# Exploratory Population Analysis with Unbalanced Optimal Transport

Samuel Gerber[1], Marc Niethammer[2], Martin Styner[2], Stephen Aylward[1]

[1] Kitware Inc, Carborro NC 27510, USA,
[2] University of North Carolina, Chapel Hill NC 27504, USA
samuel.gerber@kitware.com

**Abstract.** The plethora of data from neuroimaging studies provide a rich opportunity to discover effects and generate hypotheses through exploratory data analysis. Brain pathologies often manifest in changes in shape along with deterioration and alteration of brain matter, i.e., changes in mass. We propose a morphometry approach using unbalanced optimal transport that detects and localizes changes in mass and separates them from changes due to the location of mass. The approach generates images of mass allocation and mass transport cost for each subject in the population. Voxelwise correlations with clinical variables highlights regions of mass allocation or mass transfer related to the variable. We demonstrate the method on the white and gray matter segmentations from the OASIS brain MRI data set. The separation of white and gray matter ensures that optimal transport does not transfer mass between different tissues types and separates gray and white matter related changes. The OASIS data set includes subjects ranging from healthy to mild and moderate dementia, and the results corroborate known pathology changes related to dementia that are not discovered with traditional voxel-based morphometry. The transport-based morphometry increases the explanatory power of regression on clinical variables compared to traditional voxel-based morphometry, indicating that transport cost and mass allocation images capture a larger portion of pathology induced changes.

## 1 Introduction

Neurological disease and disorder manifest in subtle and varied changes in brain anatomy that can be non-local in nature and effect amounts of white and gray matter as well as relative positioning and shapes of local brain anatomy. To detect and quantify these changes is a primary goal of morphometry based population analysis. We propose a morphometry approach based on unbalanced optimal transport, termed UTM, that yields a voxelwise comparison but can detect global and regional deterioration of matter. The UTM formulation explicitly separates changes in matter volume from changes in matter location. This separation of effects leads to stronger correlations and more readily interpretable visualizations.

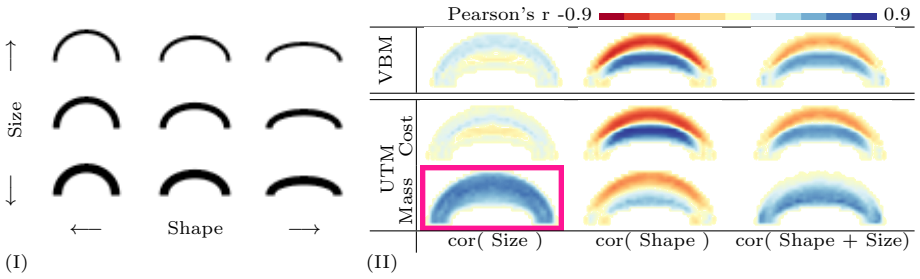


**Fig. 1.** (I) Illustration of unbalanced optimal transport on a toy data set of a 100 half–ellipses with different shapes (ratio of major to minor radius) and sizes (black pixel count). (II) Spatial correlations of size, shape and size + shape to VBM (pixel intensity) and UTM (mass allocation and transport costs). Positive and negative correlation indicate an increase and decrease of mass, cost, or intensity, respectively. Only UTM is able to detect size variation (highlighted box) as strongly positively correlated to mass allocation. Both UTM and VMB detect shape variation, with stronger correlations in transport cost than mass allocation in UTM. UTM identifies both change in shape and mass for the variable depending on shape plus size.

Optimal transport, as the name implies, solves the problem of transporting mass from a probability measure $\mu$ to a probability measure $\nu$, such that the cost of moving mass from the source $\mu$ to the target $\nu$ is minimized. Unbalanced optimal transport [11, 3] extends optimal transport to measures that do not need to have equal mass by adding a mechanism to add mass to the optimization problem. The solution of the unbalanced optimal transport yields a transport plan, or coupling, that measures local mass allocation and movement between source and target locations. We decompose these transport plans to measure for each subject, at each voxel, mass allocations and costs of mass transfer. These two measures explicitly separate changes due to differences in amount of matter from differences due to changes in location of matter related to relative position and shape of anatomies.

Voxel-based morphometry (VBM) [1] yields spatially localized changes in brain anatomy but has difficulty in discovering regionally or globally occurring changes. Figure 1 demonstrate the capability of UTM to detect changes in size on a toy example of a hundred half–ellipses with different ellipticity and thickness. Correlating the mass allocation and mass transfer images from UTM with either shape or size of the half–ellipses shows that the proposed method is capable of correctly attributing changes to either variation in shape or size. Traditional VBM results in weak correlation with changes in size and cannot attribute the source of the changes to ellipse shape or size.

Deformation-based morphometry (DBM) [2] addresses the issue of detecting global effects by comparing the parameters of non-linear spatial normalizations to a template. DBM results, typically shown as modes of variation, are more difficult to visualize and interpret. UTM combines the strengths of VBM and DBM methods. While the results are based on a voxelwise analysis and easily interpretable, the quantities compared stem from a global optimization problem which can detect global and regional effects.

A key observation is that VBM, TBM and DBM are driven by local image gradients and voxelwise measures ultimately lead to very similar results [8, Chaper 6]. UTM solves a global optimization problem based on the distribution of mass and the allocation of mass is not driven by image gradients. In UTM allocation of mass can be diffuse over a large region without the need for a smooth spatial normalization to distribute the image gradient driven warp over a larger region.

VBM has been shown to be sensitive to the particulars of the spatial normalization [4, 6]. UTM still depends on a rigid spatial normalization to bring the subjects into a common coordinate system were voxelwise comparisons are feasible. However, the optimal transport is insensitive to small miss-alignments in the spatial normalization step; small shifts in mass locations incur only small transport costs. After the spatial normalization UTM does not depend on a template, the mass allocation and transport cost images are based on an averaging of quantities derived from pairwise transport plans, which alleviates bias due to registrations to a single template.

The key contributions of this work are:

1. A voxel-wise morphometry approach capable of detecting regional and global changes.
2. A formulation of unbalanced optimal transport to detect and disentangle mass changes
3. A demonstration of the importance of incorporating mass imbalances to detect and localize neurodegenerative anatomical changes. (Section 4).from changes in mass location (Section 3).
4. An application of optimal transport to gray and white matter masks individually to avoid mass exchange between tissues and incommensurable measurements across subjects in MRI intensities (Section 4).

## 2 Related Work

Recent advances in the computation of optimal transport plans [5, 9] paved the way for a flurry of applications in machine learning and spurred interest in applications to medical image analysis. Gramfort et. al. [10] use optimal transport for improved averaging of neuroimaging data, Feydy et al. [7] use unbalanced optimal transport as a similarity measure for diffeomorphic registration and Kundu et. al. [12] formulate a DBM approach, TrBM, which replaces non-linear warps by optimal transport plans.

UTM integrates optimal transport to morphometric analysis in a different way from TrBM. TrBM yields a parametrization of the transport plans akin to DBM that captures global changes. UTM uses the transport plan to create voxelwise measures and adds mass imbalances into the transport plans, which prove to be an important indicator in neurodegenerative diseases. The TrBM is applied to graylevel MRI intensities, we choose to use the gray and white matter masks instead of MRI intensity values to avoid conversion of white to gray matter in the optimal transport optimization and to avoid difficulties in normalizing MRI intensities across subjects.

## 3 Unbalanced Optimal Transport Morphometry

Morphometry with unbalanced optimal transport follows the standard voxelwise morphometry pipeline but performs voxelwise statistical analysis on the mass allocation and transport cost images derived from the solution of unbalanced optimal transport between the subjects. Section 3.1 describes the unbalanced optimal transport problem and Section 3.2 the construction of the mass allocation and transport cost maps for each subject.

### 3.1 Unbalanced Optimal Transport

For two probability measures $\mu$ and $\nu$ on probability spaces $\mathbf{X}$ and $\mathbf{Y}$ respectively, a coupling of $\mu$ and $\nu$ is a measure $\pi$ on $\mathbf{X} \times \mathbf{Y}$ such that the marginals of $\pi$ are $\mu$ and $\nu$. The coupling $\pi$ defines a *transport plan* that captures how much mass $\pi(x, y)$ is transported from any $x \in \mathbf{X}$ to any $y \in \mathbf{Y}$. To define optimal transport and optimal couplings, we need a cost function $\mathsf{c}(x, y)$ on $\mathbf{X} \times \mathbf{Y}$ representing the work or cost needed to move a unit of mass from $x$ to $y$. An optimal coupling $\pi^*$ minimizes this cost over all choices of couplings $\mathcal{C}(\mu, \nu)$ between $\mu$ and $\nu$:

$$\pi^* = \operatorname*{argmin}_{\pi \in \mathcal{C}(\mu,\nu)} \int_{\mathbf{X}} \int_{\mathbf{Y}} \mathsf{c}(x, y) d\pi(x, y)\,. \tag{1}$$

For discrete distributions $\mu = \sum_1^n w(x_i)\delta(x_i)$ and $\nu = \sum_1^m v(y_i)\delta(y_i)$ with $\sum w(x_i) = \sum v(y_i) = 1$ the optimal transport problem can be solved by linear programming. To extend the formulation to deal with arbitrary positive measures $\mu$ and $\nu$ with mass imbalance $\Delta = |\sum_1^n v(y_i) - \sum_1^m w(x_i)|$, the linear program is modified to allow for the creation of mass by adding a new source location $z^s$ and target location $z^t$ and constraints to restrict the amount of mass added to be at most $w(x_i)$ and $v(x_i)$ at any target and source location, respectively. With these modifications the linear program reads:

$$\min_{\pi} \sum_{\substack{i=1,\dots,n \\ j=1,\dots,m}} \mathsf{c}(x_i, y_j)\pi(x_i, y_j) \,\text{s.t.}\, \begin{cases} \sum_j \pi(x_i, y_j) + \pi(x_i, z^t) = \mu(\{x_i\}) = w(x_i) \\ \sum_i \pi(x_i, y_j) + \pi(z^s, y_j) = \nu(\{y_j\}) = v(y_j) \\ \sum_j \pi(x_i, z^t) + \sum_j \pi(z^s, y_j) = \Delta \\ \pi(x_i, y_j) \geq 0\,, \pi(z^s, y_j) \geq 0\,, \pi(x_i, z^t) \geq 0 \end{cases} \tag{2}$$

The modifications result in a standard optimal transport problem that can, with only minor modifications, be solved by fast approximation algorithms for large data sets such as the Sinkhorn approach [5] or multiscale strategies [9].

An arbitrary cost $\mathsf{c}(x^s, y_j)$ and $\mathsf{c}(x_i, y^t)$ can be assigned to allocate mass, and restriction to only allocate $\Delta$ amount of mass can be relaxed or removed, striking a trade-off between the cost of creating mass and the cost of moving mass. For our application, we choose the allocation of mass at zero cost but only allow for the allocation of exactly $\Delta$ mass. This forces the transfer of all jointly available mass between source and target location and distribute the mass $\Delta$ optimally to reduce the cost of movement.

Solving the unbalanced optimal transport problem is convex and yields a global minimum without any parameter tuning, the only choice is the selection of a cost $\mathsf{c}$.

### 3.2 Construction of Mass Allocation and Transport Cost Images

For an image $X^k$ denote by $X^k(x_i)$ the associated non-negative value at voxel location $x_i$, which defines a measure $\mu^k = \sum_1^n X^k(x_i)\delta(x_i)$. To construct the voxelwise mass allocation and transport cost images we solve for optimal transport plans $\pi^*_{k,l}$ between all images $X^k$ and $X^l$. The variables $z^s$ and $z^t$ in equation 2 capture the amount of mass allocated when moving mass from $X^k$ to $X^l$. Denote by $z^s_{k,l}$ and $z^t_{k,l}$ the mass allocation variable associated with the optimal transport plan $\pi^*_{k,l}$. For subject $X_k$ the mass allocation image $M^k$ is constructed by $M^k(x_i) = \sum_l \left(\pi^*_{k,l}(z^s_{k,l}, x_i) - \pi^*_{k,l}(x_i, z^t_{k,l})\right)$ and the transport cost image by $C^k(x_i) = \sum_l \sum_j \pi^*_{k,l}(x_i, x_j)\mathsf{c}(x_i, y_j)$

The images $M^k$ and $C^k$ are smoothed with a small Gaussian to increase correlations between neighboring pixels. The smoothed images $M^k$ and $C^k$ replace the smoothed intensity or Jacobian determinant images in the statistical analysis of a VBM or TBM pipeline.

## 4 Application to OASIS Brain MRI

We apply UTM to the OASIS brain data set [13]. Code to replicate the results is located at https://github.com/KitwareMedical/UTM. The OASIS database consists of T1 weighted MRI of 416 subjects aged 18 to 96. One hundred of the subjects over the age of 60 are diagnosed with mild to moderate dementia. The images in the OASIS data set are already skull-stripped, gain-field corrected and registered to the Talaraich atlas space [14] with a 12-parameter affine transform. To focus on dementia related effects we restrict the analysis to 137 patients from age 60 to 80. This set of patients contains 66 healthy patents and 71 patients with a diagnosis of very mild to moderate dementia as established by a clinical dementia rating (CDR, 0 = normal, 0.5 = very mild, 1 = Mild, 2 = Moderate ). In addition to CDR the data set contains information about Age and a mini mental state examination score (MMSE, count of correct answers with a perfect score of 30).

We construct mass allocation and transport cost images for white and gray matter segmentation masks individually. However, the method can without modification be applied to any scalar image or with the definition of a corresponding cost function even to vector or tensor valued images. For a sensible application the only requirement is that the measurement at voxels are commensurable across subjects. For the cost $\mathsf{c}$ we use the squared euclidean distance between the voxel locations to introduce a preference of many small mass transports over single larger mass transfers. To reduce computation costs, we subsample the images to $44 \times 52 \times 44$, which yields computations times of approximately one

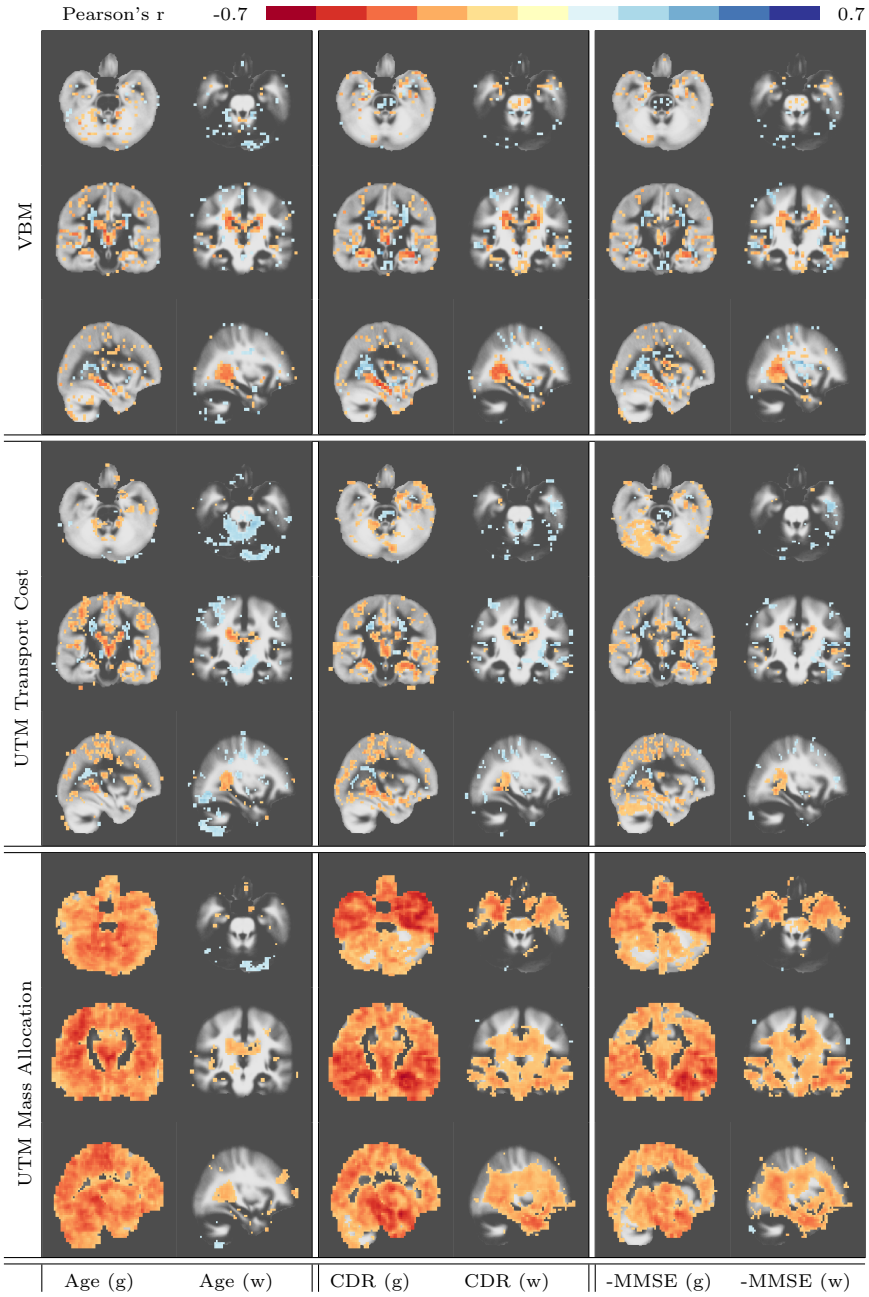


**Fig. 2.** Correlation of Age, MMSE and CDR with VBM, UTM mass imbalances and UTM transport costs on (g) gray and (w) white matter segmentations. Correlations are only shown at locations permutation tested p–value less than 0.05 (1000 permutations). The background image are average white and gray matter segmentations.

second using a multiscale optimal transport solver [9]. Both VBM intensities of the white and gray matter masks and the UTM mass allocation and transport cost images are smoothed with a Gaussian with 3mm standard deviation before computing correlations or regression models. The multiscale optimal transport solver introduces an additional smoothing of the mass allocation instead of focusing allocations at the boundary regions. This yields visually more pleasing results but does not change the overall results drastically. This effect could also be achieved by limiting the amount of mass allocated at any given location.

Figure 2 shows voxelwise correlations to Age, CDR and negative MMSE and compares to a traditional intensity VBM analysis. The most striking difference is that the loss in overall gray matter is immediately visible in the mass allocation part of UTM. The overall loss related to Age is differently distributed from the loss related to CDR and MMSE. CDR and MMSE have more pronounced gray matter deterioration in the cerebellum and in the temporal lobe, known to be involved with cognitive decline. A second striking difference is the loss of white matter around the ventricles in the area of the hippocampus that is significant in CDR and MMSE but not with respect to Age. The observed associations are consistent with findings reported in the literature on Alzheimer's disease and mild cognitive impairment. The VBM analysis is not capable to discover the global and regionally varying diminishing of gray and white matter.

We compare the explanatory power of UTM to VBM with elastic net regression [15] of transport cost and transport images and intensity images on clinical parameters. We set the elastic net regression penalty trade-off to $\alpha = 0.1$, a large ridge penalty and small sparsity penalty trade–off. Table 1 reports crossvalidated root mean square error (RMSE) using regularizations based on the minimal RMSE and the strongest regularization that results in an RMSE within one standard deviation of the minimal RMSE, as advocated by [15]. The minimal regularization appears to overfit with a much smaller $R^2$ but only a minor reduction in RMSE. Regression on UTM images outperforms VBM on all clinical variables. For the weaker regularized models UTM results in a larger reduction in RMSE combined with smaller reduction in $R^2$. These results indicate that UTM captures more neurodegenerative information than VBM.

| Model | RMSE 1se | $R^2$ 1se | RMSE min | $R^2$ min |
|---|---|---|---|---|
| VBM, Age | 4.89 | 0.24 | 4.81 | 0.95 |
| UTM, Age | **4.51** | **0.39** | 4.29 | 0.72 |
| VBM, MMSE | 3.80 | 0.21 | 3.61 | 0.97 |
| UTM, MMSE | **3.61** | **0.25** | 3.27 | 0.54 |
| VBM, CDR | 0.36 | 0.21 | 0.33 | 0.69 |
| UTM, CDR | **0.32** | **0.40** | 0.30 | 0.72 |

**Table 1.** Elastic net regression on UTM transport cost and mass allocation images versus regression on VBM intensity images. Reported are 10-fold crossvalidated root mean square error (RMSE) and $R^2$ with the amount of regularization either selected by the minimal RMSE (min) or the most parsimonious, strongest regularized, model within one standard deviation of the minimal RMSE (1se).

## 5 Conclusion

The paper demonstrates that UTM captures changes related to size not detected by traditional VBM and can attribute effects to mass transfer and mass allocation. UTM captures regional and global changes while retaining the benefits of a voxelwise visualization and analysis of results. The results in demonstrate the signficance of incorporating mass imbalance into the morphometry framework.

*Acknowledgments* This work was funded, in part, by NIH grants R01EB021391, R01HD055741, U54HD079124, NINDS grants R42NS086295, R44NS081792, NCI grant R44CA165621, NSF grant ECCS-1711776 and NIBIB grant R01EB021396.

## References

1. J. Ashburner and K. J. Friston. Voxel-based morphometrythe methods. *Neuroimage*, 11(6):805–821, 2000.
2. J. Ashburner, C. Hutton, R. Frackowiak, I. Johnsrude, C. Price, K. Friston, et al. Identifying global anatomical differences: deformation-based morphometry. *Human brain mapping*, 6(5-6):348–357, 1998.
3. J.-D. Benamou. Numerical resolution of an unbalanced mass transport problem. *ESAIM: Mathematical Modelling and Numerical Analysis*, 37(5):851–868, 2003.
4. F. L. Bookstein. voxel-based morphometry should not be used with imperfectly registered images. *Neuroimage*, 14(6):1454–1462, 2001.
5. M. Cuturi. Sinkhorn distances: Lightspeed computation of optimal transport. In *Advances in neural information processing systems*, pages 2292–2300, 2013.
6. C. Davatzikos. Why voxel-based morphometric analysis should be used with great caution when characterizing group differences. *Neuroimage*, 23(1):17–20, 2004.
7. J. Feydy, B. Charlier, F.-X. Vialard, and G. Peyré. Optimal transport for diffeomorphic registration. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 291–299. Springer, 2017.
8. R. S. Frackowiak. *Human brain function*. Academic press, 2004.
9. S. Gerber and M. Maggioni. Multiscale strategies for computing optimal transport. *Journal of Machine Learning Research*, 18(72):1–32, 2017.
10. A. Gramfort, G. Peyré, and M. Cuturi. Fast optimal transport averaging of neuroimaging data. In *International Conference on Information Processing in Medical Imaging*, pages 261–272. Springer, 2015.
11. K. Guittet. *Extended Kantorovich norms: a tool for optimization*. PhD thesis, INRIA, 2002.
12. S. Kundu, S. Kolouri, K. I. Erickson, A. F. Kramer, E. McAuley, and G. K. Rohde. Discovery and visualization of structural biomarkers from mri using transport-based morphometry. *NeuroImage*, 167:256–275, 2018.
13. D. S. Marcus, A. F. Fotenos, J. G. Csernansky, J. C. Morris, and R. L. Buckner. Open access series of imaging studies: longitudinal mri data in nondemented and demented older adults. *Journal of cognitive neuroscience*, 22(12):2677–2684, 2010.
14. J. Talairach and P. Tournoux. *Co-planar stereotaxic atlas of the human brain. 3-Dimensional proportional system: an approach to cerebral imaging*. Thieme, New York, 1988.
15. H. Zou and T. Hastie. Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society, Series B*, 67:301–320, 2005.