

Contents lists available at [ScienceDirect](https://www.sciencedirect.com)

# Medical Image Analysis

journal homepage: [www.elsevier.com/locate/media](http://www.elsevier.com/locate/media)

## WarpDrive: Improving spatial normalization using manual refinements

Simón Oxenford<sup>a,1,\*</sup>, Ana Sofía Ríos<sup>a</sup>, Barbara Hollunder<sup>a,b,c</sup>, Clemens Neudorfer<sup>a,d,e</sup>, Alexandre Boutet<sup>f,g,h</sup>, Gavin J.B. Elias<sup>f,g</sup>, Jurgen Germann<sup>f,g</sup>, Aaron Loh<sup>f,g</sup>, Wissam Deeb<sup>i,j</sup>, Bryan Salvato<sup>k</sup>, Leonardo Almeida<sup>l</sup>, Kelly D. Foote<sup>m</sup>, Robert Amaral<sup>n</sup>, Paul B. Rosenberg<sup>o</sup>, David F. Tang-Wai<sup>g,p</sup>, David A. Wolk<sup>q</sup>, Anna D. Burke<sup>r</sup>, Marwan N. Sabbagh<sup>r</sup>, Stephen Salloway<sup>s,t</sup>, M. Mallar Chakravarty<sup>n,u,v</sup>, Gwenn S. Smith<sup>n</sup>, Constantine G. Lyketsos<sup>n</sup>, Michael S. Okun<sup>m</sup>, William S. Anderson<sup>w</sup>, Zoltan Mari<sup>w,x</sup>, Francisco A. Ponce<sup>r</sup>, Andres Lozano<sup>f,g</sup>, Wolf-Julian Neumann<sup>a</sup>, Bassam Al-Fatly<sup>a</sup>, Andreas Horn<sup>a,d,e</sup>

<sup>a</sup> Movement Disorders and Neuromodulation Unit, Department of Neurology, Charité - Universitätsmedizin Berlin, Corporate Member of Freie Universität Berlin and Humboldt-Universität zu Berlin, Berlin, Germany

<sup>b</sup> Einstein Center for Neurosciences Berlin, Charité - Universitätsmedizin Berlin, Berlin, Germany

<sup>c</sup> Berlin School of Mind and Brain, Humboldt-Universität zu Berlin, Berlin, Germany

<sup>d</sup> Brain Modulation Lab, Department of Neurosurgery, Massachusetts General Hospital, Boston, MA 02114, United States

<sup>e</sup> Center for Brain Circuit Therapeutics Department of Neurology Brigham & Women's Hospital, Harvard Medical School, Boston, MA, United States

<sup>f</sup> Division of Neurosurgery, Department of Surgery, University Health Network and University of Toronto, Toronto, ON M5T2S8, Canada

<sup>g</sup> Krembil Research Institute, University of Toronto, Toronto, ON M5T2S8, Canada

<sup>h</sup> Joint Department of Medical Imaging, University of Toronto, Toronto, ON M5T1W7, Canada

<sup>i</sup> UMass Chan Medical School, Department of Neurology, Worcester, MA 01655, United States

<sup>j</sup> UMass Memorial Health, Department of Neurology, Worcester, MA 01655, United States

<sup>k</sup> University of Florida Health Jacksonville, Jacksonville, FL, United States

<sup>l</sup> Department of Neurology, University of Minnesota, Twin Cities Campus, Minneapolis, MN, United States

<sup>m</sup> Norman Fixel Institute for Neurological Diseases, Departments of Neurology and Neurosurgery, University of Florida, Gainesville, FL, United States

<sup>n</sup> Cerebral Imaging Centre, Douglas Research Centre, Montreal, QC, Canada

<sup>o</sup> Department of Psychiatry and Behavioral Sciences and Richman Family Precision Medicine Center of Excellence, School of Medicine, Johns Hopkins University, Baltimore, MD, United States

<sup>p</sup> Department of Medicine, Division of Neurology, University Health Network and University of Toronto, Toronto, ON M5T2S8, Canada

<sup>q</sup> Department of Neurology, University of Pennsylvania, Philadelphia, PA, United States

<sup>r</sup> Barrow Neurological Institute, Phoenix, AZ, United States

<sup>s</sup> Department of Psychiatry and Human Behavior and Neurology, Alpert Medical School of Brown University, Providence, RI, United States

<sup>t</sup> Memory & Aging Program, Butler Hospital, Providence, United States

<sup>u</sup> Department of Psychiatry, McGill University, Montreal, QC, Canada

<sup>v</sup> Biological and Biomedical Engineering, McGill University, Montreal, QC, Canada

<sup>w</sup> Johns Hopkins School of Medicine, Baltimore, MD, United States

<sup>x</sup> Cleveland Clinic Lou Ruvo Center for Brain Health, Las Vegas, NV, United States

### ARTICLE INFO

#### Keywords:

Image normalization  
Interactive registration  
Deep brain stimulation

### ABSTRACT

Spatial normalization—the process of mapping subject brain images to an average template brain—has evolved over the last 20+ years into a reliable method that facilitates the comparison of brain imaging results across patients, centers & modalities. While overall successful, sometimes, this automatic process yields suboptimal results, especially when dealing with brains with extensive neurodegeneration and atrophy patterns, or when high accuracy in specific regions is needed. Here we introduce WarpDrive, a novel tool for manual refinements of image alignment after automated registration. We show that the tool applied in a cohort of patients with Alzheimer's disease who underwent deep brain stimulation surgery helps create more accurate representations of the data as well as meaningful models to explain patient outcomes. The tool is built to handle any type of 3D

\* Corresponding author.

E-mail address: [simon.oxenford@charite.de](mailto:simon.oxenford@charite.de) (S. Oxenford).

<sup>1</sup> Present address: Bonhoefferweg 3, 10117 Berlin, Movement Disorders and Neuromodulation Unit, Department of Neurology, Charité -Universitätsmedizin Berlin, corporate member of Freie Universität Berlin and Humboldt-Universität zu Berlin, Berlin, Germany

<https://doi.org/10.1016/j.media.2023.103041>

Received 16 May 2023; Received in revised form 8 November 2023; Accepted 17 November 2023

Available online 19 November 2023

1361-8415/© 2023 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

imaging data, also allowing refinements in high-resolution imaging, including histology and multiple modalities to precisely aggregate multiple data sources together.

## 1. Introduction

Spatial normalization is a particularly important form of image registration in which one (moving) brain is registered to a (fixed) population average template. Over the last several decades, this process has almost exclusively been carried out using nonlinear methods with transforms represented by deformation or warp fields. These fields denote vectors mapping from each point on the moving image to the corresponding point in the fixed image. In the neuroimaging field, the Montreal Neurological Institute (MNI) space (e.g., ICBM 2009b, [Fonov et al., 2009](#)) has often been used as the fixed image and, accordingly, multiple cortical ([Desikan et al., 2006](#); [Tzourio-Mazoyer et al., 2002](#)) and subcortical ([Ewert et al., 2017](#); [Xiao et al., 2017](#)) atlases have been constructed in this space. Critically, since modern registration methods are invertible, they can be used to register such atlases back to the individual brain.

Spatial normalization is a critical step in neuroimaging analysis, as it has made it possible to compare findings across brains, cohorts, and populations. It is included as a step in common fMRI workflows ([Esteban et al., 2019](#)), and in more general-purpose neuroimaging pipelines ([Gorgolewski et al., 2011](#)). It is also a key component in the field of deep brain stimulation (DBS) imaging, if the aim is to make DBS electrode reconstructions comparable across patients ([Treu et al., 2020](#)). Unlike fMRI studies, where data is often smoothed by Gaussian kernels ranging up to 8 mm in full-width half-maximum, in DBS it is of utmost importance to accurately represent minute differences in stimulation sites ([Horn et al., 2019](#); [Bingham & McIntyre, 2022](#)). Normalization algorithms have evolved over time and have been compared with each other, showing a generally accurate performance in the cortex ([Klein et al., 2009](#)), as well as subcortex ([Ewert et al., 2019](#)).

Moving forward, the field of neuroimaging is increasingly acquiring high(er)-resolution and multi-modal data. This is seen in the field of ultra-high field fMRI, where the increasing resolution is a starting point for novel aims, such as modeling layer-specific activations ([Bandettini et al., 2021](#)). There is also an increasing interest in fusion between histology and postmortem resources with MRI ([Edlow et al., 2019](#); [Paquola et al., 2021](#)). Finally, the DBS imaging field is trying to achieve higher accuracy in smaller regions of interest: resolving directional leads with submillimeter contact-to-contact distances ([Dembek et al., 2021](#)) and using ultrahigh-field MRI for planning ([Forstmann et al., 2017](#); [Isaacs et al., 2021](#)), for example.

Unfortunately, with increasing resolution, automatic registrations seem to become more challenging and less successful, rather than easier and more accurate ([Edlow et al., 2019](#)). Furthermore, brains with substantial atrophy make registrations less straight-forward ([Avants et al., 2008](#)). Additionally, for MRI, using higher field strengths also introduces higher field inhomogeneities and hence nonlinear distortion artifacts ([Sumanaweera et al., 1994](#)).

When automatic methods fail, a promising concept is to rely on interactive approaches. For example, user-defined source and target landmarks can be selected to locally guide registrations ([Sharp et al., 2010](#); [Wu, 2014](#)). Additionally, Zhou et al. proposed a method to refine B-Spline transforms via manually dragging control points ([Zhou and Xie, 2013](#)). An alternative is to use masks, which restrict the registration to a specific region of interest ([Godley et al., 2009](#)). However—and unlike the field of image segmentation, where a wide range of interactions have been established, including classic region growing approaches ([Kikinis et al., 2014](#); [Vezhnevets and Konouchine, 2005](#)), and modern active learning frameworks ([Diaz-Pinto et al., 2022](#); [Nath et al., 2021](#)) (for a review see [Zhao and Xie, 2013](#))—interactions in the context of image registration have been more limited and less common. This

might be because of the labor-intensive nature of this task; lack of intuitive and easy-to-use tools; or unavailability within common image processing software.

Here we introduce WarpDrive, a novel tool and method that facilitates refinements of warps by manual interaction. WarpDrive also uses source and target landmark mapping, but rather as a backend to different interaction methods which we developed envisioning a more intuitive and user-friendly tool. While this concept does not solve current limitations of automatic registrations, it provides users with tools to refine their results, especially in circumscribed key areas of interest. It is also not intended to replace manual segmentations—the gold standard for patient specific analysis—but is rather conceived as a tool to accurately link patient-specific and normative spaces.

We see particular benefit for research questions that study localized brain regions (such as layer-fMRI analyses, studies of hippocampal subfields or focal brain stimulation). Indeed, our method has recently been applied in a study within the aforementioned, most challenging scenario: DBS for Alzheimer’s disease (AD) ([Ríos et al., 2022](#)), where brain atrophy is substantial but millimeters of registration accuracy matter. In it, Ríos et al. found an optimal stimulation site that robustly explained variance in clinical improvements across 46 patients. While the Ríos et al. study applied WarpDrive, a thorough methodological write-up and investigation of its effects has not been included.

Here, we revisit the same dataset to further characterize the impact WarpDrive had on the results of the analysis. Furthermore, we apply WarpDrive to another set of images from patients featuring brain atrophy on MRI scans (without DBS). Finally, we provide examples of WarpDrive refinements in registrations of high-resolution images into standard space; registrations of template space to a histological slice; and group registrations to create a brain template.

## 2. Methods

### 2.1. WarpDrive toolbox

We developed the WarpDrive toolbox for manual refinement of misalignments between source and target spaces after image registration. One of WarpDrive’s novel features is the user interface and logic through which landmarks are defined. Specifically, WarpDrive uses a set of corrections, where each correction is comprised by one or more source landmarks, their corresponding target landmarks, and a radial kernel bandwidth, which will determine the influence of the correction on the surrounding displacement field.

We next introduce how source and target landmarks can be placed using the three tools we developed for this application; and then how they translate to the output displacement field. We refer to [Fig. 1](#) for an illustration of the tools, and to supplementary material for an overview of the user interface.

#### 2.1.1. WarpDrive tools

- Point to point. This tool places a single source landmark with the first click and the corresponding target landmark with the second.
- Draw. By clicking and dragging the user can draw a freeform line, from which equidistant points are sampled to define source landmarks. In one mode of operation, the user can then draw another line, which will be again sampled defining the corresponding target landmarks. In another mode, when anatomical segmentation models are present, the target landmarks are sampled from the closest segmentation model. This mode can be helpful when recognizing structure outlines that do not precisely match an atlas model after

image normalization (e.g., there is an offset between the lateral border of a brain nucleus on the MRI and its corresponding border in the brain atlas). The user can draw the outlines in the image, and the corresponding source and target landmarks will automatically be defined, saving the user from multiple mouse interactions.

- **Smudge.** In this tool, by clicking and dragging, the user is updating a temporary displacement field ( $u$ ) following cursor displacement in real time. Specifically,

$u_t(\vec{r}) = u_{t-1}(\vec{r}) + (\|\vec{r}_{Ct-1} - \vec{r}_{Ct}\|) \times \psi(\vec{r} - \vec{r}_{Ct})$ , where  $\vec{r}_{Ct}$  is the position of the cursor at a given time  $t$ , and  $\psi$  is the gaussian radial basis function  $\psi(r) = e^{-(r\epsilon)^2}$ . When the interaction finishes, source landmarks are sampled along the trajectory in which the cursor moved; and target landmarks are defined by transforming the source landmarks with the temporary displacement field. Interaction points are not required to be aligned to the directions in which the image was sampled along. Rather, slice view orientations can be arbitrarily set, and the  $(x,y)$  in-plane position of interaction is transformed to its respective  $(r, a, s)$  spatial coordinates and  $(i, j, k)$  voxel indexes, where the warp modifications are embedded. This type of interaction provides intuitive feedback since the temporary displacement field is applied to the image in real time, providing a preview visualization of how the output will look like.

### 2.1.2. Displacement field computation

These tools provide means of placing source and target landmarks, which are fed into a registration algorithm to compute the output displacement field. For this, we base our implementation on the Plastimatch algorithm, published by Shusharina and Sharp (Shusharina and Sharp, 2012). Next, we introduce how the deformation field is calculated using the same notation as them, making emphasis on the novel

additions, and refer to their publication for details on more extensive mathematical formulation.

The value at each position  $\vec{r}$  of the displacement field  $u$  is calculated as:

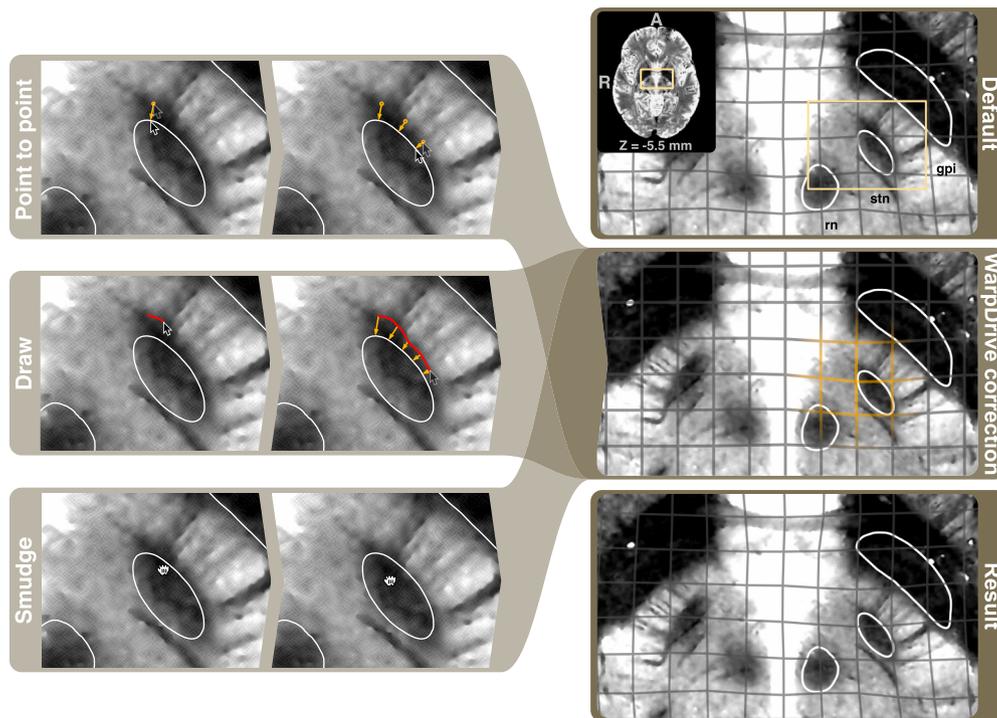
$$u(\vec{r}) = \sum_{p=x,y,z} \vec{e}_p \sum_{i=1}^M \alpha_{ip} \psi_i(\|\vec{r} - \vec{r}_{Ti}\|)$$

Where  $\vec{e}_p$  is the unit vector along the p-axis;  $M$  is the number of landmarks;  $\alpha_{ip}$  coefficients are determined such that corresponding source ( $\vec{r}_S$ ) and target ( $\vec{r}_T$ ) landmarks match:  $\vec{r}_{Si} + u(\vec{r}_{Si}) = \vec{r}_{Ti}$  (see Shusharina and Sharp, 2012 for more details); and  $\psi_i$  is the gaussian radial basis function  $\psi_i(r) = e^{-(r\epsilon_i)^2}$ .

Here, novel to this implementation, the radial kernel bandwidth  $\epsilon_i$  is specific for each correction (and therefore for each landmark that belongs to the correction). This allows to have multiple corrections that have variable influence on the output displacement field. The user manually sets this bandwidth value via a slider, or keyboard shortcuts; and a sphere outline of this radius is shown in the viewers during the interactions to indicate the area being influenced.

### 2.1.3. WarpDrive snap

Apart from the tools that allow for manual interactions with the displacement field, we also included a feature to locally improve the alignment between source and target images in an automated way. Specifically, with the Snap operation, an automated registration routine is run after a correction restricting the registration to the place where the interaction was done. Then, the output is applied to the target points belonging to the respective correction. This allows for a semi-automated correction, where the potentially misplaced target points are



**Fig. 1.** WarpDrive tools overview. WarpDrive features different tools for the user to interact with and refine the normalization. In this figure the tools are illustratively explained. With the point-to-point tool (top left) the user can place source and target points to guide refinement locally. When drawing (center left), the user draws an outline of a structure of interest which is misaligned with a target structure. Sampled points along the line are corresponded to the nearest atlas structure. When smudging (bottom left), the user clicks and drags to displace the image along the pointer movement. All these corrections will result in a new displacement field (center right) which is stacked on top of the original transform (top right), resulting in a new transform (bottom right). The WarpDrive correction is applied to the result of a default Lead-DBS normalization of the subject's images (7T 0.12 mm in-plane resolution; [https://figshare.com/articles/dataset/BWH\\_06\\_2022\\_7T/20,102,912](https://figshare.com/articles/dataset/BWH_06_2022_7T/20,102,912)). Atlas outlines, from Ewert et al. (2017): rn: Red nucleus; stn: Subthalamic nucleus; gpi: Globus pallidus internus.

automatically adjusted following an automated process.

## 2.2. ADNI cohort

118 participants (mean age:  $75 \pm 7.8$  years; 56 females) from the Alzheimer's disease Neuroimaging Initiative (ADNI) database ([adni.loni.usc.edu](http://adni.loni.usc.edu)) were retrieved. This cohort was defined based on participants for which both *Harmonized Hippocampal Protocol* (HarP) manual hippocampus segmentations (Boccardi et al., 2015), as well as *Spatial Pattern of Abnormality for Recognition of Early Alzheimer's disease* (SPARE-AD) scores (Davatzikos et al., 2009) were available.

Briefly, the SPARE-AD scores constitute an imaging derived measure of AD-like brain atrophy. More positive values indicate brains with increasing atrophy while more negative values indicate normal brain characteristics (Davatzikos et al., 2009).

As for the HarP protocol, it specifies a method for manual segmentation of the hippocampus (Boccardi et al., 2015) which was previously applied on the aforementioned ADNI cohort (Boccardi et al., 2015). The HarP protocol has been also applied on the MNI space to provide normative hippocampal segmentations (Wolf et al., 2017), which were used here to derive automatic atlas-based segmentations (see section below).

## 2.3. Atlas-based segmentation

The ADNI cohort was normalized using Advanced Normalization Tools (ANTs; <https://stnava.github.io/ANTs/>), without further manual refinement, using default parameters as implemented in the Lead-DBS toolbox (<https://www.lead-dbs.org/>). Based on the resulting warp-field, the template HarP hippocampus segmentation was transformed from MNI to patient space to derive an automatic hippocampus segmentation. In this study we refer to a segmentation derived from transforming an atlas structure to patient space as *atlas-based* or *normalization-based* segmentation. Similarities between manual versus normalization-based segmentations were quantified by calculating DICE scores (Dice, 1945). These were then correlated with the SPARE-AD brain atrophy scores.

We then applied WarpDrive on the entire ADNI cohort and obtained a new set of normalization-based segmentations and their respective DICE scores. We then compared the DICE distribution for the automated versus the WarpDrive based one considering the whole cohort; and only the bottom and top 25 % of the cohort when sorted by the SPARE-AD scores. With this analysis we wanted to distinguish the effects of WarpDrive across more vs. less atrophied brains.

## 2.4. Alzheimer's disease deep brain stimulation cohort (AD-DBS)

46 patients—from across seven international centers—with mild AD diagnosis were included in this study (patient mean age:  $67 \pm 7.9$  years; 23 females). This patient cohort has been well characterized in a prior retrospective trial (Ríos et al., 2022), to which we refer for details on patient selection and additional information. Briefly, all 46 patients underwent DBS targeting the descending columns of the fornix and were clinically evaluated using the ADAS-cog 11 score before and one year after surgery. Pre- and post-operative imaging, together with post-operative stimulation settings were retrieved for all patients, including stimulation amplitude, frequency, and active contacts.

## 2.5. Lead-DBS processing pipeline

The AD-DBS cohort was processed using the Lead-DBS toolbox including the following steps in the pipeline: (i) pre- and post-operative image registration; (ii) patient to template image normalization; (iii) brain-shift correction; (iv) DBS electrode reconstruction; and (v) electric field modeling estimation from stimulation settings. For in-depth details on how each of these steps were applied in this dataset we refer to the

original publication (Ríos et al., 2022).

Particularly, the normalization step was carried out using default settings in Lead-DBS (Ewert et al., 2019), which builds on top of an optimized multi-spectral registration implemented in ANTs (Avants et al., 2008). The specific parameters that were used are reproduced in an example `antsRegistration` call in the supplementary material. Following ANTs-based normalization, scans transformed into MNI space were refined using WarpDrive, leading to a second set of results (ANTs+WarpDrive). To do so, Ríos et al. employed a ground-truth high-resolution atlas template of the human fornix (Neudorfer et al., 2020) for WarpDrive based refinements.

We then assessed whether these refinements further optimized the registration by transforming native images, measuring their similarity against the MNI template, and comparing between ANTs and ANTs+WarpDrive groups. While before we used DICE to evaluate similarity in segmentation (binary) images, for this analysis, to evaluate similarity in gray-scale images, we used the mutual information metric (Hermosillo et al., 2002), which was also used as the objective function by the registration algorithm during the optimization process.

Another result we studied was the effect WarpDrive had on the clustering of electrode contacts. For this, we measured the distance from each contact to the average contact position in their respective hemisphere. We then compared the distributions of distance to mean active contact using a paired *t*-test between the ANTs and ANTs+WarpDrive groups. It should be noted that this is an indirect measure of registration improvement, since normalization only accounts for part of the variability of electrode dispersion, apart from variance introduced by surgical planning, frame accuracy and bias introduced by electrode reconstructions.

## 2.6. Lead-DBS sweetspot explorer

Following refinement, the cohort was then loaded into the Lead-DBS Sweetspot Explorer (Neudorfer et al., 2023), where the magnitude of simulated electric fields was correlated in a voxel-wise manner with improvements in ADAS-cog 11 scores across patients. This led to a correlation map in MNI space (i.e., *sweetspot*) where positive voxel values indicated a positive relationship between electric field magnitudes and beneficial clinical outcomes.

Ríos et al. evaluated their derived sweetspot model using k-fold cross-validations: improvement scores for iteratively left-out patients were estimated based on the similarities between their electric fields and the sweetspot (which had been iteratively calculated based on included patients). The same was done also in a circular analysis, namely, improvements for all patients were estimated using the sweetspot built from all data. The performance of this model was then assessed by evaluating the association between estimated and empirical improvements. In the present study we repeated the analysis performed by Ríos et al., but this time for the purely automated registration (ANTs) group. The circular analysis was employed, as model performance was not the essence of this study, but rather the comparison between the groups (ANTs vs. ANTs+WarpDrive). In essence, this circular analysis estimates how much variance in clinical improvements the sweetspot map can explain across the entire dataset. For this analysis, significance tests were performed using Monte Carlo permutations (Nichols and Holmes, 2002).

## 2.7. Comparison against state-of-the-art registration strategies

While our main focus was to quantify the impact of manual refinements on top of the automated normalization results (based on ANTs), we wanted to probe how well recently introduced deep learning based approaches would perform in comparison. Thus, we repeated our main analyses using EasyReg (Iglesias, 2023) and compared the outcome measures with the ones from ANTs and ANTs+WarpDrive.

## 2.8. Additional qualitative analysis

We next performed additional qualitative analysis showcasing the usability of WarpDrive for other applications. First, we showcase the registration refinement of an ultra-high resolution postmortem brain template (100  $\mu\text{m}$  isotropic resolution; Edlow et al., 2019) with the MNI template within the fornix region, demonstrating—in conjunction with the AD-DBS analysis—the use of WarpDrive to derive accurate representations of multimodal data aggregation.

Second, we demonstrate feasibility of using WarpDrive to refine the registration to a 2D histological slice. In particular, we refined the fit of the subthalamic nucleus from MNI space to a 1  $\mu\text{m}$  BigBrain slice (Amunts et al., 2013) starting from an optimized mapping between the two spaces (Xiao et al., 2019).

Third, we applied WarpDrive during the process of creating an atlas template from 12 subjects (age range 31–40) that were scanned at 7T within the Amsterdam Ultra-high field adult lifespan database (AHEAD; Alkemade et al., 2020). First, we created a template in fully automatized fashion, and then refined it, focusing on the subthalamic nucleus as our region of interest. We employed the atlas creation routine distributed in the ANTs installation (antsMultivariateTemplateConstruction2) using SyN transforms. Then, we manually refined registrations of the images to the template using WarpDrive, and particularly focused on accuracy of the subthalamic nucleus area.

Finally, we applied WarpDrive in an abdomen dataset from the 2021 Learn2Reg Challenge (Clark et al., 2013; Hering et al., 2023). We took the 8 training cases from the first task of the challenge (CT-MR thorax-abdomen intra-patient registration) and we did intra-patient registration of the CTs to MRs using ANTs quick SyN approach with neighborhood cross correlation metric (Avants et al., 2008). Then, we applied WarpDrive to the resulting deformation field, focusing on the

liver, fixing misalignments between MR and CT images. We then computed the DICE and 95 % Hausdorff distance (Huttenlocher et al., 1993) metrics between the transformed CT liver labels (based on ANTs and ANTs+WarpDrive) and their corresponding MR labels.

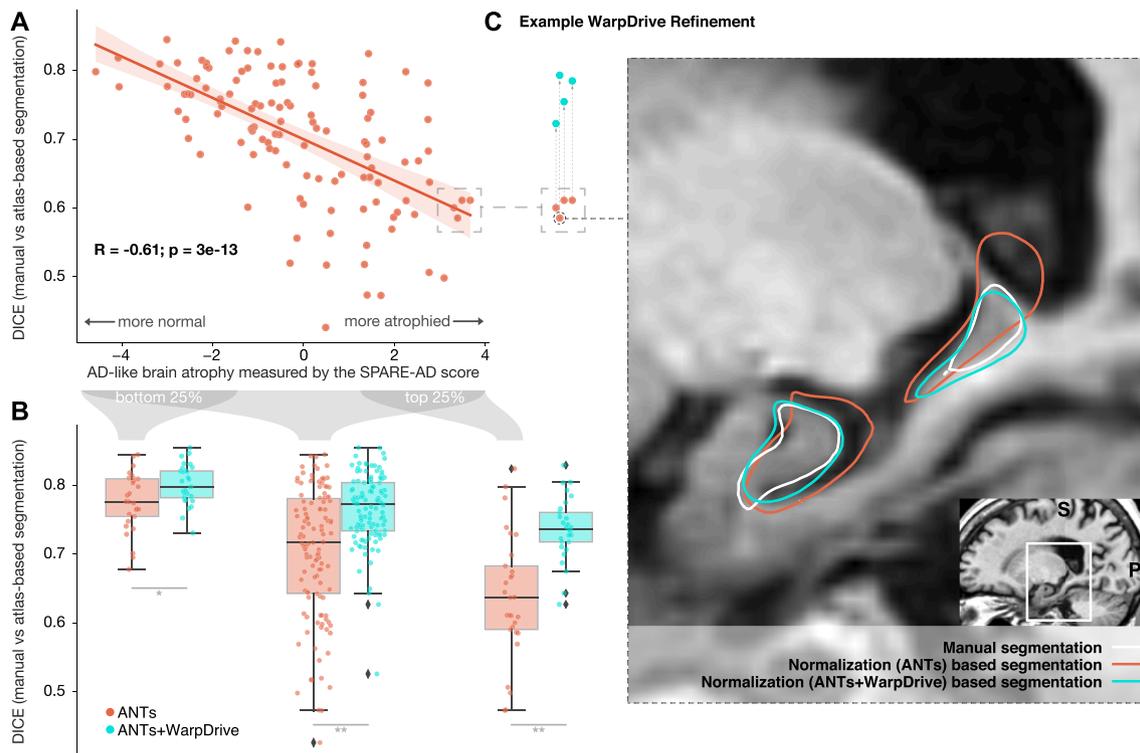
## 3. Results

### 3.1. WarpDrive toolbox

One of the main outputs of this study is the WarpDrive toolbox which we developed to manually correct mismatches after image normalization. The toolbox is made open-source and accessible through a 3D Slicer (<https://www.slicer.org/>; Fedorov et al., 2012; Kikinis et al., 2014) extension as well as within the Lead-DBS toolbox. It is composed of a set of tools that allow the user to interact with the displaced image and manually improve its alignment to template space (Fig. 1; additional details in supplementary material).

### 3.2. Automatic segmentation of atrophied brains can be improved by WarpDrive

First, we wanted to assess the necessity of applying WarpDrive, especially to brains with substantial atrophy. To do so, we retrieved a subset of the ADNI dataset in which SPARE-AD scores (which quantify AD-like brain atrophy) as well as manual segmentations of the hippocampus were available. Manual segmentations were then compared to automated (atlas-based) segmentations using the DICE score. Resulting DICE coefficients significantly correlated with the degree of atrophy represented by SPARE-AD scores (Pearson  $R = -0.61$ ;  $p = 3e - 13$ ; Fig. 2A). These data present an association between brain atrophy and the accuracy of automatic segmentations: with increasing atrophy, the



**Fig. 2.** Using WarpDrive to refine the hippocampus in brains with atrophy patterns. A Relationship between the similarities between manual versus atlas-based segmentations and the SPARE-AD value (measure of AD-like brain atrophy). A significant relation between the two is seen: as brain atrophy increases, the accuracy of the atlas-based segmentation worsens (Pearson  $R = -0.61$ ;  $p = 3e - 13$ ). Using WarpDrive to refine the normalization leads to a higher accuracy of the atlas-based segmentation (paired  $t$ -test  $T = 11.3$ ,  $p = 1e - 20$ ), with a greater improvement seen in the 25 % more atrophied brains (paired  $t$ -test  $T = 7.5$ ,  $p = 4e - 8$ ) with respect to the 25 % more normal brains (paired  $t$ -test  $T = 5.7$ ,  $p = 4e - 6$ ) (B). An example is shown in C, where the refined version matches better the manual segmentation (the user was blinded to the manual segmentations while performing the refinements).  $*p < 0.05$ ,  $**p < 1e - 6$  for paired  $t$ -tests.

accuracy declines. Especially the cases with high atrophy are potentially good candidates for WarpDrive to assist after the automatic registration. Indeed, after refining the ADNI cohort we saw a significant improvement of the DICE scores across the entire dataset (paired  $t$ -test  $T = 11.3$ ,  $p = 1e - 20$ ), with a more pronounced improvement when only taking the 25 % more atrophied brains (paired  $t$ -test  $T = 7.5$ ,  $p = 4e - 8$ ) as compared to the 25 % more normal brains (paired  $t$ -test  $T = 5.7$ ,  $p = 4e - 6$ ) (Fig. 2B). Critically, here, the user was blinded to the manual segmentation while performing the refinements.

### 3.3. WarpDrive refinements optimize registration accuracy in DBS patients

In the AD-DBS cohort published by Ríos et al., refined images were more similar to the template than without applying refinements, as measured by the mutual information metric (paired  $t$ -test  $T = 4.7$ ;  $p = 2e - 5$ ; Fig. 3A). This translated to a more accurate representation of patient data in MNI space and vice versa (e.g., transforming electrode placements to template space, or transforming atlas structures to native space, as in the highlighted example in Fig. 3).

### 3.4. Effects on electrode contact placement

The refined version of the displacement fields applied to active deep brain stimulation electrode locations translated to more clustered

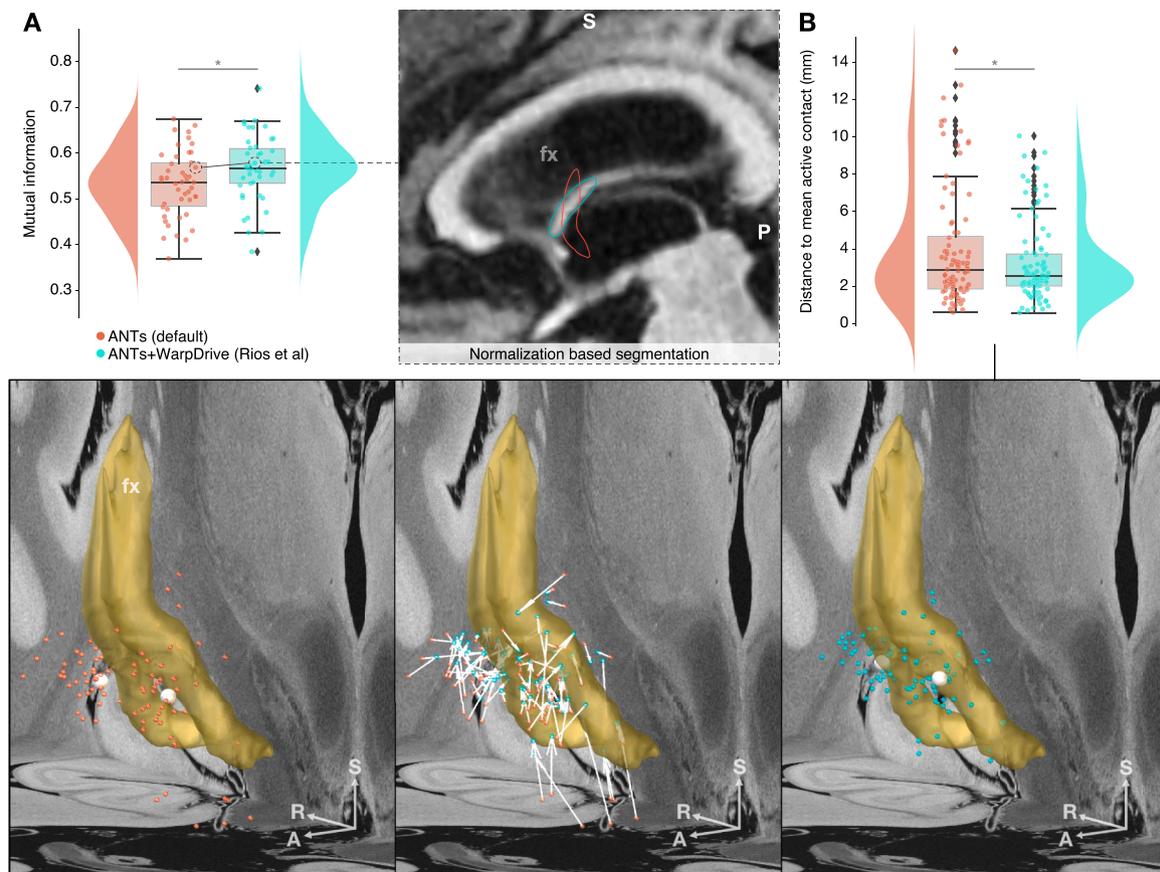
electrode positions in MNI space, as measured by comparing the distances between each active contact to their average location (Fig. 4). Better registration leading to higher clustering is expected, since surgeons attempt to target the same brain regions in each patient. Note the differences in contact position of the ventrally misregistered contacts and their relationship to the fornix following WarpDrive refinement.

### 3.5. WarpDrive as a key component in deep brain stimulation modeling

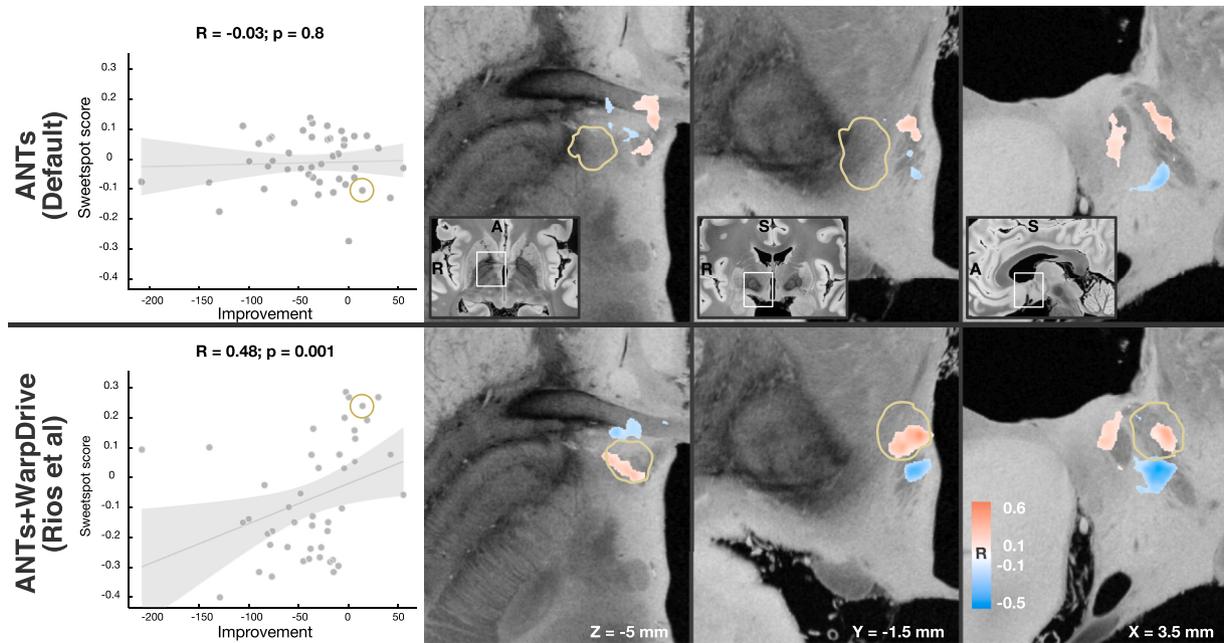
So far, we could demonstrate that WarpDrive helped create more accurate mappings between patient and template spaces. When aggregating these data and continuing ahead the processing pipeline for sweetspot analysis, WarpDrive proved to be a key step in the workflow. Fig. 4 shows the results reported by Ríos et al., which used WarpDrive refinement. To evaluate differences in sweetspot location with and without refinement, we repeated the same analysis using automated registration (ANTs) alone (omitting WarpDrive refinement). The Ríos et al. analysis shows a more specific sweetspot wherein stimulation is associated with improvement in ADAS-cog 11 score (Spearman  $R = 0.48$ ;  $p = 0.001$ ). This relationship does not hold true for the automated analysis using ANTs alone (Spearman  $R = -0.03$ ;  $p = 0.8$ ).

### 3.6. Modern registration techniques don't outperform WarpDrive

After re-running the AD-DBS analysis with EasyReg, we found that it



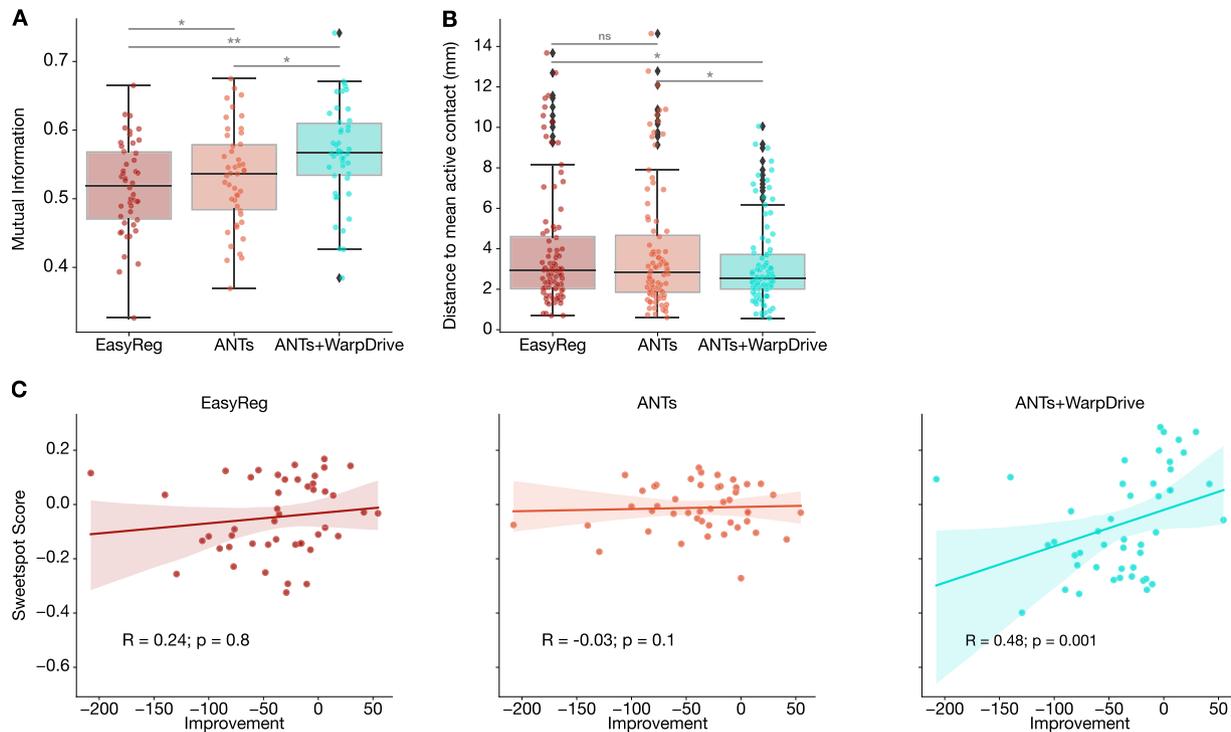
**Fig. 3.** Surrogate measures of registration improvement after using WarpDrive. Panel A shows the impact of WarpDrive in the mutual information metric: using WarpDrive, the registration accuracy is further optimized, reaching a higher mutual information value (paired  $t$ -test  $T = 4.7$ ;  $p = 2e - 5$ ). The refined transform also impacts how atlas structures map to patient space, as shown by an exemplar case in the top center. Panel B shows the impact of WarpDrive in the dispersion of warped active DBS electrode contacts placement: after WarpDrive refinement, the contacts are more clustered together (measured by the distance from the contacts to the mean contact coordinate; paired  $t$ -test  $T = 4.0$ ;  $p = 1e - 4$ ). The bottom row illustrates this by showing the active contacts in MNI space after transforming them using automated registration (ANTs; left); after additional refinement (ANTs+WarpDrive; right); and their respective displacement (center). The mean contact coordinates are represented as white spheres in the figure. fx: Fornix, from Neudorfer et al., 2020; Background brain slices defined by the postmortem 100  $\mu$ m template (Edlow et al., 2019). \* $p < 0.05$  for paired  $t$ -tests.



**Fig. 4.** Sweetspot analysis comparison with and without WarpDrive. Ríos et al. applied WarpDrive to refine the normalization of the images to template space and carried out a sweetspot analysis defining a region associated with better outcome after surgery (Spearman  $R = 0.48$ ;  $p = 0.001$ ), shown in the bottom panel. Repeating the same analysis when omitting the WarpDrive refinement step does not lead to significant associations in the analysis (Spearman  $R = -0.03$ ;  $p = 0.8$ ). The stimulation volume of an example responder patient is shown in the slices (and marked in scatter plots) highlighted by yellow outlines. In the automated analysis, their stimulation volume does not intersect with the sweetspot region, after application of WarpDrive, it does. The sweetspot is represented with red colors for association with beneficial outcome and blue colors for the opposite relationship. Background brain slices are defined by the postmortem 100  $\mu\text{m}$  template (Edlow et al., 2019).

had the lowest distribution of mutual information values (paired  $t$ -test vs ANTs:  $T = 3.8$ ,  $p = 4e - 4$ ; and vs ANTs+WarpDrive:  $T = 5.8$ ,  $p = 6e - 7$ ; Fig. 5A); and greater dispersion of contact location as compared to

ANTs+WarpDrive (paired  $t$ -test  $T = 4.2$ ,  $p = 6e - 5$ ), while insignificant difference with respect to ANTs (paired  $t$ -test  $T = 0.3$ ,  $p = 0.7$ ; Fig. 5B). As for the sweetspot analysis, the cohort processed with



**Fig. 5.** Additional comparison with novel deep learning based normalization approach. After re-doing the AD-DBS analysis with EasyReg (Iglesias, 2023), this method showed: (A) lowest mutual information metric score between the normalized and template images; (B) greater dispersion of active contact placement compared to ANTs+WarpDrive, while insignificant difference with ANTs; and (C) a positive, while not significant, association between sweetspot score and clinical improvement (Pearson  $R = 0.24$ ,  $p = 0.8$ ).  $*p < 0.05$ ,  $**p < 1e - 6$  for paired  $t$ -tests.

EasyReg showed an insignificant positive association between scores and clinical improvement (Pearson  $R = 0.24$ ,  $p = 0.8$ ; Fig. 5C).

### 3.7. Qualitative analysis results

WarpDrive can assist in the process of aligning high resolution templates to the same reference template used during group spatial normalization. A precise alignment is crucial to create visualizations of regions of interest that provide refined and more detailed anatomical insight, going beyond what can be seen on typical MNI templates. We show an example in Fig. 6, where we refine the 100  $\mu\text{m}$  postmortem template (Edlow et al., 2019) alignment to MNI space.

WarpDrive can also be used to refine registrations of volumes to histological slices: we demonstrate this application with a 1  $\mu\text{m}$  slice of the Bigbrain dataset (Amunts et al., 2013) to which a common MNI template had been registered before (Xiao et al., 2019). We manually refined this registration using WarpDrive. Particularly, the increasing resolution made some discrepancies in the alignment more evident, which we were able to adjust (supplementary Fig. 1). Furthermore, WarpDrive was also able to assist in the atlas creation process, where focusing the refinements on a target region of interest translated to clearer and crisper borders of the subthalamic nucleus (supplementary Fig. 2). Finally, we show the usability of WarpDrive in the context of abdomen registration, improving the alignment of intra-patient multi-modal images focusing on the liver (supplementary Fig. 3).

## 4. Discussion

We propose a method for manual refinement of deformation fields based on user interactions with images. The tool provides visual feedback about the alignment of images in conjunction with atlases and parcellations in MNI space. This is a step forward in the field of spatial normalization, which has been around for over two decades and has achieved excellent results in many contexts, but sometimes, refinements are still critical to achieve an accurate correspondence between images. Without such a tool, when automatic algorithms fail (or produce poor fits in specific regions), the user is often left without choices to address the issue. WarpDrive closes this gap, i.e., constitutes a tool that allows interaction with normalization results following application of automated registration routines.

It should be noted that since WarpDrive starts from an initial transform, this first alignment impacts the usability of the tool: if the discrepancies are too big, WarpDrive is likely not the best tool to manually correct them; on the other end, if the alignment is already perfect, the application of WarpDrive is unnecessary. In other words, WarpDrive was

designed for typical registration results we often see in neuroimaging, which are overall of decent fit but may require small adjustments here or there.

We studied an AD-DBS cohort, which is characterized by significant levels of atrophy and regions of interest are particularly small. When analyzing such datasets, the key question is to resolve the small differences in electrode placements with high accuracy, in order to contrast stimulation sites of responders with the ones of non-responders (Treu et al., 2020). However, since all electrodes are targeting the same brain region, and given the levels of atrophy, these analyses range among the most challenging applications of spatial normalizations. While the original paper by Ríos et al. had used WarpDrive already, the report did not quantify its impact on results. The present manuscript addressed this limitation by comparing results between analyses carried out *with* and *without* the additional WarpDrive step. This showed a strong impact of the WarpDrive method on the result reported by Ríos et al., demonstrating how critical this processing step had been.

Moreover, we provide evidence for the utility of WarpDrive in correcting registrations of brains with substantial morphological changes, specifically atrophy in the context of neurodegeneration. We show that spatial normalization (after automated processing without additional refinement) negatively correlates with the degree of atrophy and that applying WarpDrive to atrophied brains can reduce this bias.

Following this direction, we consider WarpDrive a tool to assist in creating accurate spatial representations of data. Not only can patient imaging be better aligned to an average reference brain. Just as much, other high-resolution resources can be precisely registered to template (or patient) spaces (Al-Fatly et al., 2023). Our examples of refinements in the 100  $\mu\text{m}$  postmortem template (Edlow et al., 2019) and 1  $\mu\text{m}$  BigBrain slice (Amunts et al., 2013) showcase WarpDrive's utility for such cases.

While only qualitatively presented here, these concepts could be further explored in the growing field of histology to MRI registration, bridging micro to macro scales (Adler et al., 2014; Amunts et al., 2013; Iglesias et al., 2018). Another example of application we foresee is in fMRI studies that focus on specific neuroanatomical subregions of small scale (such as activations within hippocampal subfields or layer-specific activation studies). Here, WarpDrive could help refine registrations of regions of interest to potentially enhance the power of statistical analyses on the group level, or to better compare anatomical details of brain activations. WarpDrive may be readily applied to non-human or inanimate imaging applications, and although these concepts are not studied here, we foresee great potential for WarpDrive in these applications as well.

Finally, as the applications of WarpDrive grow, we also envision

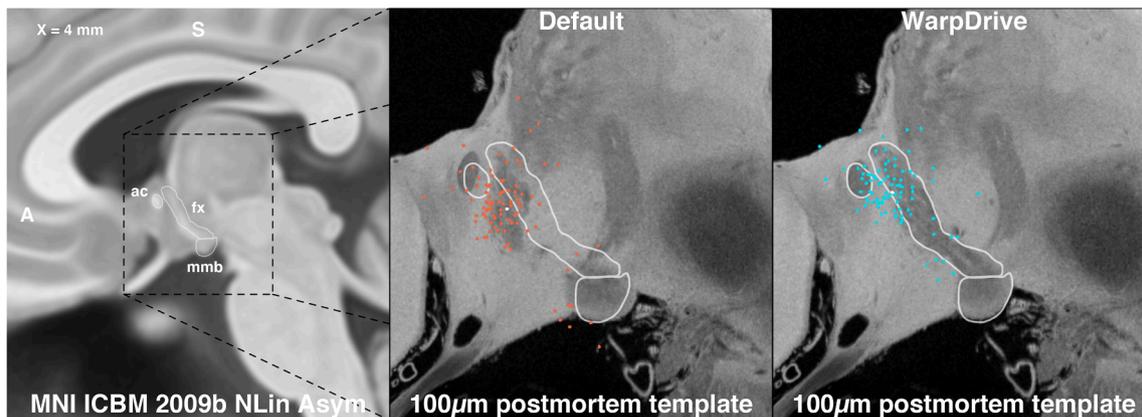


Fig. 6. WarpDrive can assist in creating accurate representations of data in template space. In this figure we show the default (center) and WarpDrive-refined (right) version of the 100  $\mu\text{m}$  postmortem template (Edlow et al., 2019). We also include the active contacts and their mean (white dot) mapped to a sagittal slice (similarly without and with application of WarpDrive). This way, multiple data can be precisely transformed to a template space and create accurate representations of them. Atlas outlines, from Neudorfer et al., 2020: ac: Anterior commissure; fx: Fornix; mmb: Mammillary bodies.

improvements in the tool itself. While in this work we proposed different ways of interactions to apply corrections to a given displacement field, further interaction modes or enhancements to the present ones could make WarpDrive more versatile and user-friendly. For example, a tool where the user could radially shrink or expand the warpfield locally could be useful for some applications. Furthermore, drawing with assistance could be convenient as well, where the outline drawn automatically follows contours, as it does in the magnetic selection tool in common photo editing software. Finally, it would be interesting to investigate how different registration metrics perform when using the Snap mode—and perhaps at some point delegate this operation to an active learning model, which could learn from the user’s corrections and feedback.

#### 4.1. Limitations

While WarpDrive is a potentially powerful method, it introduces observer bias to a process that had been observer-independent before. This might be counterintuitive when the goal is to ensure an automatic processing pipeline for image analyses. But the automation should be balanced with the accuracy of the achieved results. Furthermore, while not explored here, it is possible to apply WarpDrive as an intermediate step. Within Lead-DBS, it is implemented in such a way that automatic registrations can first be refined using WarpDrive. Then, the automatic algorithm (ANTs) can be run again, starting off from the refined solution. Technically, this process could even be iteratively repeated as often as desired (i.e., handing off the registration from algorithm to human and back), albeit it remains unclear, whether this would lead to superior results. The introduced *Snap* feature is also a semi-automatized routine which could objectively adjust the corrections.

Second, WarpDrive alone will not necessarily improve the correspondence of images. Since it is applied by a user, the user requires anatomical knowledge and methodological insights into neuroimaging. In other words, it is possible to make results worse using WarpDrive. The quality of the images is also an essential factor of impact on the quality of WarpDrive refined results. Obviously, it is important to see and distinguish the structures of interest to be able to refine their mapping.

Third, the process can introduce so-called “high frequency deformations”, i.e. “swirling” or strong distortions. These could be corrected for by smoothing the warpfield and/or by re-applying the registration routine taking the refined transform as starting point. However, further work is needed to clarify limitations of these concepts, depending on the field of application (different for fMRI, histology, DBS, etc.). Hence, we must emphasize that, while potentially powerful, WarpDrive is a manual method that should be applied by expert neuroimagers on high quality data and with the necessary anatomical knowledge.

Finally, WarpDrive is not suited to refine whole-brain registrations. While technically possible (one could iteratively refine the entire brain), it would be a very labor-intensive process. Rather, the key application we see is to refine certain areas of interest. DBS imaging, layer-specific fMRI or registrations of smaller tissue blocks are clear applications in which WarpDrive could be useful. Instead, using the tool to refine registrations that are significantly off at various locations throughout the brain is not advised.

## 5. Conclusion

We introduce WarpDrive, a method to refine pair-wise nonlinear registrations between images. While in this study we focus mainly on human brain MRI images, the tool is not tied to a specific modality or species and can be applied to any type of 3D imaging data. It can be particularly useful to refine region-specific registrations in the domains of ultra-high-field imaging, cross-modality registrations, and high-fidelity neuroimaging, such as deep brain stimulation. Indeed, we show how WarpDrive can help create accurate anatomical

representations and models of multimodal imaging data in an AD-DBS cohort. The tool is openly available in form of a 3D Slicer extension and within the Lead-DBS software.

#### Data availability

Raw data from the ADNI cohort can be accessed from the database (<https://adni.loni.usc.edu/>). Raw data from the Alzheimer’s disease deep brain stimulation cohort cannot be shared to ensure patient data protection. Anonymized derivatives together with code to recreate the result figures can be accessed here: [https://github.com/simoxen/WarpDrive\\_Supplementary](https://github.com/simoxen/WarpDrive_Supplementary).

Data were processed using the openly available WarpDrive (<https://github.com/netstim/SlicerNetstim>) and Lead-DBS (<https://github.com/netstim/leaddbs>) software.

#### Ethics statement

Post hoc data analysis carried out in the present study was approved by the ethics board of Charité—Universitätsmedizin Berlin (master vote EA2/186/18).

#### Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

A.H. reports a relationship with Boston Scientific Corporation that includes: speaking and lecture fees. L.A. reports a relationship with Boston Scientific Corporation that includes: consulting or advisory. L.A. reports a relationship with Medtronic that includes: consulting or advisory. K.F. reports a relationship with Medtronic that includes: consulting or advisory. K.F. reports a relationship with Boston Scientific Corporation that includes: consulting or advisory. K.F. reports a relationship with Abbott that includes: funding grants. K.F. reports a relationship with Functional Neuromodulation that includes: funding grants. D.W. reports a relationship with Functional Neuromodulation that includes: funding grants. D.W. reports a relationship with Avid Lily that includes: funding grants. D.W. reports a relationship with Merck that includes: funding grants. D.W. reports a relationship with Janssen that includes: consulting or advisory. D.W. reports a relationship with GE Healthcare that includes: consulting or advisory. D.W. reports a relationship with Biogen that includes: consulting or advisory. D.W. reports a relationship with Neuronix that includes: consulting or advisory. S.S. reports a relationship with Lilly that includes: consulting or advisory. S.S. reports a relationship with Roche that includes: consulting or advisory. S.S. reports a relationship with Novartis that includes: consulting or advisory. S.S. reports a relationship with Biogen that includes: consulting or advisory. G.S. reports a relationship with NIH that includes: funding grants. M.S. reports a relationship with Allergan that includes: consulting or advisory. M.S. reports a relationship with Biogen that includes: consulting or advisory. M.S. reports a relationship with Roche that includes: consulting or advisory. M.S. reports a relationship with Cortexyme that includes: consulting or advisory. M.S. reports a relationship with Bracket that includes: consulting or advisory. M.S. reports a relationship with Brain Health Inc that includes: consulting or advisory. M.S. reports a relationship with uMethod Health that includes: consulting or advisory. C.L. reports a relationship with Functional Neuromodulation that includes: funding grants. C.L. reports a relationship with Avanir that includes: funding grants. C.L. reports a relationship with Eli Lilly that includes: funding grants. C.L. reports a relationship with NFL Benefits Office that includes: funding grants. M.O. reports a relationship with NIH that includes: funding

grants. M.O. reports a relationship with Tourette Association of America Grant that includes: funding grants. M.O. reports a relationship with Parkinson's Alliance that includes: funding grants. M.O. reports a relationship with Smallwood Foundation that includes: funding grants. M.O. reports a relationship with Parkinson's Foundation Medical Director that includes: consulting or advisory. M.O. reports a relationship with Books4Patients that includes: consulting or advisory. M.O. reports a relationship with American Academy of Neurology that includes: consulting or advisory. M.O. reports a relationship with Peerview that includes: consulting or advisory. M.O. reports a relationship with WebMD Medscape that includes: consulting or advisory. M.O. reports a relationship with Mededicus that includes: consulting or advisory. M.O. reports a relationship with Movement Disorders Society that includes: consulting or advisory. M.O. reports a relationship with Taylor and Francis that includes: consulting or advisory. M.O. reports a relationship with Demos that includes: consulting or advisory. M.O. reports a relationship with Robert Rose that includes: consulting or advisory. M.O. reports a relationship with Medtronic that includes: non-financial support. A.L. reports a relationship with Medtronic that includes: funding grants. A.L. reports a relationship with Functional Neuromodulation that includes: funding grants. A.L. reports a relationship with Medtronic that includes: consulting or advisory. A.L. reports a relationship with St. Jude that includes: consulting or advisory. A.L. reports a relationship with Boston Scientific that includes: consulting or advisory. A.L. has patent #US Patent 8346,365 licensed to Functional Neuromodulation.

#### Data availability

Anonymized derivatives together with code to recreate the result figures can be accessed here: [https://github.com/simonoxen/WarpDrive\\_Supplementary](https://github.com/simonoxen/WarpDrive_Supplementary).

#### Acknowledgements

Part of this study was presented and worked on during the 35th and 38th NA-MIC Project Weeks (Kapur et al., 2016). We would like to thank the NA-MIC community, specially to Dr. Andras Lasso for the help and discussions on 3D Slicer modules implementation.

A.H. was supported by the German Research Foundation (Deutsche Forschungsgemeinschaft, 424778381 – TRR 295), Deutsches Zentrum für Luft- und Raumfahrt (DynaSti grant within the EU Joint Programme Neurodegenerative Disease Research, JPND), the National Institutes of Health (R01 13478451, 1R01NS127892-01, 2R01 MH113929 & UM1NS132358) as well as the New Venture Fund (FFOR Seed Grant). B. H. was supported by a scholarship from the Einstein Center for Neurosciences Berlin.

Data collection and sharing for this project was funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI) (National Institutes of Health Grant U01 AG024904) and DOD ADNI (Department of Defense award number W81XWH-12-2-0012). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: AbbVie, Alzheimer's Association; Alzheimer's Drug Discovery Foundation; Araclon Biotech; BioClinica, Inc.; Biogen; Bristol-Myers Squibb Company; CereSpir, Inc.; Cogstate; Eisai Inc.; Elan Pharmaceuticals, Inc.; Eli Lilly and Company; EuroImmun; F. Hoffmann-La Roche Ltd and its affiliated company Genentech, Inc.; Fujirebio; GE Healthcare; IXICO Ltd.; Janssen Alzheimer Immunotherapy Research & Development, LLC.; Johnson & Johnson Pharmaceutical Research & Development LLC.; Lumosity; Lundbeck; Merck & Co., Inc.; Meso Scale Diagnostics, LLC.; NeuroRx Research; Neurotrack Technologies; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Piramal Imaging; Servier; Takeda Pharmaceutical Company; and Transition Therapeutics. The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the

Foundation for the National Institutes of Health ([www.fnih.org](http://www.fnih.org)). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Therapeutic Research Institute at the University of Southern California. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California.

#### Supplementary materials

Supplementary material associated with this article can be found, in the online version, at [doi:10.1016/j.media.2023.103041](https://doi.org/10.1016/j.media.2023.103041).

#### References

- Adler, D.H., Pluta, J., Kadivar, S., Craige, C., Gee, J.C., Avants, B.B., & Yushkevich, P.A. (2014). Histology-derived volumetric annotation of the human hippocampal subfields in postmortem MRI. *84*, 505–523. [10.1016/j.neuroimage.2013.08.067](https://doi.org/10.1016/j.neuroimage.2013.08.067).
- Al-Fatly, B., Giesler, S.J., Oxenford, S., Li, N., Dembek, T.A., Achtzehn, J., Krause, P., Visser-Vandewalle, V., Krauss, J.K., Runge, J., Tadic, V., Bäumer, T., Schnitzler, A., Vesper, J., Wirths, J., Timmermann, L., Kühn, A.A., Koy, A., 2023. Neuroimaging-based analysis of DBS outcomes in pediatric dystonia: Insights from the GEPSTIM registry. *NeuroImage. Clin.* 39, 103449. <https://doi.org/10.1016/j.nicl.2023.103449>.
- Alkemade, A., Mulder, M.J., Groot, J.M., Isaacs, B.R., van Berendonk, N., Lute, N., Isherwood, S.J., Bazin, P.L., Forstmann, B.U., 2020. The Amsterdam ultra-high field adult lifespan database (AHEAD): a freely available multimodal 7 Tesla submillimeter magnetic resonance imaging database. *NeuroImage* 221, 117200. <https://doi.org/10.1016/j.neuroimage.2020.117200>.
- Amunts, K., Lepage, C., Borgeat, L., Mohlberg, H., Dickscheid, T., Rousseau, M.E., Bludau, S., Bazin, P.L., Lewis, L.B., Oros-Peusquens, A.M., Shah, N.J., Lippert, T., Zilles, K., Evans, A.C., 2013. BigBrain: an ultrahigh-resolution 3D human brain model. *Science* 340 (6139), 1472–1475. <https://doi.org/10.1126/science.1235381>.
- Avants, B., Epstein, C., Grossman, M., Gee, J., 2008. Symmetric diffeomorphic image registration with cross-correlation: evaluating automated labeling of elderly and neurodegenerative brain. *Med. Image Anal.* 12 (1), 26–41. <https://doi.org/10.1016/j.media.2007.06.004>.
- Boccardi, M., Bocchetta, M., Apostolova, L.G., Barnes, J., Bartzokis, G., Corbetta, G., DeCarli, C., deToledo-Morrell, L., Firbank, M., Ganzola, R., Gerritsen, L., Henneman, W., Killiany, R.J., Malykhin, N., Pasqualetti, P., Pruessner, J.C., Redolfi, A., Robitaille, N., Soinen, H., 2015a. Delphi definition of the EADC-ADNI Harmonized Protocol for hippocampal segmentation on magnetic resonance. *Alzheimer's Dement.* 11 (2), 126–138. <https://doi.org/10.1016/j.jalz.2014.02.009>.
- Boccardi, M., Bocchetta, M., Morency, F.C., Collins, D.L., Nishikawa, M., Ganzola, R., Grothe, M.J., Wolf, D., Redolfi, A., Pievani, M., Antelmi, L., Fellgiebel, A., Matsuda, H., Teipel, S., Duchesne, S., Jack Jr., C.R., Frisoni, G.B., 2015b. Training labels for hippocampal segmentation based on the EADC-ADNI harmonized hippocampal protocol. *Alzheimer's Dement.* 11 (2), 175–183. <https://doi.org/10.1016/j.jalz.2014.12.002>.
- Clark, K., Vendt, B., Smith, K., Freymann, J., Kirby, J., Koppel, P., Moore, S., Phillips, S., Maffitt, D., Pringle, M., Tarbox, L., Prior, F., 2013. The cancer imaging archive (TCIA): maintaining and operating a public information repository. *J. Digit. Imaging* 26 (6), 1045–1057. <https://doi.org/10.1007/s10278-013-9622-7>.
- Davatzikos, C., Xu, F., An, Y., Fan, Y., Resnick, S.M., 2009. Longitudinal progression of Alzheimer's-like patterns of atrophy in normal older adults: the SPARE-AD index. *Brain* 132 (8), 2026–2035. <https://doi.org/10.1093/brain/awp091>.
- Desikan, R.S., Ségonne, F., Fischl, B., Quinn, B.T., Dickerson, B.C., Blacker, D., Buckner, R.L., Dale, A.M., Maguire, R.P., Hyman, B.T., Albert, M.S., Killiany, R.J., 2006. An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *NeuroImage* 31 (3), 968–980. <https://doi.org/10.1016/j.neuroimage.2006.01.021>.
- Diaz-Pinto, A., Mehta, P., Alle, S., Asad, M., Brown, R., Nath, V., Ihsani, A., Antonelli, M., Palkovics, D., Pinter, C., Alkalay, R., Pieper, S., Roth, H.R., Xu, D., Dogra, P., Vercauteren, T., Feng, A., Quraini, A., Ourselin, S., Cardoso, M.J., 2022. DeepEdit: deep editable learning for interactive segmentation of 3D medical images. H. V. Nguyen, S. X. Huang, & Y. Xue Data Augmentation, Labelling, and Imperfections. Springer Nature Switzerland, pp. 11–21. [https://doi.org/10.1007/978-3-031-17027-0\\_2](https://doi.org/10.1007/978-3-031-17027-0_2).
- Dice, L.R., 1945. Measures of the amount of ecologic association between species. *Ecology* 26 (3), 297–302. <https://doi.org/10.2307/1932409>.
- Edlow, B.L., Mareyam, A., Horn, A., Polimeni, J.R., Witzel, T., Tisdall, M.D., Augustinack, J.C., Stockmann, J.P., Diamond, B.R., Stevens, A., Tirrell, L.S., Folkert, R.D., Wald, L.L., Fischl, B., Van Der Kouwe, A., 2019. 7 Tesla MRI of the ex vivo human brain at 100 μm resolution. *Sci. Data* 6 (1). <https://doi.org/10.1038/s41597-019-0254-8>.
- Ewert, S., Horn, A., Finkel, F., Li, N., Kühn, A.A., Herrington, T.M., 2019. Optimization and comparative evaluation of nonlinear deformation algorithms for atlas-based segmentation of DBS target nuclei. *NeuroImage* 184, 586–598. <https://doi.org/10.1016/j.neuroimage.2018.09.061>.
- Ewert, S., Plettig, P., Li, N., Chakravarty, M., Collins, L., Herrington, T., Kühn, A., Horn, A., 2017. Toward defining deep brain stimulation targets in MNI space: A subcortical atlas based on multimodal MRI, histology and structural connectivity. *NeuroImage*. <https://doi.org/10.1016/j.neuroimage.2017.05.015>.

- Fedorov, A., Beichel, R., Kalpathy-Cramer, J., Finet, J., Fillion-Robin, J.C., Pujol, S., Bauer, C., Jennings, D., Fennessy, F., Sonka, M., Buatti, J., Aylward, S., Miller, J.V., Pieper, S., Kikinis, R., 2012. 3D slicer as an image computing platform for the quantitative imaging network. *Magn. Reson. Imaging* 30 (9), 1323–1341. <https://doi.org/10.1016/j.mri.2012.05.001>.
- Fonov, V.S., Evans, A.C., McKinstry, R.C., Almlí, C.R., Collins, D.L., 2009. Unbiased nonlinear average age-appropriate brain templates from birth to adulthood. *NeuroImage* 47, S102. [https://doi.org/10.1016/S1053-8119\(09\)70884-5](https://doi.org/10.1016/S1053-8119(09)70884-5).
- Godley, A., Ahunbay, E., Peng, C., Li, X.A., 2009. Automated registration of large deformations for adaptive radiation therapy of prostate cancer. *Med. Phys.* 36 (4), 1433–1441. <https://doi.org/10.1118/1.3095777>.
- Hering, A., Hansen, L., Mok, T.C.W., Chung, A.C.S., Siebert, H., Häger, S., Lange, A., Kuckertz, S., Heldmann, S., Shao, W., Vesal, S., Rusu, M., Sonn, G., Estienne, T., Vakalopoulou, M., Han, L., Huang, Y., Yap, P.T., Brudfors, M., Heinrich, M.P., 2023. Learn2Reg: comprehensive multi-task medical image registration challenge, dataset and evaluation in the era of deep learning. *IEEE Trans. Med. Imaging* 42 (3), 697–712. <https://doi.org/10.1109/TMI.2022.3213983>.
- Hermosillo, G., Chef'd'Hotel, C., Faugeras, O., 2002. Variational methods for multimodal image matching. *Int. J. Comput. Vis.* 50 (3), 329–343. <https://doi.org/10.1023/A:1020830525823>.
- Huttenlocher, D.P., Klanderman, G.A., Rucklidge, W.J., 1993. Comparing images using the Hausdorff distance. *IEEE Trans. Pattern Anal. Mach. Intell.* 15 (9), 850–863. <https://doi.org/10.1109/34.232073>.
- Iglesias, J.E., 2023. A ready-to-use machine learning tool for symmetric multi-modality registration of brain MRI. *Sci. Rep.* 13 (1) <https://doi.org/10.1038/s41598-023-33781-0>. Article 1.
- Iglesias, J.E., Insausti, R., Lerma-Usabiaga, G., Bocchetta, M., Van Leemput, K., Greve, D. N., van der Kouwe, A., Fischl, B., Caballero-Gaudes, C., Paz-Alonso, P.M., 2018. A probabilistic atlas of the human thalamic nuclei combining ex vivo MRI and histology. *Neuroimage* 183, 314–326. <https://doi.org/10.1016/j.neuroimage.2018.08.012>.
- Kapur, T., Pieper, S., Fedorov, A., Fillion-Robin, J.C., Halle, M., O'Donnell, L., Lasso, A., Ungi, T., Pinter, C., Finet, J., Pujol, S., Jagadeesan, J., Tokuda, J., Norton, I., Estepar, R.S.J., Gering, D., Aerts, H.J.W.L., Jakab, M., Hata, N., Kikinis, R., 2016. Increasing the impact of medical image computing using community-based open-access hackathons: the NA-MIC and 3D slicer experience. *Med. Image Anal.* 33, 176–180. <https://doi.org/10.1016/j.media.2016.06.035>.
- Kikinis, R., Pieper, S.D., Vosburgh, K.G., 2014. 3D Slicer: a platform for subject-specific image analysis, visualization, and clinical support. *Intraoperative Imaging and Image-Guided Therapy*. Springer, pp. 277–289.
- Nath, V., Yang, D., Landman, B.A., Xu, D., Roth, H.R., 2021. Diminishing uncertainty within the training pool: active learning for medical image segmentation. *IEEE Trans. Med. Imaging* 40 (10), 2534–2547. <https://doi.org/10.1109/TMI.2020.3048055>.
- Neudorfer, C., Butenko, K., Oxenford, S., Rajamani, N., Achtzehn, J., Goede, L., Hollunder, B., Ríos, A.S., Hart, L., Tasserie, J., Fernando, K.B., Nguyen, T.A.K., Al-Fatly, B., Vissani, M., Fox, M., Richardson, R.M., van Rienen, U., Kühn, A.A., Husch, A.D., Horn, A., 2023. Lead-DBS v3.0: mapping deep brain stimulation effects to local anatomy and global networks. *Neuroimage*, 119862. <https://doi.org/10.1016/j.neuroimage.2023.119862>.
- Neudorfer, C., Germann, J., Elias, G.J.B., Gramer, R., Boutet, A., Lozano, A.M., 2020. A high-resolution *in vivo* magnetic resonance imaging atlas of the human hypothalamic region. *Sci. Data* 7 (1), 305. <https://doi.org/10.1038/s41597-020-00644-6>.
- Nichols, T.E., Holmes, A.P., 2002. Nonparametric permutation tests for functional neuroimaging: a primer with examples. *Hum. Brain Mapp.* 15 (1), 1–25. <https://doi.org/10.1002/hbm.1058>.
- Ríos, A.S., Oxenford, S., Neudorfer, C., Butenko, K., Li, N., Rajamani, N., Boutet, A., Elias, G.J.B., Germann, J., Loh, A., Deeb, W., Wang, F., Setsompop, K., Salvato, B., Almeida, L.B., Foote, K.D., Amaral, R., Rosenberg, P.B., Tang-Wai, D.F., Horn, A., 2022. Optimal deep brain stimulation sites and networks for stimulation of the fornix in Alzheimer's disease. *Nat. Commun.* 13 (1) <https://doi.org/10.1038/s41467-022-34510-3>. Article 1.
- Sharp, G.C., Li, R., Wolfgang, J., Chen, G., Peroni, M., Spadea, M.F., Mori, S., Zhang, J., Shackelford, J., Kandasamy, N., 2010. Plastimatch—an open source software suite for radiotherapy image processing. In: *Proceedings of the 16th International Conference on the use of Computers in Radiotherapy (ICCR)*. Amsterdam, Netherlands.
- Shusharina, N., Sharp, G., 2012. Analytic regularization for landmark-based image registration. *Phys. Med. Biol.* 57 (6), 1477. <https://doi.org/10.1088/0031-9155/57/6/1477>.
- Sumanaweera, T.S., Adler, J.R.J., Napel, S., Glover, G.H., 1994. Characterization of Spatial Distortion in Magnetic Resonance Imaging and Its Implications for Stereotactic Surgery. *Neurosurgery* 35 (4), 696.
- Treu, S., Strange, B., Oxenford, S., Neumann, W.J., Kühn, A., Li, N., Horn, A., 2020. Deep brain stimulation: imaging on a group level. *Neuroimage* 219, 117018. <https://doi.org/10.1016/j.neuroimage.2020.117018>.
- Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., Mazoyer, B., Joliot, M., 2002. Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage* 15 (1), 273–289. <https://doi.org/10.1006/nimg.2001.0978>.
- Vezhnevets, V., & Konouchine, V. (2005). GrowCut: Interactive multi-label ND image segmentation by cellular automata. 1(4), 150–156.
- Wolf, D., Bocchetta, M., Preboske, G.M., Boccardi, M., Grothe, M.J., 2017. Reference standard space hippocampus labels according to the European Alzheimer's disease consortium—Alzheimer's disease neuroimaging initiative harmonized protocol: utility in automated volumetry. *Alzheimer's Dement.* 13 (8), 893–902. <https://doi.org/10.1016/j.jalz.2017.01.009>.
- Wu, B., 2014. Interactive Medical Image Registration With Multigrid Methods and Bounded Biharmonic Functions. University of Pennsylvania [Ph.D.]. <https://www.proquest.com/docview/1614194341/abstract/F32096CC5C9B4F4BPQ/1>.
- Xiao, Y., Fonov, V., Chakravarty, M.M., Beriault, S., Al Subaie, F., Sadikot, A., Pike, G.B., Bertrand, G., Collins, D.L., 2017. A dataset of multi-contrast population-averaged brain MRI atlases of a Parkinson's disease cohort. *Data Brief* 12, 370–379. <https://doi.org/10.1016/j.dib.2017.04.013>.
- Xiao, Y., Lau, J.C., Anderson, T., Dekraker, J., Collins, D.L., Peters, T., Khan, A.R., 2019. An accurate registration of the BigBrain dataset with the MNI PD25 and ICBM152 atlases. *Sci. Data* 6 (1). <https://doi.org/10.1038/s41597-019-0217-0>.
- Zhao, F., Xie, X., 2013. An overview of interactive medical image segmentation. *Ann. BMVA* 2013 (7), 1–22.
- Zhou, W., Xie, Y., 2013. Interactive multigrid refinement for deformable image registration. *BioMed Res. Int.*, e532936 <https://doi.org/10.1155/2013/532936>, 2013.