Harmonization of MRI and PET Data: Challenges, Solutions, and Relevance to an Era of New AD therapeutics

Introduction

Dana L. Tudorascu, PhD

University of Pittsburgh

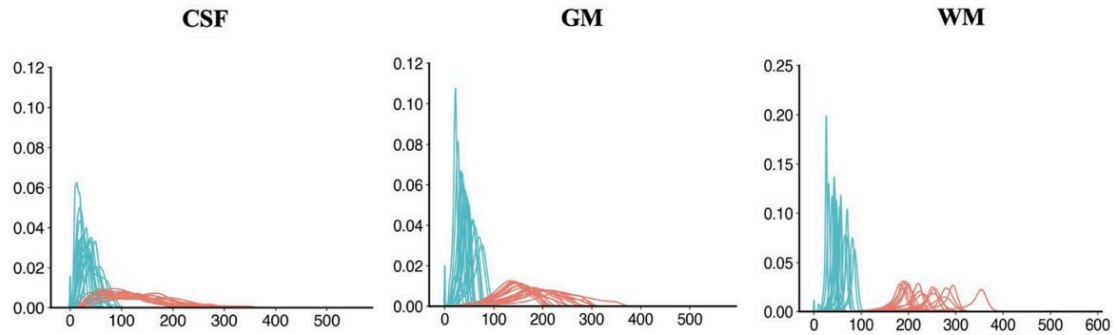
State of the field (data collection and technical issues)

Moved from single site neuroimaging studies to multisite studies, vast majority observational studies not randomized clinical trials.

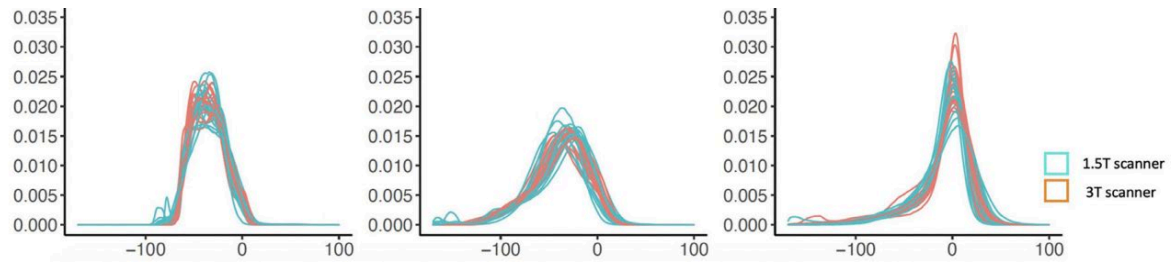
1. MRI
   - Different scanners for different sites for cross-sectional studies.
   - Different scanners/different sites and within subject change of scanners during longitudinal studies.
   - Different sampling population at each site.

2. PET
   - Different scanners for different sites for cross-sectional studies.
   - Different tracers/different sites.
   - Within subject longitudinal studies with different tracers at the same site.
   - Longitudinal studies with different within subject tracers or imaging with one tracer at earlier times and different scanner different tracer at later times (or vice versa).

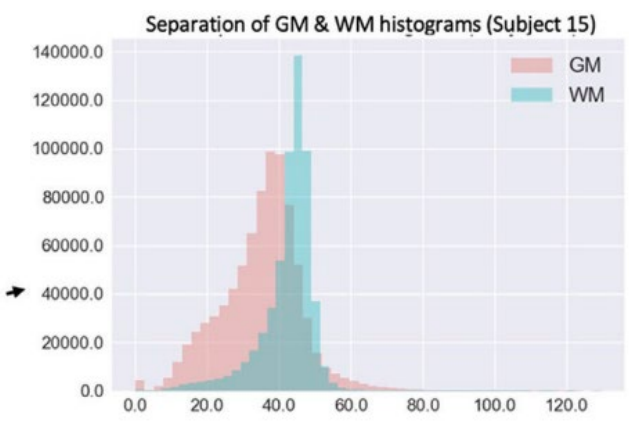
Why are different scanners a problem? (18 elderly scanned on 1.5T and then 3T, 3mo)

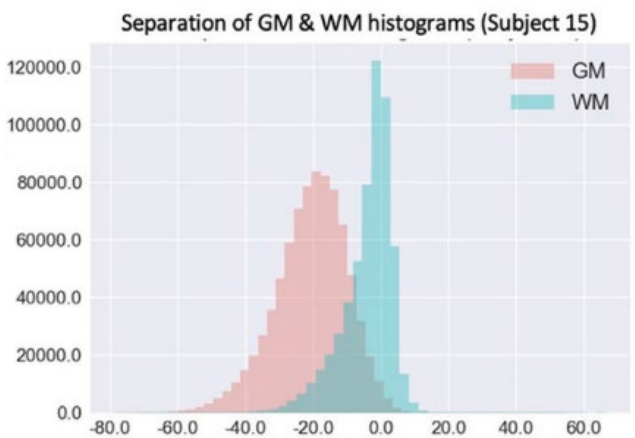


(a) RAW



(c) RAVEL-Normalized

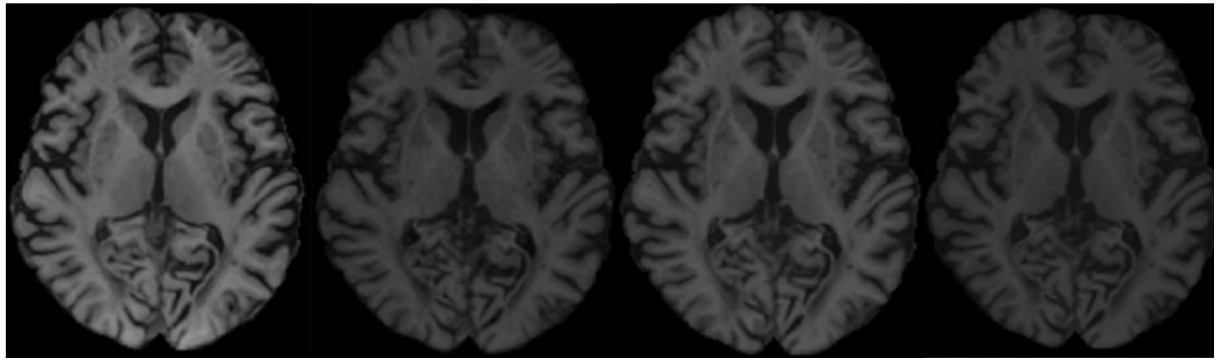




Mean (SD) of percent differences				
	RAW	**ComBat**	**RAVEL**	**RAVEL-ComBat**
ROIs	**Cortical Thickness**			
Left				
Entorhinal	7.43 (27.59)	6.48 (37.93)	2.64 (28.57)	6.12 (35.71)
Fusiform	11.11 (16.67)	4.22 (32.0)	4.74 (5.88)	4.22 (9.52)
Inferior Parietal	2.35 (13.33)	6.76 (38.1)	1.88 (0.0)	2.75 (16.67)
Inferior Temporal	10.78 (19.05)	2.42 (40.0)	4.6 (10.53)	3.24 (19.23)
Middle Temporal	3.17 (4.17)	0.77 (4.55)	0.78 (8.7)	1.56 (14.29)
Right				
Entorhinal	8.0 (12.5)	5.92 (29.17)	2.6 (5.26)	5.25 (36.36)
Fusiform	9.48 (0.0)	2.06 (20.0)	4.2 (5.88)	2.48 (0.0)
Inferior Parietal	0.47 (21.05)	3.18 (30.0)	1.4 (16.67)	0.46 (11.11)
Inferior Temporal	11.16 (25.0)	4.15 (5.88)	5.0 (16.67)	3.31 (6.67)
Middle Temporal	1.96 (10.53)	0.38 (31.25)	0.39 (5.26)	1.16 (33.33)
	Volumes			
Left				
Entorhinal	4.88 (6.25)	12.99 (11.76)	0.0 (9.38)	7.59 (2.94)
Inferior Temporal	9.81 (52.43)	5.07 (53.0)	3.74 (29.09)	3.04 (35.58)
Middle Temporal	1.82 (26.58)	3.78 (25.69)	1.79 (10.84)	2.61 (14.97)
Amygdala	9.29 (15.38)	6.29 (19.05)	4.9 (7.69)	2.74 (4.55)
Hippocampus	2.08 (4.35)	4.75 (4.26)	0.79 (0.0)	2.03 (2.13)
Right				
Entorhinal	5.62 (5.88)	7.74 (6.06)	0.0 (5.71)	5.1 (0.0)
Inferior Temporal	12.71 (31.71)	8.27 (10.95)	5.26 (15.5)	4.93 (3.62)
Middle Temporal	2.32 (38.52)	0.62 (24.46)	1.04 (19.58)	0.2 (13.38)
Amygdala	3.23 (0.0)	1.28 (5.26)	1.92 (4.76)	0.0 (0.0)
Hippocampus	2.26 (10.0)	2.44 (8.7)	0.76 (5.0)	0.99 (8.89)

MRI data from different scanners (same participant, same slice, 3T, within one week, N=20 total)

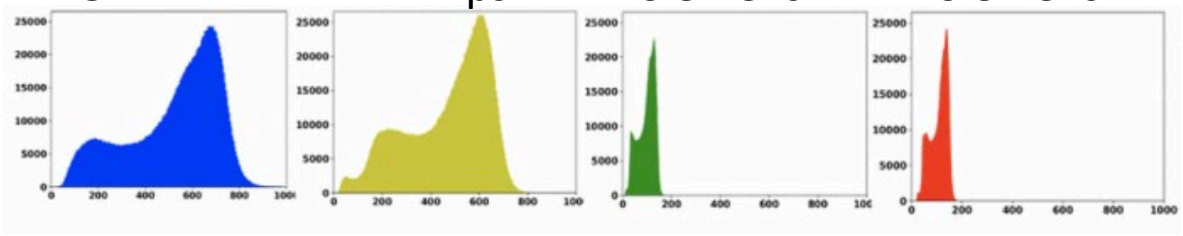
GE Philips Siemens P Siemens T



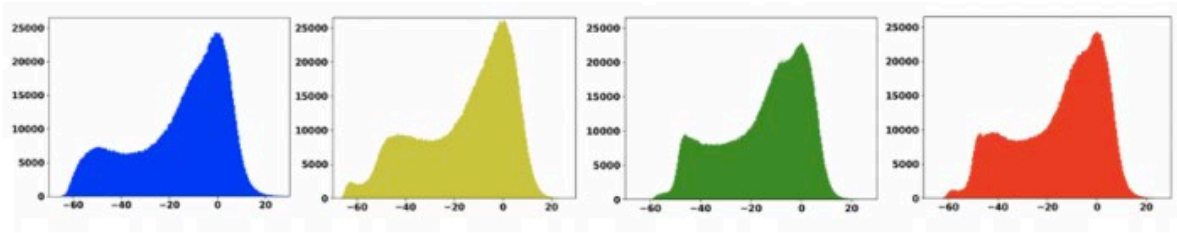
(b) Mean (SD) of absolute differences over all pairs.

Mean (SD) of absolute differences over all scanner pairs				
	Cortical Thickness (mm)		Volume (cm^3)	
Method	Entorhinal	Inferior Temporal	Hippocampus	Amygdala
RAW	0.62 (0.42)	0.46 (0.36)	0.30 (0.23)	0.25 (0.20)
WS	1.00 (0.73)*	0.63 (0.48)*	0.43 (0.52)*	0.23 (0.30)
RAVEL	0.84 (0.57)*	0.56 (0.41)*	0.41 (0.29)*	0.24 (0.21)
CALAMITI	0.87 (0.60)*	0.45 (0.32)	0.71 (0.54)*	0.30 (0.26)
MISPEL	**0.44** (0.34)*	**0.27** (0.28)*	0.30 (0.25)	**0.17** (0.15)*

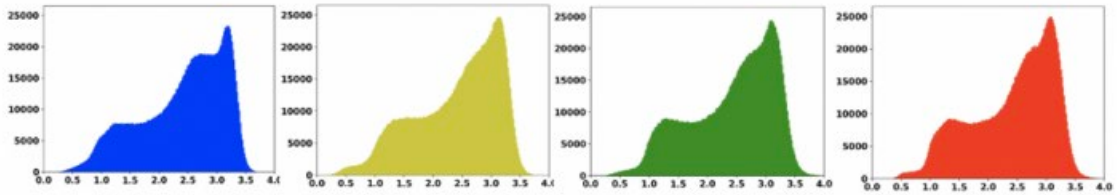
GE Philips Siemens P Siemens T



(a) **Technical variability in *Original* images.**



(c) **Histograms of WS-normalized images.**



(d) **Histograms of MISPEL-harmonized images.**

PET study (ABC-DS data)

PiB in Ventral Striatum region:

1. **Harmonized data:** to detect an effect size of 0.752 (Combat), n=58 participants.
2. **Non-harmonized data**: to detect an effect size of 0.719 (RAW), n=64 participants.
   - **For Tau in BRAAK4 :**
3. **Harmonized data**: to detect an effect size of 0.253 (Combat), n=120 participants
4. **Non-Harmonized data**: to detect an effect size of 0. 231(Raw), n=144 participants.
   - **For Tau in BRAAK 5:**
5. **Harmonized data**: to detect an effect size of 0.248 (Combat), n=124 participants
6. **Non-Harmonized data**: to detect an effect size of 0. 238(Raw), n=136 participants.
   - Sample size computed to achieve 80% power at a significance level of 0.05, two-sided.

- **In therapeutics trials this is going to make a huge difference.**

- Smoothing is believed to mitigate the PET scanner effects- however it increases the contamination from off target areas into ROI of interest potentially causing false positives (MK 6240), contamination of cerebellum and other off target areas.

ResearchGroup	Site	Sample_size
DS	UK	29
	UP	51
	WI	51
sibcontrol	UK	9
	UP	12
	WI	12

Caveat: We only looked at differences between controls and DS, very preliminary results.

What have we learned and how do we move forward?

- The effect sizes that we should be looking for should be the scanner variability + the targeted clinical difference (effect size).
- The methodological research for data harmonization is lacking behind in terms of funding compared to big data collection studies.
- Bigger efforts are needed to be able to support major data collection and the variability present due to technical differences.
- Statisticians should be included in the design phase of a study, so that the design is appropriate for the specific study.
- The NIH should work closely with a committee of specialists in the field about harmonization (statisticians, computer scientists, engineers) for the scientific community to establish standards when it comes to accounting for these differences.
- Our grant scoring criteria for multisite studies should include an evaluation for a technical variability mitigation plan, same as it includes for accounting for biological factors.

Some of our work on harmonization

Statistical Methods for Processing Neuroimaging Data from Two Different Sites with a Down Syndrome Population Application

Davneet S. Minhas, Zixi Yang, John Muschelli, Charles M. Laymon, Joseph M. Mettenburg, Matthew D. Zammit, Sterling Johnson, Chester A. Mathis, Ann D. Cohen, Benjamin L. Handen, William E. Klunk, Ciprian M. Crainiceanu, Bradley T. Christian & Dana L. Tudorascu

Conference paper | First Online: 05 June 2020

1249 Accesses

Part of the Communications in Computer and Information Science book series (CCIS,volume 1239)

> IEEE Int Conf Comput Vis Workshops. 2021 Oct;2021:3277-3286.
doi: 10.1109/ICCVW54120.2021.00367. Epub 2021 Nov 24.

Multi-scanner Harmonization of Paired Neuroimaging Data via Structure Preserving Embedding Learning

Mahbaneh Eshaghzadeh Torbati [1], Dana L Tudorascu [1], Davneet S Minhas [1], Pauline Maillard [2], Charles S DeCarli [2], Seong Jae Hwang [1]

Affiliations + expand

PMID: 34909551 PMCID: PMC8668020 DOI: 10.1109/ICCVW54120.2021.00367

Free PMC article

> Neuroimage. 2021 Dec 15;245:118703. doi: 10.1016/j.neuroimage.2021.118703. Epub 2021 Nov 1.

A multi-scanner neuroimaging data harmonization using RAVEL and ComBat

Mahbaneh Eshaghzadeh Torbati [1], Davneet S Minhas [2], Ghasan Ahmad [3], Erin E O'Connor [3], John Muschelli [4], Charles M Laymon [3], Zixi Yang [3], Ann D Cohen [5], Howard J Aizenstein [5], William E Klunk [5], Bradley T Christian [6], Seong Jae Hwang [7], Ciprian M Crainiceanu [4], Dana L Tudorascu [8]

Affiliations + expand

PMID: 34736996 PMCID: PMC8820090 DOI: 10.1016/j.neuroimage.2021.118703

Free PMC article

MISPEL:A deep learning approach for harmonizing multi-scanner matched neuroimaging data

Mahbaneh Eshaghzadeh Torbati, Davneet S. Minhas, Charles M. Laymon, Pauline Maillard, James D.Wilson, Chang-Le Chen, Ciprian M. Crainiceanu, Charles S. DeCarli, Seong Jae Hwang, Dana L.Tudorascu

doi: https://doi.org/10.1101/2022.07.27.501786

ComBat Harmonization: Empirical Bayes versus Fully Bayes Approaches

Maxwell Reynolds, Tigmanshu Chaudhary, Mahbaneh Eshaghzadeh Torbati, Dana L.Tudorascu, Kayhan Batmanghelich, the Alzheimer's Disease Neuroimaging Initiative

doi: https://doi.org/10.1101/2022.07.13.499561