Supplementary material

Summary of participating methods

This section briefly summarizes the methods of various participating teams. Invited methods are indicated with *. Table S1 gives an overview of all participating methods.

Team cwmok[1]

Mok and Chung[1] proposed a 3-step registration method, which comprises an affine pre-alignment, a convolutional neural network with forward-backward consistency constraint, and a nonlinear instance optimization. First, possible linear misalignments caused by the tumour mass effect were eliminated with the descent-based affine registration method. Second, conditional deep Laplacian pyramid image registration network with forward-backward consistency (DIRAC)[2,3] was leveraged to jointly estimate the regions with missing correspondence and bidirectional nonlinear displacement fields for the pre-operative and follow-up scans. During the training phase, regions with missing correspondence were excluded from the similarity measure. This reduced the effect of the pathological regions on the registration algorithm in an unsupervised manner. Finally, non-rigid instance optimization with forward-backward consistency constraint was introduced to correct solutions from the previous step that were biased because of insufficient training and discrepancy in distributions. This step further improved the robustness and registration accuracy of initial solutions. The non-parametric deformation was controlled by the forward-backward consistency constraint as in the previous step and was updated using an Adam optimizer together with multi-level continuation to avoid local minima. The implementation of DIRAC is available at https://github.com/cwmok/DIRAC.

Team UZL[4]

UZL utilised a combination of hand-crafted features, a single-level discrete correlation layer subject to a convex optimisation scheme, and a subsequent Adam instance optimisation for fine-grained displacement prediction. First, MIND-SSC features were extracted on T1 contrast-enhanced baseline and follow-up images. The features were then used to compute a correlation-based cost tensor containing sum-of-squared-differences. Next, an iterative convex optimisation regarding feature matching and global smoothness was performed. To account for tumor-related mass effects and missing correspondences the authors employed a large search range and enforced inverse consistency. An Adam instance optimisation further refined the intermediate displacement field with diffusion regularisation and B-spline-interpolation.

Team AGHSSO[5]

The proposed method consists of: (i) preprocessing and modality aggregation, (ii) iterative affine registration, (iii) dense displacement field calculation by LapIRN, (iv) iterative, instance optimization-based nonrigid registration, and (v) displacement field fine-tuning by optimizing an objective function that was weighted based on the inverse consistency. The method addressed both the challenges related to the pre- to post- operative registration, namely the large, nonrigid deformations and the missing tissues. The large and complex deformations were addressed by the LapIRN network with large enough receptive field, while the missing tissues were handled by the proposed inverse consistency-based objective function weighting. The objective function (both the similarity measure and the regularization term) were weighted by the corresponding inverse consistency error during the fine-tuning step. As a result, the registration accuracy close to the missing tumor and its cavity was improved.

Team BDAV USYD[6]

This team adopted the recently proposed Non-Iterative Coarse-to-fine registration Network (NICE-Net)[7] as the backbone and extended it by introducing dual deep supervision. The NICE-Net consists of a feature learning encoder and a coarse-to-fine registration decoder. The feature learning encoder has two identical, weight-shared paths to extract features from the fixed and moving images separately, which were then propagated to the coarse-to-fine registration decoder. The decoder performs multiple steps of coarse-to-fine registration in a single network iteration. Dual deep supervision, including a deep self-supervised loss based on image similarity (local normalized cross-correlation) and a deep weakly-supervised loss based on manually annotated landmarks (mean square error), was embedded into the NICE-Net (referred as NICE-Net-ds). As the provided training set was relatively small (140 intra-patient image pairs), the NICE-Net-ds was first pretrained with inter-patient image pairs (280×279 pairs) to avoid overfitting. Then, the NICE-Net-ds was further trained for intra-patient registration with dual deep supervision. During inference, pair-specific fine-tuning was performed to improve the network's adaptability to testing variations. In addition, as the MRI scans provided by the challenge organizers had been rigidly registered to the same anatomical template, this method solely optimized for deformable image registration without considering affine registration.

Team CaMed[8]

An enhanced unsupervised learning-based method was developed for reliable and accurate registration of patient-specific brain MRI scans containing pathologies. The proposed method extends our previously developed registration framework iRegNet[9]. In particular, incorporating an unsupervised learning-based paradigm as well as several minor modifications to the network pipeline, allowed the enhanced iRegNet method to achieve respectable results. Similar to the baseline iRegNet, the registration procedure consists of two steps: First, I_B and I_F were fed into our convolutional neural network (CNN) that then predicts ϕ. Second, I_F was transformed into a warped image ($I_F.\phi$) using a spatial re-sampler.

- **Team Kurtlab**[10]

A two-stage cascaded network was developed consisting of the Inception and the TransMorph architecture. A series of Inception modules were initially used to fuse the 4 image modalities inputs and extract their most relevant information. In short, the Inception module was used to process each contrast separately, extracting the relevant information before concatenating them. The concatenated data was then passed through more Inception modules that merged the contrasts together and output new moving and target images. This approach had several advantages. First, it added more training parameters corresponding to the data merging layers. It also helped reduce the memory requirements by merging 4 volumetric images into a single one. The output of the Inception model was then sent into a variant of the TransMorph architecture to generate displacement fields for transforming the post-surgery images to their corresponding pre-surgery ones. TransMorph, a hybrid TransformerConvNet, was able to determine which parts of the input sequence were essential based on contextual information through the use of self-attention mechanisms. Finally, the loss function was composed of a standard image similarity measure and a diffusion regularizer.

- **Team YKW**[11]

The QPDIR algorithm was an intensity-based algorithm, which transforms the computation of deformation field to an optimization problem aimed at minimizing terms related to image dissimilarity and regularization. The terms were computed based on performing an exhaustive search among image blocks. The optimization was performed using a gradient-free quadratic penalty method. The whole optimization problem was decomposed to several sub-problems and each of them can be solved by straightforward block coordinate decent iteration. The QPDIR algorithm consisted of 3 steps. Firstly, the objective function was formulated by combining an image dissimilarity term and a regularization term determined through exhaustive search. The gradient-free quadratic penalty method was then employed to optimize the objective function. Next, the search window size was gradually reduced, and step 1 was repeated until the search window size reached a sufficiently small scale. Finally, the full displacement field was computed by applying moving least square (MLS). It's worth noting that multi-modal registration was achieved by fusing the results of single-modality registration.

- **Team MeDAL** and **MeDAL(pc)**[12]

The proposed method consists of two stages, a segmentation stage, and a subsequent registration stage. The segmentation stage consists of two U-Nets with shared parameters[13]. The U-Nets were similar, with each one of them containing three levels with residual blocks at every level and the number of feature maps starts from 8. Each U-Net segments the regions of interests (ROI), which are patches of sizes $(9 \times 9 \times 9)$ around the landmarks of the input volume. These landmarks can be found in either the moving volume or the fixed volume. The segmentation network was followed by an attention block in which the output of the U-Net (a binary segmentation map) was multiplied by the input volume to produce an attentive volume. The concatenated outputs (the attentive fixed and moving volumes) of the segmentation network serve as an input to the registration network. To tackle the problem of class imbalance between the foreground and the background of the segmentation mask, we used the focal loss between the segmentation masks and the predicted segmentation maps[14]. The architecture of the registration network was the same as the U-Net architecture used for segmentation. The difference was that the network for registration outputs a deformation field instead of a segmentation map. The deformation field was used for deforming the moving volume to match the fixed one. The loss function of the registration network was a combination of two losses, the similarity loss, and the smoothness loss.

MeDAL(pc): In the original challenge submission, the team had problems with loading the landmarks properly. After fixing the issues, the team was allowed to provide a post-challenge submission of their method, indicated with *(pc)*.

- **Team Fraunhofer MEVIS**[15]

The proposed method was an iterative variational image registration approach based on[16], in which the registration of two volumes can be modeled as the minimization of a discretized objective function. The final solution consists of a parametric and a non-parametric step. In the parametric step, the registration task was based on the estimation of a limited number of transformation parameters. In particular, a rigid registration, restricted to the search of rotations and translations parameters, was computed. In the parametric approach, the objective function only consists of the distance measure, computed between the fixed (post-operative) and warped moving (pre-operative) image. As a distance measure, we utilized the normalized gradient field (NGF)[17]. The transformation matrix obtained in the parametric registration initializes the non-parametric step, in which a deformation vector field was computed. In the deformable solution, the NGF was also utilized as a distance measure. To limit the possible registration solutions and make the deformation field more plausible, two regularization terms were added to the objective function. The first one was the curvature regularizer[18], which penalizes deformation fields having large second derivatives. Additionally, the volume change control was utilized to reduce foldings in the deformation field[19]. In the parametric and non-parametric steps, the registration was conducted on three levels using images at different scales. The deformation was initially computed on the coarsest level, where the images were downsampled by a factor equal to $2^{(L-1)}$. On a finer level, the previously computed deformations were utilized as an initial guess by warping the moving image. At each level, the moving and fixed images were downsampled. Furthermore, the choice of the optimal transformation parameters was conducted by

using the quasi-Newton l-BGFS, due to its speed and memory efficiency[20]. Each step of the proposed registration method can process only one MRI sequence at a time. Thus, we first verified on the training set of the BraTS-Reg challenge dataset which sequence was the best to guide the registration in both the rigid and non-rigid approaches. Our final solution used the T2 images in the parametric step, whereas the T1c acquisitions of the corresponding volumes guided the deformable registration.

Team SuperX

This method consists of two steps i) A rigid registration method, the Nelder-Mead method (also named downhill simplex method) and ii) Affine transformation for rigid registration algorithm which was applied on floating image before non-rigid registration was performed using free-formed deformations. In the non-rigid registration step, two optimization methods were investigated. The first one was particle swarm optimization, which is a computational method that optimizes a problem by iteratively improving a candidate solution according to a given measure of quality. The second optimization method that was examined was the downhill simplex method which is a commonly used local fast optimization. According to our experimental results, these methods achieved similar accuracy. We finally used the downhill simplex method to minimize the correlation coefficient as an image similarity function by downhill simplex optimization. In the non-rigid image registration, we used free-form deformation, where the force exerted on each floating voxel drives it to the correct position to match the reference volume. Our algorithm used scaling factors of 0.125, 0.25, and 0.5 to generate a multi-level pyramid and accelerate the running time. We initially optimized at a lower resolution level and then scaled up the warp field from lower to higher resolutions. According to Information Theoretic Similarity Measures in Non-Rigid Registration, the force field was based on joint entropies.

Team ANTs*

The BraTS-Reg22 data were processed using previously vetted and frequently used registration parameter sets[21,22], which have been packaged within the different ANTsX platforms, specifically ANTs, ANTsPy, and ANTsR (https://github.com/ANTsX). Each of these parameter sets consists of multiple transformation stages for determining anatomical correspondence. Initially, linear transform parameters are estimated, including center of (intensity) mass alignment followed by optimization of both rigid and affine transforms using mutual information as the similarity metric[23]. The final deformable alignment utilized symmetric normalization (SyN) with Gaussian[24] or B-spline[25] regularization and neighborhood cross-correlation[24] or mutual information similarity metric. The effects of image modality choice including all single modalities and combinations of modality pairs (specifically, T1-contrast enhanced/T2 and T1-contrast enhanced/FLAIR) were also explored. Although performance with the training data was similar across the different configurations, SyN with Gaussian regularization and neighborhood cross-correlation (radius = 2 voxels) using T1-contrast enhanced images was selected for a single submission during the validation phase. Further details on this internal evaluation, including the precise ANTsPy calls, can be found at a dedicated GitHub repository (https://github.com/ntustison/BraTS-Reg22).

SynthMorph* and **HyperMorph***

These approaches were build on VoxelMorph[26], a widely-used learning-based framework for pairwise deformable registration that aligns a moving image m with a fixed image f by predicting a dense correspondence ϕ. This framework leverages a convolutional architecture g_θ, with trainable parameters θ, that takes as input $\{m, f\}$ and outputs a stationary velocity field, integrated via squaring-and-scaling[27] to yield the diffeomorphic map ϕ_θ. VoxelMorph generally optimizes a loss combining an image matching term $\mathscr{L}_{sim}$ with a regularization term $\mathscr{L}_{reg}$ to encourage smooth deformations:

$$\mathscr{L}(m, f, \phi_\theta) = \mathscr{L}_{sim}(m \circ \phi_\theta, f) + \lambda \mathscr{L}_{reg}(\phi_\theta),$$

where $m \circ \phi_\theta$ represents m transformed by ϕ_θ, and λ is a regularization weight. In this challenge, we used normalized cross-correlation for $\mathscr{L}_{sim}$ and defined $\mathscr{L}_{reg}(\phi) = \frac{1}{2}||\nabla u||^2$, where u was the displacement of deformation ϕ. Network g_θ implements a U-Net[28], using convolutional kernels of size 3^3 and Leaky-ReLU activations. We treat m as the baseline and f as the follow-up scan for each subject, stacking all four contrasts to compose multi-channel input images.

VoxelMorph generally performs better on affinely aligned images and, like most learning-based approaches, does not generalize well to modalities outside the training domain. Furthermore, registration quality was often sensitive to the selected values of hyperparameters, such as λ. To address this set of challenges, two methods that implement and build on the VoxelMorph framework were evaluated: SynthMorph and HyperMorph.

SynthMorph is a fully-convolutional architecture for joint affine and deformable registration trained on a wide range of synthetic brain images, enabling the network to generalize across MRI contrasts, resolutions, and anatomies[29]. First, an affine network h_η, with parameters η, predicted separate feature maps for m and f, and aligned the barycenters of these maps using a least-squares regression, yielding a linear transform $T_\eta = h_\eta(m, f)$ that was applied to m[30]. Then, network g_θ was used to predict the deformation field $\phi_\theta = g_\theta(T_\eta \circ m, f)$.

A pre-trained SynthMorph model was fine-tuned to the challenge data using a semi-supervised approach, optimizing an additional loss term $\mathscr{L}_{sup}$ that measures the mean squared error between moving and fixed landmarks $\{x_m, x_f\}$:

$$\mathscr{L}(T_\eta, \phi_\theta, m, f, x_m, x_f) = \mathscr{L}_{sim}(m \circ T_\eta \circ \phi_\theta, f) + \lambda_1 \mathscr{L}_{reg}(\phi_\theta) + \lambda_2 \mathscr{L}_{sup}(x_m, x_f),$$

with $\lambda_1 = 32$ and $\lambda_2 = 1$. The networks h_η and g_θ were initialized with pre-trained weights and the affine component h_η before fine-tuning g_θ. The model comprised 20 total convolutions of 256 channels each. During training a learning rate of 10^{-5} was used and data augmentation was applied following[31].

HyperMorph facilitates training a *single* model encompassing a landscape of possible values for λ[32,33]. This avoids the need to train several separate models and allows for fine-scale hyperparameter choice during inference. In HyperMorph, a hypernetwork learns to model the effect of varying λ to predict the corresponding parameters of g_θ.

Using a subset of the challenge training data, first HyperMorph was trained and the effect of λ on a held-out validation set was evaluated. Based on a combination of evaluating both visual deformation quality and distance between annotated landmarks, $\lambda = 2.4$ was selected and the base VoxelMorph model was trained until convergence. The pre-trained affine component of SynthMorph was employed as a pre-processing step. The model g_θ comprised 12 total convolutions, each with 64 channels, and was trained with a learning rate of 10^{-4}, using the augmentation strategy noted above.

- **Team GradICON***

GradICON's training protocol and hyperparameters[34] were adopted. Its ability to generalize was assessed by investigating its performance without explicitly modeling image differences due to tumor resection with two significant changes in the original approach. Initially, the number of input channels of the first convolutional layer was increased to match the number of modalities. This adjustment enabled the utilization of visual cues across different modalities. The image similarity was computed by defining the local normalized cross correlation (LNCC) as an average over the LNCCs for each modality (channel). The second modification consisted of a new training strategy to alleviate overfitting caused by the small available training dataset. The network was pre-trained following the original training process in[34] with inter-patient pairs from the train set. Subsequently, the entire network (Stage1 and Stage2) was fine-tuned using intra-patient pairs from the train set. The input images were normalized to [0, 1] per modality. In the pre-train phase, random pairs of pre-operative and follow-up images were picked as training pairs. In the fine-tune phase, the paired images provided by the challenge were used.

In both phases, the network was trained with the GradICON's default hyperparameters (i.e., learning rate of $5e^{-5}$, regularizer weight $\lambda = 1.5$, and ADAM optimizer). Due to memory limitations, a batch size of 2 was used when training Stage 1[34]. A batch size of 1 was used for Stage 2[34] in the inter-patient (pre-training) phase and in the intra-patient training (fine-tuning) phase. The network was trained for 20,000 iterations for pre-training and 10,000 iterations for fine-tuning. The protocol described in[34] was followed for inference where a displacement field was directly predicted using the trained network from the input images without any pre-alignment. Subsequently, 50 iterations of instance optimization were used. Due to memory limitations, this optimisation was only over Stage 2 with a learning rate of $5e^{-5}$ keeping Stage 1 frozen. The implementation was available at https://github.com/uncbiag/ICON.

- **Team MICS***

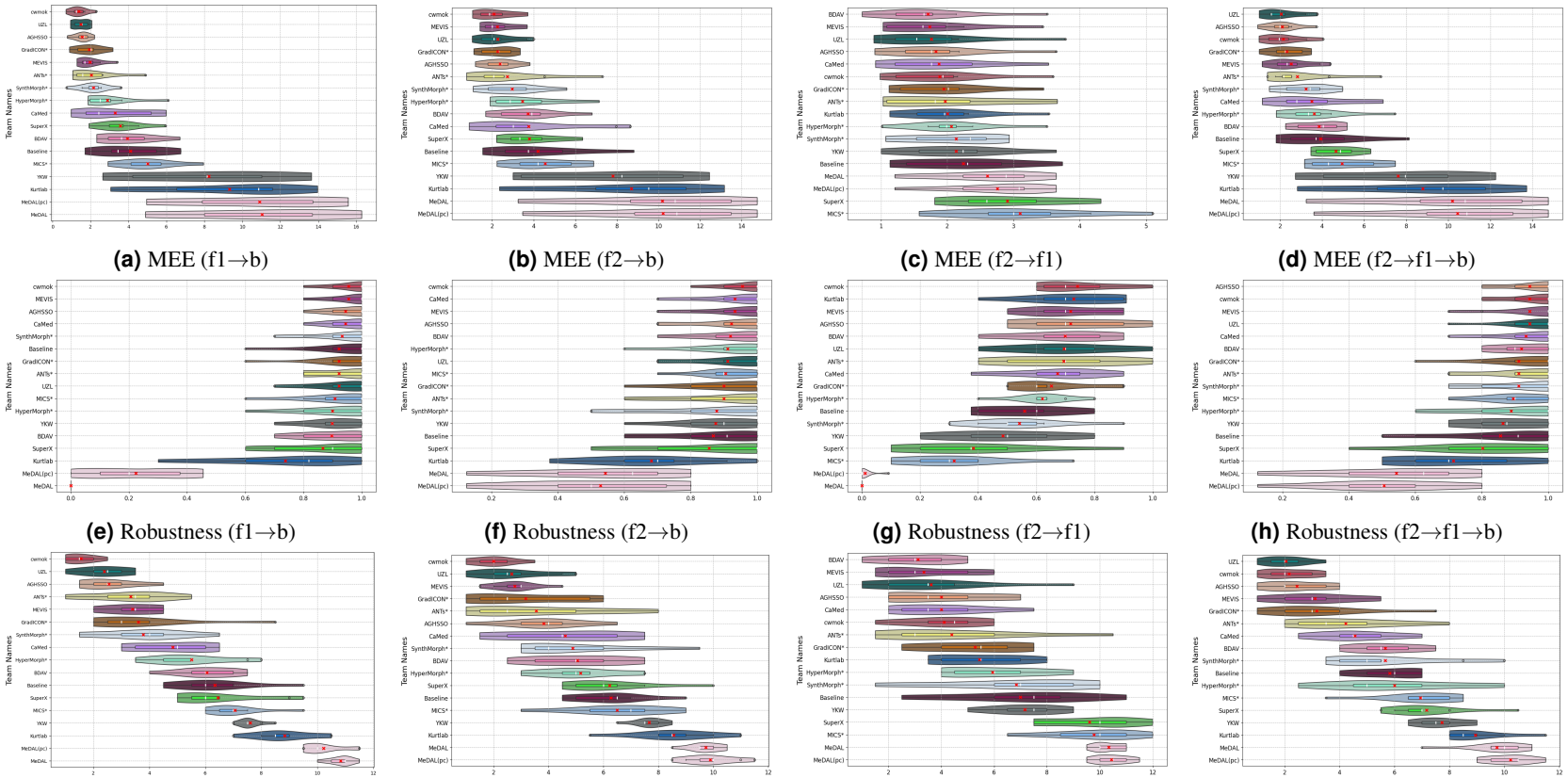
The method was based on the multi-task registration framework presented in[35]. In summary, the architecture utilized a shared UNet-like encoder that maps both inputs (moving M and reference R) to a latent code using a late fusion strategy. The final layer of a UNet-like registration decoder, was based on the formulation presented in[36]. Specifically, a cumulative sum across all dimensions was performed on the output of a logistic growth function (i.e., $f(x) = L/1 + e^{-k(x-x_0)}$, with $L = 8$, $k = 1$, $x_0 = 1$) that forces the values to be strictly positive an on a specific range. By enforcing these displacements to have positive values and subsequently applying an integration operation along each dimension, the spatial sampling coordinates can be retrieved. Such an approach ensures the generation of smooth deformations that avoid self-crossings, while allowed the control of maximum displacements among consecutive pixels using appropriate parameterization of the logistic growth function. The training was performed in two steps. In the first step a random pairing scheme was utilized to generate different input pairs drawn from different cases. In the second step, the model that was derived was fine-tuned only on baseline-follow-up pairs for each patient individually. The losses that were utilized for these two training steps were $L1$ and $L2$ respectively:

$$L1 = L_{rec} + L_{ncc} + L_{jac} + L_{cycl}$$
$$L2 = L_{rec} + L_{ncc} + L_{jac} + L_{cycl} + L_{lndmrk}$$

where L_{rec} was the mean squared error between the target and deformed image, L_{ncc} was the normalized cross correlation loss between the target and deformed image, L_{jac} was the Jacobian determinant of the deformation field, L_{lndmrk} was the mean absolute error between the deformed and target landmark locations and lastly, L_{cycl} was the mean squared error between the moving image and the output of a sequential application of a forward and backward deformation. It should be noted that, in both training stages a cycling scheme was used where all the aforementioned losses were averaged by considering both inputs as moving/reference images, respectively.

Table S1. Summary of participating methods. BCD - block coordinate descent, CR - curvature regularizer, DF - displacement field, DR - diffusive regularization, GD - gradient descent, GE - gradient error, (G)IC - (gradient) inverse consistency, HM - histogram matching, IN - intensity normalization, IO - Instance Optimization, JL - jacobian loss, LL - landmark loss, (L)(N)CC - (local) (normalized) cross correlation, MI - mutual information, MSE - mean squared error, N/A - no information provided, NGF - normalized gradient fields, RL - reconstruction loss, SM - Symmetric Cropping, VCC - volume change control.

Team	MRI sequences	Loss function	Optimizer	Affine prealignment	Landmarks	Pre-processing	Post-processing	Data augmentation
AGHSSO	T1, T1CE, T2, FLAIR	NCC, DR, IC	Adam	Yes	No	IN	N/A	N/A
ANTs*	T1CE	CC	GD	Yes	No	N/A	N/A	N/A
BDAV	T1, T1CE	NCC, MSE	Adam	No	Yes	IN	N/A	random pairing
CaMed	T1CE	GE, NCC	Adam	Yes	Yes	IN	N/A	random cropping patches
cwmok	T1, T1CE, T2, FLAIR	NGF, LNCC		Yes	No	IN	IO	rotation, translation
GradICON*	T1, T1CE, T2, FLAIR	LNCC, GIC	Adam	No	No	IN	IO	translation
HyperMorph*	T1, T1CE, T2, FLAIR	NCC	Adam	Yes	No	IN	N/A	linear transforms, Gaussian noise, gamma-exponentiating
Kurtlab	T1, T1CE, T2, FLAIR	MSE, DR	Adam	Yes	No	IN, HM	Gaussian smoothing	flipping, rotation
MeDAL	T1, T1CE	NCC, MI, L2	Adam	No	Yes	SM, IN	N/A	rotation, scaling, translation.
MEVIS	T1CE, T2	NGF, CR, VCC	quasi-Newton l-BGFS	No	No	N/A	N/A	N/A
MICS*	T1, T1CE, T2, FLAIR	RL, NCC, JL, LL, MSE	Adam	No	Yes	N/A	N/A	N/A
SuperX	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
SynthMorph*	T1, T1CE, T2, FLAIR	NCC, MSE, DR	Adam	No	Yes	IN	N/A	spatial deformation, shift, rotation, scaling, shear, Gaussian noise, gamma augmentation, bias field, blurring
UZL	T1CE	MIND-MSE, DR	Adam	No	No	N/A	N/A	N/A
YKW	T1	DF Similarity/Regularity	BCD	No	No	SM	N/A	N/A



(a) MEE (f1→b) **(b)** MEE (f2→b) **(c)** MEE (f2→f1) **(d)** MEE (f2→f1→b)

(e) Robustness (f1→b) **(f)** Robustness (f2→b) **(g)** Robustness (f2→f1) **(h)** Robustness (f2→f1→b)

(i) BraTS-Reg score (f1→b) **(j)** BraTS-Reg score (f2→b) **(k)** BraTS-Reg score (f2→f1) **(l)** BraTS-Reg score (f2→f1→b)

Figure S1. Comparative performance analysis of various participating methods using MEE (top row), Robustness (middle row), and BraTS-Reg score metrics (bottom row) for four longitudinal registration tasks. (f1 = follow-up 1, f2 = follow-up 2, b = baseline.)

References

1. Mok, T. C. & Chung, A. C. Robust image registration with absent correspondences in pre-operative and follow-up brain mri scans of diffuse glioma patients. In *International MICCAI Brainlesion Workshop*, 231–240 (Springer, 2022).
2. Mok, T. C. & Chung, A. Unsupervised deformable image registration with absent correspondences in pre-operative and post-recurrence brain tumor mri scans. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 25–35 (Springer, 2022).
3. Mok, T. C. W. & Chung, A. C. S. Conditional deformable image registration with convolutional neural network. In *Medical Image Computing and Computer Assisted Intervention – MICCAI 2021*, 35–45 (Springer International Publishing, Cham, 2021).
4. Großbröhmer, C., Siebert, H., Hansen, L. & Heinrich, M. P. Employing convexadam for brats-reg. In *International MICCAI Brainlesion Workshop*, 252–261 (Springer, 2022).
5. Wodzinski, M., Jurgas, A., Marini, N., Atzori, M. & Müller, H. Unsupervised method for intra-patient registration of brain magnetic resonance images based on objective function weighting by inverse consistency: Contribution to the brats-reg challenge. In *International MICCAI Brainlesion Workshop*, 241–251 (Springer, 2022).
6. Meng, M., Bi, L., Feng, D. & Kim, J. Brain tumor sequence registration with non-iterative coarse-to-fine networks and dual deep supervision. In *International MICCAI Brainlesion Workshop*, 273–282 (Springer, 2022).
7. Meng, M., Bi, L., Feng, D. D. F. & Kim, J. *Non-iterative Coarse-to-Fine Registration Based on Single-Pass Deep Cumulative Learning*, 88–97 (2022).
8. Zeineldin, R. A., Karar, M. E., Mathis-Ullrich, F. & Burgert, O. Self-supervised iregnet for the registration of longitudinal brain mri of diffuse glioma patients. In *International MICCAI Brainlesion Workshop*, 25–34 (Springer, 2022).
9. Zeineldin, R. A. *et al.* iregnet: Non-rigid registration of mri to interventional us for brain-shift compensation using convolutional neural networks. *IEEE Access* **9**, 147579–147590, DOI: 10.1109/ACCESS.2021.3120306 (2021).
10. Abderezaei, J., Pionteck, A., Chopra, A. & Kurt, M. 3d inception-based transmorph: Pre- and post-operative multi-contrast mri registration in brain tumors. In *International MICCAI Brainlesion Workshop*, 45–34 (Springer, 2022).
11. Yan, K. & Yan, Y. Applying quadratic penalty method for intensity-based deformable image registration on brats-reg challenge 2022. In *International MICCAI Brainlesion Workshop*, 3–14 (Springer, 2022).
12. Almahfouz Nasser, S., Kurian, N. C., Meena, M., Shamsi, S. & Sethi, A. Wssamnet: Weakly supervised semantic attentive medical image registration network. In *International MICCAI Brainlesion Workshop*, 15–24 (Springer, 2022).
13. Schwarz, L. A. Non-rigid registration using free-form deformations. *Tech. Univ. München* **6**, 4 (2007).
14. Machado, I. *et al.* Deformable mri-ultrasound registration via attribute matching and mutual-saliency weighting for image-guided neurosurgery. In *Simulation, Image Processing, and Ultrasound Systems for Assisted Diagnosis and Navigation*, 165–171 (Springer, 2018).
15. Canalini, L. *et al.* Iterative method to register longitudinal mri acquisitions in neurosurgical context. In *International MICCAI Brainlesion Workshop*, 262–272 (Springer, 2022).
16. Modersitzki, J. *FAIR: flexible algorithms for image registration* (SIAM, 2009).
17. Haber, E. & Modersitzki, J. Intensity gradient based registration and fusion of multi-modal images. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 726–733 (Springer, 2006).
18. Fischer, B. & Modersitzki, J. Curvature based image registration. *J. Math. Imaging Vis.* **18**, 81–85 (2003).
19. Rühaak, J. *et al.* Estimation of large motion in lung ct by integrating regularized keypoint correspondences into dense deformable registration. *IEEE transactions on medical imaging* **36**, 1746–1757 (2017).
20. Liu, D. C. & Nocedal, J. On the limited memory bfgs method for large scale optimization. *Math. programming* **45**, 503–528 (1989).
21. Avants, B. B. *et al.* A reproducible evaluation of ants similarity metric performance in brain image registration. *Neuroimage* **54**, 2033–2044 (2011).
22. Avants, B. B. *et al.* The insight toolkit image registration framework. *Front. neuroinformatics* **8**, 44 (2014).
23. Viola, P. & Wells III, W. M. Alignment by maximization of mutual information. *Int. journal computer vision* **24**, 137–154 (1997).

24. Avants, B. B., Epstein, C. L., Grossman, M. & Gee, J. C. Symmetric diffeomorphic image registration with cross-correlation: evaluating automated labeling of elderly and neurodegenerative brain. *Med. image analysis* **12**, 26–41 (2008).

25. Tustison, N. J. & Avants, B. B. Explicit b-spline regularization in diffeomorphic image registration. *Front. neuroinformatics* **7**, 39 (2013).

26. Balakrishnan, G., Zhao, A., Sabuncu, M. R., Guttag, J. & Dalca, A. V. Voxelmorph: A learning framework for deformable medical image registration. *IEEE Transactions on Med. Imaging* **38**, 1788–1800 (2019).

27. Dalca, A. V., Balakrishnan, G., Guttag, J. & Sabuncu, M. R. Unsupervised learning for fast probabilistic diffeomorphic registration. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 729–738 (2018).

28. Ronneberger, O., Fischer, P. & Brox, T. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 234–241 (2015).

29. Hoffmann, M. *et al.* Synthmorph: learning contrast-invariant registration without acquired images. *IEEE Transactions on Med. Imaging* **41**, 543–558 (2022).

30. Hoffmann, M., Hoopes, A., Fischl, B. & Dalca, A. V. Anatomy-specific acquisition-agnostic affine registration learned from fictitious images. In *Medical Imaging 2023: Image Processing* (SPIE, 2023).

31. Hoffmann, M., Hoopes, A., Greve, D. N., Fischl, B. & Dalca, A. V. Anatomy-aware and acquisition-agnostic joint registration with synthmorph. *arXiv preprint arXiv:2301.11329* DOI: 10.48550/ARXIV.2301.11329 (2023).

32. Hoopes, A., Hoffmann, M., Fischl, B., Guttag, J. & Dalca, A. V. Hypermorph: Amortized hyperparameter learning for image registration. In *Information Processing in Medical Imaging: 27th International Conference, IPMI 2021, Virtual Event, June 28–June 30, 2021, Proceedings 27*, 3–17 (Springer, 2021).

33. Hoopes, A., Hoffmann, M., Greve, D. N., Guttag, J. & Dalca, A. V. Learning the effect of registration hyperparameters with hypermorph. *Mach. Learn. for Biomed. Imaging* **1** (2022).

34. Tian, L. *et al.* Gradicon: Approximate diffeomorphisms via gradient inverse consistency. *arXiv preprint arXiv:2206.05897* (2022).

35. Estienne, T. *et al.* Deep learning-based concurrent brain registration and tumor segmentation. *Front. computational neuroscience* 17 (2020).

36. Christodoulidis, S. *et al.* Linear and deformable image registration with 3d convolutional neural networks. In *Image Analysis for Moving Organ, Breast, and Thoracic Images: Third International Workshop, RAMBO 2018, Fourth International Workshop, BIA 2018, and First International Workshop, TIA 2018, Held in Conjunction with MICCAI 2018, Granada, Spain, September 16 and 20, 2018, Proceedings 3*, 13–22 (Springer, 2018).