A Framework For Memory Performance Prediction From Brain Volume In Preterm-Born Adolescents

Hassna Irzan[1,2], **Helen O'Reilly**[3], **Sebastien Ourselin**[2,1], **Neil Marlow**[3], **Andrew Melbourne**[2,1]

[1]Dept. Medical Physics and Biomedical Engineering, University College London

[2]Biomedical Engineering and Imaging Sciences, Kings College London

[3]Institute for Women's Health, University College London

Abstract

With advances in medical care, higher numbers of extremely preterm-born babies are now surviving, however the rate of neurodevelopmental and neurological complications has not improved at the same rate and the relative rate of disabilities and health problems is increasing, with associated high costs for health care systems and education. Understanding brain development after early birth is of great importance to be able to make informed decisions. Many studies have associated different areas of the preterm brain with poor cognitive performance, however it is less clear whether these associations persist into adult life. In this study, we investigate how well cortical volumes describe memory performance in 133 19 year-old adolescents, 61% of whom were born extremely preterm. We employ LASSO to identify brain regions that better explain memory performance. The brain regions identified by LASSO explained 27% and 32% of the variance in the visual working memory scores and the visual short term memory respectively. Furthermore, the correlation between the predicted scores and validation scores is statistically significant and it is 58% for the visual working memory task and 56% for the visual short term memory task.

1 Introduction

In the last decades, studies have shown a general decreased mortality and incidence of disability for low gestational age at birth [1], however cortical grey matter, white matter changes, even without evidence of focal lesion, and cognitive deficits have been found to be one of the most occurring abnormalities [1]. Recent studies on a cohort of adolescents who were born extremely preterm concluded that the general cognitive functioning does not improve or deteriorate from infancy through to 19 years of age [2]. The extent to which alterations of regional brain volumes adversely affect the neuropsychological performance of preterm subjects is still unclear [3]. Studies on preterm subjects have demonstrated linear relationship between the volume of some brain regions and cognitive performance [4]. In our cohort, preterm subjects present abnormal brain volumes and memory performance deficit. We aim to investigate the relation between the visual memory performance and cortical brain volumes. Specifically we analyse visual short term memory (V-STM) and working memory (V-WM) outcome and brain volume. Firstly, we explore the one- to-one relationship between memory scores and some ROI such as hippocampus and thalamus. Secondly, we investigate if a linear combination of more than one regional volume can predict memory

performance. This is achieved by performing feature selection on the cortical brain volumes using LASSO. The aim of the latter is to investigate the differential pattern of brain regions that is linked to the gap in memory performance between term born and preterm born subjects.

2 Materials And Methods

2.1 Participants

Eighty-one adolescents (48 female, 33 male) born extremely preterm (EP), along with 52 term-born socioeconomically matched peers (TC) (33 female, 19 male) were considered in this study. The EP subjects were born, in UK and Ireland, between March 1 and December 31, 1995, from 20 to 25 weeks of gestation and 6 days. The mean gestational age at birth for the EP subjects is 24.9±0.8 and the birth weight is 728.4 ± 126.3 g.

2.2 MRIdata

The subjects were scanned with a 3T Philips Achieva; 3D T1-weighted volume were acquired at TR = = 6.93 ms, TE = 3.14 ms and 1 mm isotropic resolution. The T1-weighted images were bias field corrected by using N4ITK algorithm. Geodesic Information Flows (GIF) [5] framework was applied to extract brain parcellations. This framework produces structural parcellations (144 brain regions) by propagating voxel-wise annotations between several atlases based on a local distance metric. We used the structural parcellations of GIF to derive the regional volumes from each subjects brain. The volumes are normalised with respect to the total volume of the brain. Note that for each brain region we consider left and right volumes separately.

2.3 Memory assessment

In this work we investigate V-STM and VS-WM selected from the Automated Working Memory Assessment (AWMA, Alloway 2007) Dot Matrix and Spatial Recall. The temporary storage of information in the brain is explained by two models: working memory and short term memory [6]. The first deals with complex span tasks, which involves maintaining a collection of information and performing some cognitive task on it; the latter is measured by a simple span tasks, which require to maintain some information over a small amount of time [6]. However, these models are distinct from structural and connectivity-based models of memory, and thus there might be an overlap between the brain regions processing them [6]. A normality test suggested that the distributions of the data are unlikely Gaussians, hence the statistical analysis is conducted using Mann-Whitney U-test non-parametric test.

2.4 Regression analysis

To examine the relationship between the volume of the cortical regions and the memory performance we apply two strategies. In the first experiment, we compute the correlation between the memory scores (V-STM and VS-WM) and the brain volume of the thalamus (Thal) and the hippocampus (HPC). In the previous studies on this cohort, thalamus and hippocampus showed reduced volumes and thalamus abnormal shape [7]. We investigate if the abnormalities seen in such regions are related to V-STM and S-WM deficits. The

statistical analysis is carried out, correcting for gender and prematurity, by using Mann-Whitney U-test with p values by two-sided χ^2. We correct for multiple comparisons using Bonferroni correction. In the second experiment, we aim to study if a linear combination of more than one regional volume can predict the memory performance. We use a data-driven approach in which we rely on Least Absolute Shrinkage and Selection Operator (LASSO) to identify the most relevant cortical volumes that can predict the memory scores. LASSO regression is a linear regression with L1 regularisation method and it optimises the following objective function:

$$\hat{\beta} = \underset{\beta}{argmin}\left(\frac{1}{2n} \parallel Y - \beta X \parallel_2^2 + \lambda \parallel \beta \parallel_1 \right) \tag{1}$$

where Y represents the memory scores across the subjects, X is the whole-brain cortical volumes and n is the number of samples. The first addend is the normal least squares, and the second is the penalty term equivalent to the absolute value of the magnitude of coefficients β and the parameter λ is the shrinkage parameter and it is an unknown.

First, we standardise the data, then we randomly split the data into a 70% training set, 15% testing set and 15% validation set. We then use 10-folding cross validation to estimate coefficients $\hat{\beta}$ and the parameter λ. We estimate A without any *prior* knowledge, other than $\lambda \quad 0$. In our identification strategy the best λ is the one that minimises, on average, the Mean Squared Error (MSE) metric for the testing set on each fold, varying λ. Specifically, we use a grid of shrinkage parameters λ for fitting, then we perform cross validation on each testing fold and compute MSE as:

$$MSE = \frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \tag{2}$$

where y_i is the real observation and $\hat{Y}_i$ is its estimate. We then evaluate the model performance by comparing the predicted memory scores $\hat{Y}$ and the measured memory scores Y_{actual} of the validation by correlation coefficient and Explained Variance Score (EVS). This is defined as:

$$EVS = 1 - \frac{Var(Y_{actua} - \hat{Y})}{Var(Y_{actual})} \tag{3}$$

One of the advantages of LASSO is that it can be used even when the number of variables is higher than the number of observations. This because the regularisation term sets most coefficients to zero, promoting sparsity and avoiding overfitting. This suits our case study because we have higher number of cortical regions than subjects.

3   Results

Memory impairment is observed in the extremely-preterm born group. Table 1 shows that the preterm subjects have scored less than full-term subjects independently from gender, and that the gap is statistically significant. In addition, it reports thalamic and hippocampal (only female) volume loss. With the exception of right hippocampus ($\rho = -0.32$, $p = 3.8e^{-3}$) when

controlling for prematurity, V-STM scores do not correlate with the individual ROI volumes. When comparing EP and TC females, V-WM scores show significant correlation with the volume of the right Thal ($p = 0.31$, $p = 5.0e^{-3}$), left Thal ($p = 0.28$, $p = 1.2e^{-2}$) and left HPC ($p = 0.24$, $p = 3e^{-2}$). The results are not statistically significant after Bonferroni correction. Table 1 displays volumes of ROI and visual memory data. For the LASSO regression experiment, we found that the penalisation term λ for V-STM is 1.36 and for V-WM is 1.44. For each memory task the coloured brain areas in figure 1 have been selected by LASSO, while the black regions were not selected. Each brain area has a unique colour label, this helps highlighting the overlapping regions between memory tests. Table 2 contains the LASSO performance for both memory tasks. The correlation between the predicted scores and validation scores is statistically significant and it is 58% for the V-WM task and 56% for the V-STM task. The brain volume of the selected regions explains 27% and 32% of the variation in the V-WM scores and V-STM scores respectively. Although, the covariance is higher when comparing predicted and real V-WM score (57.92 vs 43.01), the variance of their difference is higher for the V-WM sample (229.33 vs 98.01). In figure 2 the predicted memory scores are plotted along the y-axis and the real memory scores along the x-axis for the both the memory tasks. The blue line represents the ideal case in which the LASSO regression is able to perfectly predict the measured scores.

4   Discussion

We investigated patterns of brain volumes and their associated memory performance in individuals at adolescence after extremely preterm birth. To link memory scores and cortical volumes, first we computed a one-to-one correlation, then we performed feature selection with LASSO regression.

Thalamus is known to be an important region in this cohort [7]. Despite its role in relaying sensory signals including visual information it doesn't appear to be correlated to visual memory tasks.

In the LASSO analysis, similar numbers of regions were selected as predictive for each memory test, with several overlapping regions. A complex of regions from the dorsolateral frontal cortex, primary motor cortex and posterior parietal cortex were selected for the working memory task (figure 1). These regions have previously been found to be related to working memory tasks by other imaging studies [6]. For the V-STM tasks, the selected regions are distributed along superior frontal, inferior temporal and parietal lobe (figure 1). There are several regions that have been selected from both tasks such as right fusiform gyrus, left anterior orbital gyrus, left central operculum, right supplementary motor cortex, left planum polare, right precentral gyrus, left middle occipital gyrus, right precentral gyrus medial segment and left amygdala. It has been anticipated that the same brain regions, mainly in the frontal lobe, mediate the cognitive load of both working memory and short term working memory [6]. In addition, the distribution of the selected regions for both memory tests recalls the dorsal and ventral visual processing streams as presented by some imaging studies [8]. Furthermore, the right hemispheric asymmetry in our results has been observed by previous studies [9].

Overall, the visual memory tasks used involved recall of spatial relations which would explain the involvement of the parietal lobe given its role in the dorsal visual pathway. It is possible that the areas identified by LASSO merely mirror the processing of visual information in the dorsal visual pathway, and the frontal involvement indicates the executive function aspects of short term and working memory. LASSO regression is able to describe some of the data variation, however the volume alterations were not sufficient to explain the overall poor visual memory outcome. Certainly the neuroanatomical substrates associated with memory processing are more complex, and most likely, brain volumes difference *per se* does not fully explain functional change. It is also possible that such complex cognitive function might depend on a global neural workspace rather than segregated brain areas. In the future, our work will explore the link between cognitive performance and other brain features, and make use of a wider range of MR contrasts to explore both structural and functional connectivity.

Acknowledgements

References

[1]. Ball G, Boardman JP, Rueckert D, Aljabar P, Arichi T, Merchant N, Gousias IS, David Edwards A, Counsell SJ. The effect of preterm birth on thalamic and cortical development. Cerebral Cortex. 2012; 22(5):1016–1024. [PubMed: 21772018]

[2]. Linsell L, Johnson S, Wolke D, O'Reilly H, Morris JK, Kurinczuk JJ, Marlow N. Cognitive trajectories from infancy to early adulthood following birth before 26 weeks of gestation: a prospective, populationbased cohort study. Archives of Disease in Childhood. 2018; 103(4):363–370. [PubMed: 29146572]

[3]. Nosarti C, AlAsady MHS, Frangou S, Stewart AL, Rifkin L, Murray RM. Adolescents who were born very preterm have decreased brain volumes. Brain. 2002; 125(7):1616–1623. [PubMed: 12077010]

[4]. Keunen K, Išgum I, van Kooij BJM, Anbeek P, van Haastert IC, Koopman-Esseboom C, Fieret-van S Petronella C, Nievelstein RAJ, Viergever MA, de Vries LS, Groenendaal F, et al. Brain volumes at term-equivalent age in preterm infants: Imaging biomarkers for neurodevelopmental outcome through early school age. The Journal of Pediatrics. 2016; 172:88–95. [PubMed: 26774198]

[5]. Cardoso MJ, Modat M, Wolz R, Melbourne A, Cash D, Rueckert D, Ourselin S. Geodesic information flows: Spatially-variant graphs and their application to segmentation and fusion. IEEE Transactions on Medical Imaging. 2015; 34(9):1976–1988. [PubMed: 25879909]

[6]. Aben B, Stapert S, Blokland A. About the distinction between working memory and short-term memory. Frontiers in Psychology. 2012; 3:301. [PubMed: 22936922]

[7]. Orasanu, E; Melbourne, A; Eaton-Rosen, Z; Atkinson, D; Lawan, J; Beckmann, J; Marlow, N; Ourselin, S. Local shape analysis of the thalamus in extremely preterm born young adults. 2016.

[8]. Ungerleider LG, Courtney SM, Haxby JV. A neural system for human visual working memory. Proceedings of the National Academy of Sciences. 1998; 95(3):883.

[9]. Sheremata SL, Somers DC, Shomstein S. Visual short-term memory activity in parietal lobe reflects cognitive processes beyond attentional selection. The Journal of Neuroscience. 2018; 38(6):1511. [PubMed: 29311140]
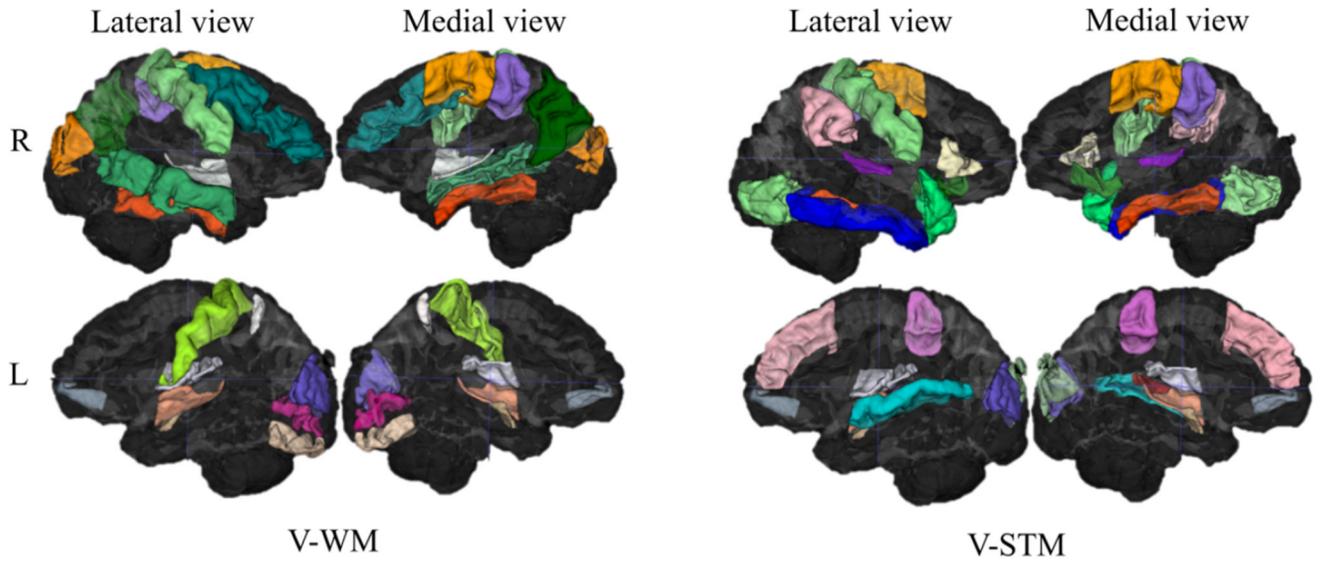
Fig. 1. Regions selected by LASSO for V-WM (right) and for V-STM (left) are coloured, the black areas were not selected.
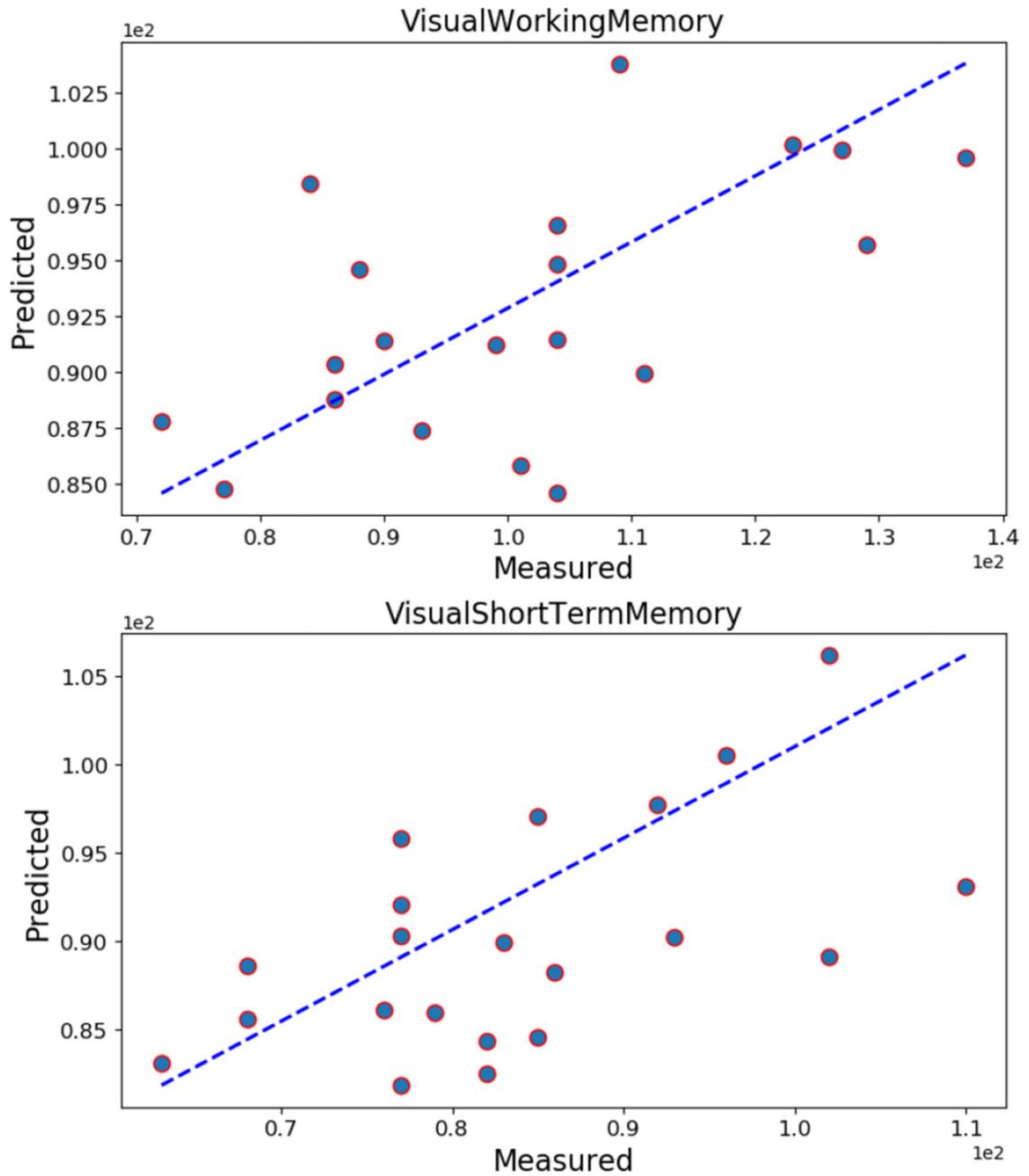
Fig. 2. The actual and the predicted memory scores are plotted along the *x*-axis and *y*-axis respectively. The blue line represents the perfect matching between predicted and real memory scores.

Table 1

Volumes and psychological data are split by prematurity and gender. The sample size is indicated by *n*. The statistical analysis is performed using Manne-Whitney U-test with *p* values by two-sided X^2

	All subjects			Female			Male		
Psychological test	**EP n = 81**	**TC n = 52**	***p***	**EP n = 48**	**TC n = 33**	***p***	**EP n = 33**	**TC n = 19**	***p***
V-STM, $\mu\pm\sigma$	81.0±12.6	97.1±15.5	$7.6\mathrm{e}^{-7}$	80.6±11.7	95.0±15.9	$5.1\mathrm{e}^{-5}$	81.6±13.9	100.5±15.3	$5.7\mathrm{e}^{-5}$
V-WM, $\mu\pm\sigma$	86.3±11.3	103.3±15.7	$7.2\mathrm{e}^{-8}$	85.9±11.1	102.2±16.5	$2.2\mathrm{e}^{-6}$	86.9±11.7	105.1±14.6	$4.6\mathrm{e}^{-5}$
ROI Volumes									
L HPC, $\mu\pm\sigma$	$2.8\mathrm{e}^{-3}$ ±$2.8\mathrm{e}^{-4}$	$2.8\mathrm{e}^{-3}$±$2.5\mathrm{e}^{-4}$	0.90	$2.7\mathrm{e}^{-3}$±$2.7\mathrm{e}^{-4}$	$2.9\mathrm{e}^{-3}$±$2.3\mathrm{e}^{-4}$	$1.5\mathrm{e}^{-3}$	$2.7\mathrm{e}^{-3}$±$2.9\mathrm{e}^{-4}$	$2.7\mathrm{e}^{-3}$ ±$2.2\mathrm{e}^{-4}$	0.3
R HPC, $\mu\pm\sigma$	$3.0\mathrm{e}^{-3}$ ±$3.2\mathrm{e}^{-4}$	$3.0\mathrm{e}^{-3}$±$2.5\mathrm{e}^{-4}$	0.99	$3.0\mathrm{e}^{-3}$±$3.1\mathrm{e}^{-4}$	$3.1\mathrm{e}^{-3}$±$2.5\mathrm{e}^{-4}$	$3.1\mathrm{e}^{-2}$	$2.9\mathrm{e}^{-3}$±$3.6\mathrm{e}^{-4}$	$2.9\mathrm{e}^{-3}$ ±$2.3\mathrm{e}^{-4}$	0.5
L Thal, $\mu\pm\sigma$	$4.2\mathrm{e}^{-3}$ ±$3.7\mathrm{e}^{-4}$	$4.4\mathrm{e}^{-3}$±$3.3\mathrm{e}^{-4}$	0.01	$4.1\mathrm{e}^{-3}$±$3.6\mathrm{e}^{-4}$	$4.4\mathrm{e}^{-3}$±$3.6\mathrm{e}^{-4}$	$2.1\mathrm{e}^{-4}$	$4.1\mathrm{e}^{-3}$±$3.1\mathrm{e}^{-4}$	$4.4\mathrm{e}^{-3}$ ±$2.7\mathrm{e}^{-4}$	$1.0\mathrm{e}^{-3}$
R Thal, μ±σ	$4.1\mathrm{e}^{-3}$ ±$3.9\mathrm{e}^{-4}$	$4.3\mathrm{e}^{-3}$±$2.4\mathrm{e}^{-4}$	0.02	$4.0\mathrm{e}^{-3}$±$3.3\mathrm{e}^{-4}$	$4.3\mathrm{e}^{-3}$±$2.4\mathrm{e}^{-4}$	$1.7\mathrm{e}^{-4}$	$3.9\mathrm{e}^{-3}$±$4.3\mathrm{e}^{-4}$	$4.2\mathrm{e}^{-3}$ ±$2.4\mathrm{e}^{-4}$	$1.8\mathrm{e}^{-3}$

Table 2

Linear correlation and EV of the model prediction and the validation scores

Memory test	ρ	p-value	EV
V-WM	0.58	$7.5\mathrm{e}^{-3}$	27%
V-STM	0.56	$8.0\mathrm{e}^{-3}$	32%