Reappraisal of the benign external hydrocephalus syndrome. A prospective study.

Running Title: Reappraisal of the benign external hydrocephalus syndrome**.**

Authors' full name:

Federica Maruccia, MSc[1,2], Laura Gomáriz, MSc[1], Katiuska Rosas, MD[3], Turgut Durduran, PhD[2,4], Fernando Paredes-Carmona, M.D.[5], Juan Sahuquillo*, MD, PhD [1,3,6] and Maria A. Poca*, MD, PhD [1,3,6]

* These authors contributed equally

1. Neurotrauma and Neurosurgery Research Unit, Vall d'Hebron Institut de Recerca (VHIR), Vall d'Hebron Hospital Universitari, Vall d'Hebron Barcelona Hospital Campus, Passeig Vall d'Hebron 119-129, 08035 Barcelona, Spain; 2. ICFO - Institut de Ciencies Fotoniques, The Barcelona Institute of Science and Technology, Barcelona, Spain; 3. Department of Neurosurgery, Vall d'Hebron Hospital Universitari, Vall d'Hebron Barcelona Hospital Campus, Passeig Vall d'Hebron 119-129, 08035 Barcelona, Spain; 4. Institució Catalana de Recerca i Estudis Avançats (ICREA), Barcelona, Spain; 5. Department of Pediatrics, Hospital Universitari Arnau de Vilanova de Lleida, Spain; 6. Universitat Autònoma de Barcelona, 08193 Bellaterra, Spain.

Neuroimaging Studies

The size of the extraventricular CSF compartments was measured along the frontal convexities at the coronal slices in transfontanellar ultrasound or MRI to calculate: the craniocortical width, the sinocortical width (the distance from the lateral wall of the superior sagittal sinus to the surface of the cerebral cortex), and the anterior part of the interhemispheric fissure (**Fig. 1**). In a review of the literature, Zahl et al found that the upper limits of a normal craniocortical width range from 4 to 10 mm in infants <1 year of age, and from 3.3 to 5 mm in neonates. [11] The defined upper limit of the normal interhemispheric fissure width ranges from 6 to 8.5 mm, whereas a comparable spectrum for the sinocortical width is from 2 to 10 mm.[11] Prassopoulos and Cavouras showed that the extent of the subarachnoid spaces is not gender dependent, either in children younger or older than 3 years. [8] Consequently, to establish the diagnosis of BEH we required that at least one of three measurements of the frontal subarachnoid space was greater than 10 mm.

Ventricular volume —in transfontanellar ultrasound, CT scan or MRI— was estimated in all children using the Evans' index (EI) [5, 9] calculated as the maximum width between the frontal horns of the lateral ventricles and the maximum transverse inner diameter of the skull at the same axial slice in the CT scan and MRI and in the same coronal slice in transfontanellar ultrasound (**Fig. 1**). Only patients with an Evans Index ≤ 0.35 were included in this study.

Psychomotor development assessment

The Bayley-III scales. Bayley's third edition was published in 2006 and was designed to assess the developmental functioning of infants and toddlers from 16 days to 42 months 15 days of age, and the raw data was compared to a standardized norm of US children. [1–3] The Bayley-III scale has five distinct scales: Cognitive, Language, Motor, Social-Emotional, and Adaptive Behavior. [1–3] In our study, we used the Cognitive, Language, and Motor scales. In Bayley-III, the examinee's chronological age (adjusted for prematurity if required) defines a starting point designated by a letter A through Q. This letter is used to determine the starting item for the Cognitive, Language, and Motor scales. [1–3] The Language scale consists of the receptive and expressive language subscales, and the motor scale has a fine and gross motor subscale. In the three selected scales, Bayley-III yields three composite scores reflecting infants' cognitive, language, and motor development. The composite language score combines expressive and receptive language together, and the composite motor score includes both gross and fine motor scales that are converted to a composite standardized score with a mean of 100 and an SD of 15. [1, 2] To improve the sensitivity of Bayley-III, we also used the expressive and receptive language and fine and gross motor development subscales independently as they might identify differences in the language and the motor development, which might be hidden in composite scores. In the subscales, the standardized scores have a normative mean of 10 and an SD of 3 [2].

Standardization sample: In Bayley-IIII, the standardization sample for the Cognitive, Language, and Motor scales included 1700 children aged 1 month to 42 months, divided into 17 separate age groups, with 100 individuals in each group. [1] Only children who were born at 36 to 42 weeks gestation and who were typically developing were included in the standardization sample.

Test situation. Children were always tested in the presence of at least one caregiver. The test was performed in properly-ventilated, well-illuminated clinical rooms. An appropriate setting was created before the entrance of the patients, separate from any distracting material. Toys were used to make the child and the caregiver feeling comfortable before starting the assessment. Cognitive, receptive, and expressive language and fine motor items of the Bayley III were tested with the child sitting alone or on the caregiver's lap in front of a table. Gross motor items of the Bayley-III were assessed on the floor. The situation was adjusted according to the individual child's needs. The examiner started with a preliminary interview and the assessment explanation to the caregiver, and then, the test was performed, giving periodical breaks to the child.

Statistical analysis

Multiple logistic regression model (MLR). MLR was used to explore the relationships between predictors and the outcome variable from the effects of covariates. The outcome variable was described as any detected delay in the simple or composite scales of the Bayley-III with the criteria already described. The absence of delay in any scale was coded as 0 and as 1 when the child presented at least a delay in any single or composite scale. We applied the general linear model (glm) function in R using the binomial family. Preselected input variables were introduced in the model according to the 'purposeful method' suggested by Hosmer et al. [7] Independent predictors for any delay were age, weight at birth, prematurity, and most of the demographic variables recorded in this cohort. All variables with $p < 0.25$ in

the univariate analysis were then entered in an MLR analysis. [7] Variables that were not statistically significant at p <0.05 were eliminated, and a new model was generated without them. In the third step, variables excluded in the univariate analysis were added individually to the final model to test statistical significance. In the final model, the original coefficients, their statistical significance, the 95% confidence intervals (CI), and the odds ratio (OR) were reported. A 2-tailed p-value <0.05 was considered statistically significant for the MLR. Nagelkerke pseudo-R-squared values were used as a goodness-of-fit measurement for the model.

The variables that were clinically relevant and/or statistically significant ($p < 0.25$) in univariate analysis were included in the first model. Thirteen covariates were included in the first MLR model: age in months, gestational age, sex, Apgar score, birth weight, prematurity, macrocephaly frequency, Evan's index, ethnicity, father and mother education in years, and father and mother's HC. In a second iteration, a new reduced model fit was tested using the covariates that were statistically significant ($p \leq 0.05$) in the initial model. Deleting, refitting, and verifying covariates were continued until all relevant variables were included in the model, with those excluded being clinically and/or statistically irrelevant. The final model included only non-complicated prematurity as the only independent factor for the detection of any developmental delay in infants or children with BEH.

Limitations and future directions

A limitation of our study is the use of the Bayley-III scales since it was standardized on an American pediatric population. Differences in cross-cultural performance in Bayley-II and III have been described when comparing infants from different countries. [6] Chinta et al. have shown that Bayley-III underestimates developmental delay in healthy 3-year-old Australian children. [4] Ideally, differences in cross-cultural performance require a validation of the scales with a country-specific cohort of healthy children. Although the validity of the scales has been verified in a cohort of Spanish children with autism, [10] ideally Bayley-III should be re-standardized on a local population of healthy children. The main risk of this lack of local controls is that children with delays are not detected, and therefore they are disregarded as candidates for early intervention and follow-up.

Future studies on BEH need to use similar inclusion criteria, standards for clinical thresholds with homogeneous and standardized tools for evaluating developmental delay and for long term follow-ups. Many studies have shown that prematurity and body weight at birth are risk factors for neurodevelopmental deficits compared to full-term children. Our findings that uncomplicated prematurity is a risk factor for delay in patients with BEH should be replicated in a larger cohort of patients. To confirm the presence of a long-term delay in premature and full-term BEH children, we need a larger number of patients with information about long-term follow-up.

BIBLIOGRAPHY

1. Albers CA, Grieve AJ. (2007) Test Review: Bayley, N. (2006). Bayley Scales of Infant and Toddler Development– Third Edition. J Psychoeduc Assess 25(2):180–190. doi: 10.1177/0734282906297199

2. Bayley N. (2006) Bayley Scales of Infant and Toddler Development-Third Edition. Harcourt Assessment, SAn Antonio, TX, USA

3. Bayley N. (2015) Spanish Adaptation of the Bayley Scales of Infant and Toddler Development. Pearson Educación SA, Madrid

4. Chinta S, Walker K, Halliday R, Loughran-Fowlds A, Badawi N (2014) A comparison of the performance of healthy Australian 3-year-olds with the standardised norms of the Bayley Scales of Infant and Toddler Development (version-III). Archives of Disease in Childhood 99:621–624. doi: 10.1136/archdischild-2013-304834

5. Evans WA (1942) AN ENCEPHALOGRAPHIC RATIO FOR ESTIMATING VENTRICULAR ENLARGEMENT AND CEREBRAL ATROPHY. Arch NeurPsych 47:931. doi: 10.1001/archneurpsyc.1942.02290060069004

6. Fischer VJ, Morris J, Martines J. (2014) Developmental screening tools: feasibility of use at primary healthcare level in low- and middle-income settings. J Health Popul Nutr 32(2):314–26

7. Hosmer DW, Lemeshow S, Sturdivant RX. (2013) Applied logistic regression. 3 ed. Wiley series in probability and statistics. Wiley;

8. Prassopoulos P, Cavouras D (1994) CT evaluation of normal CSF spaces in children: relationship to age, gender and cranial size. European Journal of Radiology 18:22–25. doi: 10.1016/0720-048X(94)90359-X

9. Sari E, Sari S, Akg[illegible]n V, [illegible]zcan E, [illegible]nce S, Babacan O, Saldir M, A[illegible]ikel C, Basbozkurt G, Yesilkaya S, Kilic C, Kara K, Vurucu S, Kocaoglu M, Yesilkaya E (2015) Measures of Ventricles and Evans' Index: From Neonate to Adolescent. Pediatr Neurosurg 50:12–17. doi: 10.1159/000370033

10. Torras-Mañá M, Gómez-Morales A, González-Gimeno I, Fornieles-Deu A, Brun-Gasca C (2016) Assessment of cognition and language in the early diagnosis of autism spectrum disorder: usefulness of the Bayley Scales of infant and toddler development, third edition. Journal of Intellectual Disability Research 60:502–511. doi: 10.1111/jir.12291

11. Zahl SM, Egge A, Helseth E, Wester K (2011) Benign external hydrocephalus: a review, with emphasis on management. Neurosurg Rev 34:417–432. doi: 10.1007/s10143-011-0327-4