 2014
10. Lam HP, Heilman CB: Ventricular access device versus ventriculosubgaleal shunt in post hemorrhagic hydrocephalus associated with prematurity. J Matern Fetal Neonatal Med 22:1097–1101, 2009
11. Wang JY, Amin AG, Jallo GI, Ahn ES: Ventricular reservoir versus ventriculosubgaleal shunt for posthemorrhagic hydrocephalus in preterm infants: infection risks and ventriculoperitoneal shunt rate. J Neurosurg Pediatr 14:447–454, 2014
12. Wellons JC III, Shannon CN, Holubkov R, Riva-Cambrin J, Kulkarni AV, Limbrick DD Jr, et al. : Shunting outcomes in posthemorrhagic hydrocephalus: results of a Hydrocephalus Clinical Research Network prospective cohort study. J Neurosurg Pediatr 20:19–29, 2017
13. Adams-Chapman I, Hansen NI, Stoll BJ, Higgins R: Neurodevelopmental outcome of extremely low birth weight infants with posthemorrhagic hydrocephalus requiring shunt insertion. Pediatrics 121:e1167–e1177, 2008
14. Whitelaw A: Repeated lumbar or ventricular punctures in newborns with intraventricular hemorrhage. Cochrane Database Syst Rev (1):CD000216, 2001
15. Whitelaw A, Lee-Kelland R: Repeated lumbar or ventricular punctures in newborns with intraventricular haemorrhage. Cochrane Database Syst Rev 4:CD000216, 2017
16. AI can’t replace doctors. Buit can make them better. MIT Technology Review n.d. https://www.technologyreview.com/s/612277/ai-cant-replace-doctors-but-it-can-make-them-better/ (access the 29th of June 2021)
17. Artificial Intelligence Global Adoption Trends and Strategies n.d. <https://www.idc.com/getdoc.jsp?-containerId=US45120919> (access the 28th of June 2021)
18. Beunza J, Puertas E, Condés E. Manual práctico de inteligencia artificial en entornos sanitarios. Barcelona: Elsevier; 2020.
19. Palpan Flores A, Saceda Gutiérrez J, Brin Reyes JR, Sierra Tamayo J, Carceller Benito F. Risk factors associated with conversion of an Ommaya reservoir to a permanent cerebrospinal fluid shunt in preterm posthemorrhagic hydrocephalus. J Neurosurg Pediatr. 2020 Jan 17:1-8. doi: 10.3171/2019.11.PEDS19320. Epub ahead of print. PMID: 31952037.
20. Arbabshirani MR, Fornwalt Bk, Mongelluzo GJ, Suever JD, Geise BD, Patel AA, et al. Advanced machine learning in action: identification of intracranial hemorrhage on computed tomography scans of the head with clinical workflow integration. NPJ Digit Med 2018;1:9. Doi:10.1038/s41746-017-0015z.
21. Aprendizaje automático. Coursera n.d. https://es.corusera.org/learn/machine -learning (access 29 june 2021)
22. Yang XT, Feng DF, Zhao L, Sun ZL, Zhao G. Application of the Ommaya Reservoir in Managing Ventricular Hemorrhage. World Neurosurg. 2016 May;89:93-100. doi: 10.1016/j.wneu.2015.12.040. Epub 2015 Dec 25. PMID: 26724628.
23. Agajany N, Gigi M, Ross J, Roth J, Eshel R, Constantini S, Bassan H. The impact of neonatal posthemorrhagic hydrocephalus of prematurity on family function at preschool age. Early Hum Dev. 2019 Oct;137:104827. doi: 10.1016/j.earlhumdev.2019.104827. Epub 2019 Jul 30. PMID: 31374456.
24. Pendleton C, Cristofalo EA, Biondo GN, Jallo GI, Quiñones-Hinojosa A, Ahn ES. Posthemorrhagic hydrocephalus in preterm neonates: socioeconomic characteristics in a single-institution experience. Pediatr Neurosurg. 2012;48(2):80-5. doi: 10.1159/000342539. Epub 2012 Sep 28. PMID: 23038417.

# FIGURE LEGEND

Figure 1: Treatment Algorithm. Clinical worsening was defined as a bulging fontanelle, bradycardia, or arterial hypertension. Poor sonographic response was defined as a ventricular index greater than the preoperative value.

Table 1: kNN: k-Nearest Neighbors, D1: Fold1 ,D2: Fold 2,D3: Fold 3, D4: Fold 4

Table 2: No need for legend

Table 3: No need for legend

Table 4: No need for legend

Table 5: No need for legend

Table 6: No need for legend