

Studying the effect of scanner change on measures of gray matter for Alzheimer's disease studies

Agnes Pérez-Millan^{1,2}, José Contador¹, Adrià Tort¹, Sergi Borrego-Écija¹, Beatriz Bosch¹, Mircea Balasa¹, Bàrbara Segura^{3,4}, Gemma Monté-Rubio³, Núria Bargalló¹, Albert Lladó^{1,4}, Raquel Sanchez-Valle^{1,4}, Roser Sala-Llloch^{2,5}

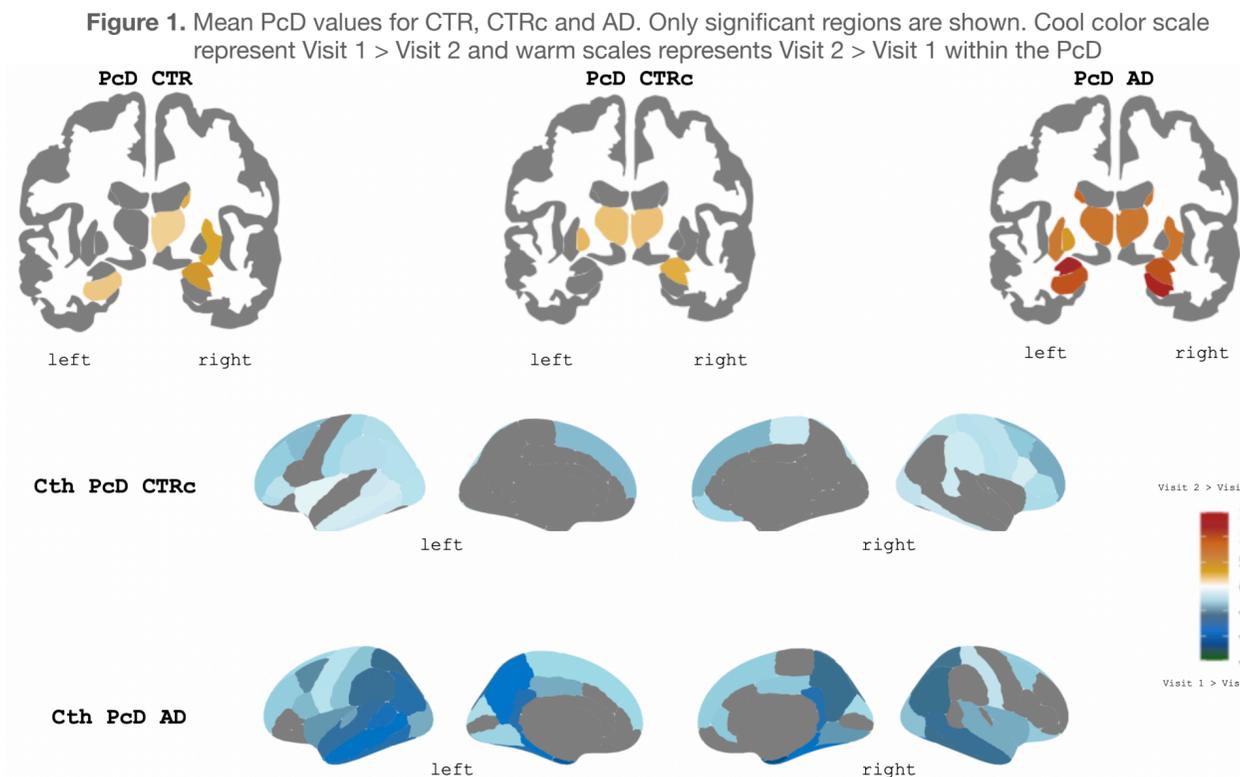
(1) Alzheimer's Disease and Other Cognitive Disorders Unit, Neurology Service, Hospital Clínic de Barcelona, Institut d'Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS), Fundació Clínic per a la Recerca Biomèdica, Universitat de Barcelona, Barcelona, 08036, Spain.
 (2) Institute of Neurosciences. Department of Biomedicine, Faculty of Medicine, University of Barcelona, Barcelona, 08036, Spain.
 (3) Institute of Neurosciences. Medical Psychology Unit, Department of Medicine, University of Barcelona, Barcelona, 08036, Spain.
 (4) Centro de Investigación Biomédica en Red de Enfermedades Neurodegenerativas. CIBERNED, Spain.
 (5) Biomedical Research Networking Center of Bioengineering, Biomaterials and Nanomedicine (CIBER-BBN), Barcelona, Spain.

INTRODUCTION: Longitudinal magnetic resonance imaging (MRI) studies allow the study of brain biomarkers for the study of neurological disorders such as Alzheimer's disease (AD). During the course of these studies, scanner changes may affect the reliability of MRI-derived metrics. Here we aim to study the effects of changing from a 3T-Siemens Trio Trim (namely scanner 1) to a 3T-Siemens Prisma (namely scanner 2) on gray matter (GM) measures obtained from T1-weighted acquisitions.

METHODS: We studied 3 MRI-longitudinal datasets with 2 time points separated 2 years, all having available T1-weighted scans, with equal or equivalent acquisition protocols: A group of 14 healthy controls scanned with scanner 1 at both timepoints (CTR), a group of 17 healthy controls scanned at baseline with scanner 1 and after 2 years with scanner 2 (CTRc), and a group of 12 AD patients, scanned with scanner 1 (Table 1). All AD patients showed pathological levels of core CSF biomarkers (A+ T+) with neurodegeneration (N+), while all CTR subjects were within the normal biomarker range (NIA-AA 2018). We used cross-sectional and longitudinal processing streams of FreeSurfer to obtain summary measures, including subcortical GM volume, cortical thickness (CTh), cortical surface area and cortical GM volume, using available atlas of cortical parcellations and subcortical segmentations. We studied the intra-class correlation coefficient (ICC) to assess the reliability of the metrics for CTR and CTRc groups. We calculated the percent difference (PcD) to obtain the relative difference of GM measures between timepoints, for all groups. With the PcD measure, we aimed to study the magnitude of the changes due to the scanner in comparison with the effect of AD atrophy. We further focused on hippocampal atrophy, as it is considered a biomarker for AD.

Table 1. Sample ages. CTR: Healthy Controls, CTRc: Healthy Controls with scanner change, AD: Alzheimer's Disease. There were no differences between groups.

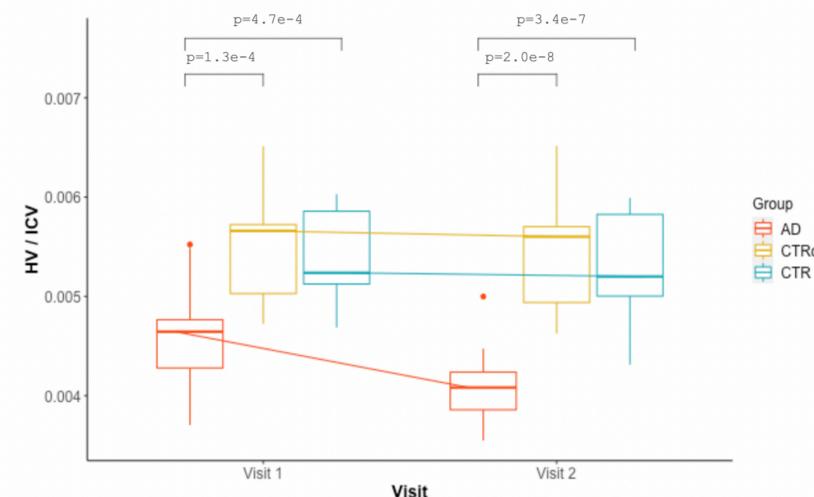
	CTR N=14	CTRc N=17	AD N=12
Age at visit 1 Median [interquartile range, IQR] years	58.4 [54.7, 61.4] years	59.7 [49.0, 60.9] years	60.4 [58.0, 65.8] years
Age at visit 2 Median [interquartile range, IQR] years	60.6 [56.9, 63.5] years	61.7 [50.0, 63.0] years	62.4 [59.9, 67.8] years



RESULTS: We found that the longitudinal FreeSurfer stream provided higher mean ICC values. Of all the measures studied, ROI-CTh were the most affected by scanner change, with a mean ICC across regions of 0.8 and 0.92 for the CTRc and CTR groups respectively (36 out of 68 regions were significantly different). We found differences in ICC between CTR and CTRc in 16/68 regions for the surface area measure and in 20/68 regions for the volume measure. However, mean ICC values were nearly 1 for both datasets for these metrics, indicating good reliability. As regards PcD, we consistently observed that longitudinal differences were stronger in AD than in CTR or CTRc groups (Figure 1). Mean PcD values for CTh (across parcellations) were: -0.20% for CTR, -1.53% for CTRc and -4.57% for AD. And mean PcD for subcortical GM volumes were (across atlas regions): -3.11% for CTR, -2.76% for CTRc and -10% for AD. Hippocampal atrophy was similar between CTR and CTRc with values nearly zero with no significant differences between them. In AD, hippocampal volume loss was stronger, showing significant differences with CTR and CTRc (Figure 2).

CONCLUSIONS: We provide an analysis of longitudinal GM measures derived from T1-weighted MRI scans on 3 different datasets, accounting for scanner changes and AD progression. In general, we found small test-retest changes and high reproducibility rates in healthy controls when analyzed with the FreeSurfer longitudinal stream, even when changing scanner. CTh was the most affected measure (although showing ICC nearly 1) by scanner changes, suggesting that further approaches could improve this measure. By using an AD group, we found high rates of changes in comparison with both the CTR and CTRc groups, suggesting that the change of scanner is not a handicap in order to identify pathological brain processes in T1-weighted data.

Figure 2. Boxplot of the HV normalized by the intracranial volume (ICV) for each visit. Pairwise comparisons show significantly different in some comparisons.



ACKNOWLEDGEMENTS: The authors thank patients for their participation in the study. This work was supported by Spanish Ministry of Science and Innovation-Instituto de Salud Carlos III and Fondo Europeo de Desarrollo Regional (FEDER), Unión Europea, "Una manera de hacer Europa" (PI19/00449 to Dr. Lladó) and Departament de Salut - Generalitat de Catalunya (PERIS 2016-2020 SLT008/18/00061 to Dr A. Lladó).