

5

REtroSpective Evaluation of Cerebral Tumors (RESECT): a clinical database of pre-operative MRI and intra-operative ultrasound in low-grade glioma surgeries

10 ^{1,2}Yiming Xiao, ^{1,2}Maryse Fortin, ^{3,4,6}Geirmund Unsgård, ^{1,2}Hassan Rivaz, and ^{5,6}Ingerid Reinertsen

¹PERFORM Centre, Concordia University, Montreal H4B 1R6, Canada

²Department of Electrical and Computer Engineering, Concordia University, Montreal H3G 1M8, Canada

³Department of Neurosurgery, St. Olavs University Hospital, Trondheim NO-7006, Norway

15 ⁴Department of Neuroscience, Norwegian University of Science and Technology, Trondheim NO-7491, Norway

⁵SINTEF, Department of Medical Technology, Trondheim NO-7465, Norway

⁶Norwegian National Advisory Unit for Ultrasound and Image Guided Therapy, St. Olavs University Hospital, Trondheim NO-7006, Norway

20

Corresponding Author

Yiming Xiao

Rm 2.211, PERFORM Centre
Concordia University
7200 Sherbrooke St. W.
Montreal, Quebec
Canada H4B 1R6

25

E-mail: yiming.xiao@concordia.ca

Telephone: (514) 848-2424 ext: 4453

30

Abstract

35 **Purpose:** The advancement of medical image processing techniques, such as image registration, can effectively help improve the accuracy and efficiency of brain tumor surgeries. However, it is often challenging to validate these techniques with real clinical data due to the rarity of such publicly available repositories.

Acquisition and validation methods: Pre-operative magnetic resonance images (MRI), and intra-
40 operative ultrasound (US) scans were acquired from 23 patients with low-grade glioma who underwent surgeries at St. Olavs University Hospital between 2011 and 2016. Each patient was scanned by Gadolinium-enhanced T1w and T2-FLAIR MRI protocols to reveal the anatomy and pathology, and series of B-mode ultrasound images were obtained before, during, and after tumor resection to track the surgical progress and tissue deformation. Retrospectively, corresponding
45 anatomical landmarks were identified across US images of different surgical stages, and between MRI and US, and can be used to validate image registration algorithms. Quality of landmark identification was assessed with intra- and inter-rater variability.

Data format and access: In addition to co-registered MRIs, each series of US scans are provided as a reconstructed 3D volume. All images are accessible in MINC2 and NIFTI formats, and the
50 anatomical landmarks were annotated in MNI tag files. Both the imaging data and the corresponding landmarks are available online as the RESECT database at <https://archive.norstore.no>.

Potential impact: The proposed database provides real high-quality multi-modal clinical data to validate and compare image registration algorithms that can potentially benefit the accuracy and
55 efficiency of brain tumor resection. Furthermore, the database can also be used to test other image processing methods and neuro-navigation software platforms.

Key words: database, intra-operative ultrasound, MRI, low-grade glioma, brain tumor, registration

60 I. Purpose

Gliomas are primary central nervous system (CNS) tumors that originate from the glial cells and infiltrate the surrounding tissues. As currently the most common brain tumors in adults¹, they are categorized in grade I-IV based on histological characteristics outlined by the World Health Organization (WHO)². While low-grade gliomas (LGG) are grade I and II gliomas that grow more slowly and have higher survival rate than the high-grade ones (grade III and IV), the grade II gliomas will eventually progress to high grade types and lead to death. Evidence³⁻⁵ has suggested that tumor resection can effectively improve the patient's survival rate, but the procedure can be difficult without good image-guidance. One major reason is that soft tissue deformation can displace the surgical target and vital structures (e.g., blood vessels) shown in pre-operative images due to multiple factors⁶⁻⁸ (e.g., drug administration, intracranial pressure change, tissue removal) while these displacements may not be directly visible in the surgeon's field of view.

Intra-operative imaging modalities such as intra-operative magnetic resonance images (MRI)⁸ and ultrasound (US)^{6, 9} can be used to track the progress of the resection and tissue deformation. Although intra-operative MRI provides image contrasts that are easier to comprehend, intra-operative US is more commonly seen thanks to its low cost, high portability and flexibility. Additionally, tumors can often be delineated in US images even when they are not distinguishable from normal tissues under the microscope. This can facilitate accurate resection and result in better surgical outcomes. To update the surgical plan in the presence of continuous tissue shift, automatic image registration algorithms can be used to recover the deformation by aligning the pre-operative images with intra-operative images. In contrast to the surgical judgement by direct visual comparison between pre- and intra-operative images, automatic image registration can offer more intuitive and potentially more accurate clinical assessments of tumor removal while avoiding displaced vital structures, such as blood vessels, the ventricles, and critical motor and sensory cortex⁹.

85 In the past years, a large number of image registration algorithms¹⁰⁻¹² have been proposed in the literature, which could greatly benefit the patients and clinicians in the operating room (OR). However, validation and comparison of these methods with real clinical data have been challenging, and thus posing the difficulty in transferring these potentially beneficial registration

90 algorithms into the operating room (OR). This is largely due to the fact that many technical institutes do not have affiliated hospitals or direct access to real well-annotated clinical data. To solve this issue, previously, there have been a number of publicly available clinical datasets that can be used for the evaluation of image registration algorithms. As for brain image registration, the Retrospective Image Registration Evaluation (RIRE) project, formerly known as the
95 Retrospective Registration Evaluation Project (RREP)¹³, is one of the earliest data repositories to offer open access to real clinical images with registration ground truths. The project contains Computed Tomography (CT), MRI and Positron emission tomography (PET) images, and was designed to validate and compare inter-modality image registration techniques. The registration ground truths were defined using a prospective, marker-based technique, and were hidden from
100 the users for fair comparison. The Brain Web (www.bic.mni.mcgill.ca/brainweb) database contains 20 raw brain images simulated from healthy subjects with segmented brain structures. The LONI-LPBA40¹⁴ database and the Non-rigid Image Registration Evaluation Project (NIREP) (www.nirep.org) contain healthy brain MRIs with regions of interest (ROIs) annotated by neurologists (32 ROIs in the NIREP database and 56 ROIs in the LONI-LPBA40 database). They
105 can be used to benchmark inter-subject non-linear registration methods through examining the overlap of the segmented labels after registration. The Alzheimer's Disease Neuroimaging Initiative (ADNI) database¹⁵ offers a large number of multi-contrast MRIs and PET images of both healthy and diseased brains, aiming to help provide a better understanding of the disease. As the hippocampus segmentations are available for many ADNI subjects, they can also be used for
110 evaluating inter-subject registration algorithms against the impact of multi-site MRI scanner differences and disease-induced structural changes.

Despite the progress in promoting public databases to validate medical image registration, datasets that provide both MRIs and intra-operative images specifically for brain cancer are still scarce. So far, the only public data repository of such type is the BITE dataset¹⁶ hosted by the Montreal
115 Neurological Institute. It contains both MRIs and intra-operative US scans of 14 patients with brain cancer, along with intra- and inter-modality homologous landmarks that were selected manually. Although the dataset has been instrumental for the medical imaging community in different contexts¹⁷⁻²¹, the technology used to collect the US scans (ATL/Philips HDI5000 ultrasound scanner with a P7-4 phased-array transducer) is no longer concurrent in the clinic, and newer US
120 scanners have provided better image quality. As a result, there is a need for similar datasets that

contain high quality US scans obtained with more up-to-date US technology. In addition, as the surgical flow, equipment, familiarity with intra-operative US, imaging data acquisition protocols can differ at different medical centers, it is beneficial to obtain multi-center clinical data of this type for validation purposes other than the BITE database from the MNI.

125

In our REtroSpective Evaluation of Cerebral Tumors (RESECT) database, we have included pre-operative MR and intra-operative US images from 23 patients who have received LGG resection surgeries (no follow-up re-operations) at St. Olavs University Hospital. The new dataset has five major differences from the BITE dataset¹⁶. First, the quality of both US and MR images are better in the new dataset. US images were obtained with a complete tumor coverage using a more recent Sonowand Invite (Trondheim, Norway) intraoperative imaging system and linear probes with a higher operating frequency than the one used for the BITE dataset. This system is specifically designed for neurosurgical applications and thus produces improved US images of high quality. MR images for all patients except three were acquired on a 3T scanner, as opposed to 1.5T in BITE dataset¹⁶, and 3T scanners typically produces lower noise and better tissue contrast than the 1.5T ones. The image quality differences between two datasets are evident via direct visual inspection. Second, the pre-operative MR images were systematically acquired the day before surgery, whereas in the BITE dataset, the MR acquisition time varied and averaged 17 days¹⁶. Third, instead of acquiring data purely for research purposes with intra-operative US training offered by a researcher on site¹⁶, the US data were captured by clinicians with rich experience in intra-operative US, which has been routinely adopted in neurosurgery at St. Olavs University Hospital since 1997²². During each procedure, the obtained US data were actively employed to guide the resection. Fourth, due to the relatively low quality of US images, an average of 9 landmarks were provided in the BITE database¹⁶. In this work, we provide substantially more landmarks, which can lead to better assessment of registration accuracy. Finally, Surgicel (Ethicon, Somerville, NJ), a blood clotting agent that is placed around the resection cavity generates strong shadowing artifacts in the BITE dataset¹⁶. No Surgicel was used in the operations at St. Olavs University Hospital.

130

135

140

145

In addition to the imaging data, homologous landmarks were manually selected across US images of different surgical stages, and between MRIs and US scans showing before and after resection. These can be used to validate intra-modality and inter-modality registration algorithms for various

150

levels of tissue deformation at different surgical stages. Aside from validating medical image registration methods, the images provided in the dataset can also be used for developing other image processing methods (e.g., segmentation²³ and denoising²⁴) and image visualization strategies²⁵. As the multi-modal data reflects the real clinical workflow, they can also be used to test neuro-navigation platforms, such as 3D Slicer²⁶, CustusX²⁷, and IBIS²⁸, and to train surgeons for US-guided tumor resection as US interpretation is often the first obstacle for many novel users.

II. Acquisition and validation methods

II.A. Clinical data

We included 23 clinical cases of low-grade gliomas (Grade II) from the adult patients of brain tumors, who underwent surgeries between 2011 and 2016 at St. Olavs University Hospital. The subjects were selected with good image quality (no severe image artifacts) and US volumes that have good coverage of the resection sites. There is no selection bias and the dataset includes tumors at various locations within the brain. In this dataset, we focused on low-grade gliomas as it is where intra-operative ultrasound is the most useful. The collection and publication of the data was approved by the Regional Committee for Medical and Health Research Ethics of Central Norway, ref 2013/270. All patients signed written informed consent.

II.B. Pre-operative MR imaging

Pre-operative MR images were acquired the day before surgery. The MR protocol included a sagittal T1w Gd-enhanced sequence (TE=2.96 ms, TR=2000 ms, flip angle= 8 deg., 192 sagittal slices, acquisition matrix = 256x256, voxel size=1.0x1.0x1.0 mm³), and a sagittal T2w fluid-attenuated inversion recovery, or FLAIR (TE=388ms, TR=5000 ms, flip angle=120 deg., 192 sagittal slices, acquisition matrix = 256x256, voxel size=1.0x1.0x1.0 mm³) sequence both acquired on a 3T Magnetom Skyra (Siemens, Erlangen, Germany) with a 32-channel head coil. The combined imaging time for both T1w and T2 FLAIR scans was 12 min. For patients 2, 14 and 15, pre-operative MR images were acquired on a 1.5T Magnetom Avanto (Siemens, Erlangen, Germany) with a 12-channel head coil, and included a sagittal T1w Gd-enhanced sequence (TE=2.30 ms, TR=2500 ms, flip angle= 7 deg., 176 sagittal slices, slice thickness=1 mm, acquisition matrix=512x496, in-plane resolution=0.5x0.5mm²), and a sagittal FLAIR (TE=333ms, TR=6000 ms, flip angle=120 deg., 176 sagittal slices, slice thickness=1 mm, acquisition matrix=256x224, in-plane resolution=1.0x1.0mm²). The combined imaging time for both T1w and

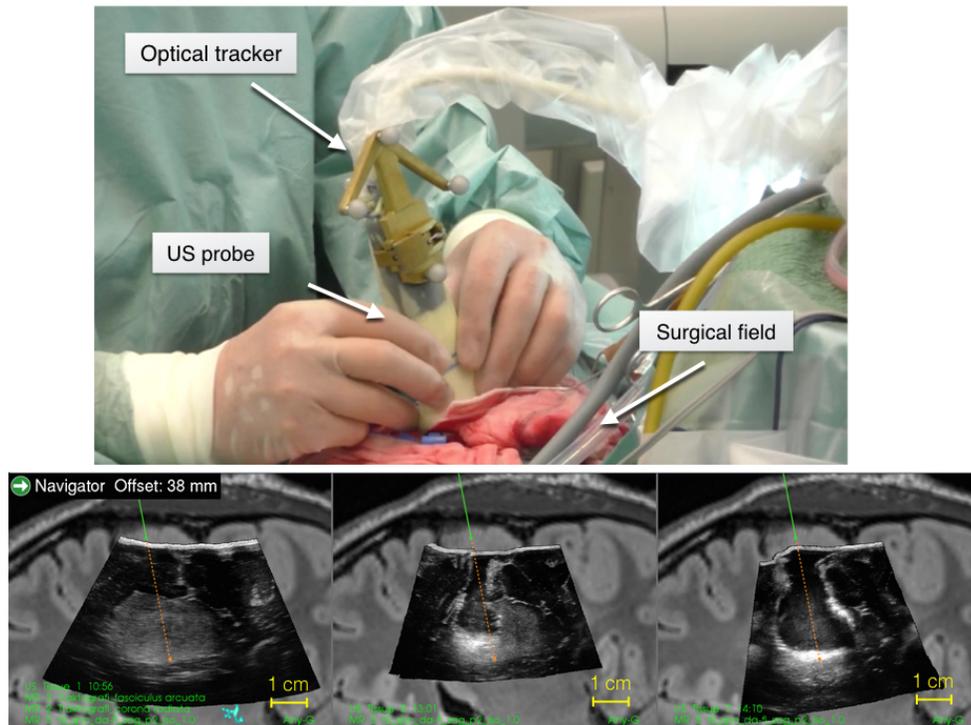
T2 FLAIR scans was 15min. Five fiducial markers were glued to the patient's head prior to scanning. The markers were used for image-to-patient registration after head immobilization on the operating table. MR images were transferred over the hospital network to the neuro-navigation system before surgery.

II.C. Intra-operative ultrasound imaging

The intra-operative 3D ultrasound images were acquired using the Sonowand Invite neuronavigation system (Sonowand AS, Trondheim, Norway). The most commonly used probe was the 12FLA-L linear probe with a frequency range of 6-12 MHz and a footprint of 48x13 mm. For smaller superficial tumors, we used the 12FLA flat linear array probe with a frequency range of 6-12 MHz and a footprint of 32x11 mm. The ultrasound probes were all factory calibrated and equipped with removable sterilizable reference frames for optical tracking. A Polaris infra-red camera (NDI, Waterloo, Canada) built in the Sonowand Invite neuronavigation system was used to capture the position and pose of the ultrasound probe via the optical trackers attached (see Fig. 1). As a result, the US images can be aligned with the surgical plans obtained previously so that tissue shift can be truthfully reflected in the intra-operative US images. The raw ultrasound data were reconstructed to 3D volumes using the built-in proprietary reconstruction method in the Sonowand Invite system, with a reconstruction resolution in the range of $0.14 \times 0.14 \times 0.14 \text{ mm}^3$ to $0.24 \times 0.24 \times 0.24 \text{ mm}^3$ depending on the probe types and imaging depth. The probes were covered with sterile probe drapes prior to image acquisition. The first ultrasound acquisition was usually performed before opening the dura. In cases where the image quality was not satisfactory due to artifacts from calcifications for example, an additional acquisition was performed after dura opening. For acquisitions during and after resection, the cavity was cleaned and filled with saline water. All patients were positioned in order to have an approximately horizontal craniotomy for the cavity to retain saline water and thus optimize ultrasound image quality. The setup for US acquisition and the demonstration of tracked US in the stages of before, during and after tumor resection as displayed in the surgical navigation system are shown in *Fig. 1*. For publication, we selected the first ultrasound volume acquired on the dura or the cortical surface, one ultrasound volume acquired during resection where there is still residual tumor, and the last ultrasound volume acquired for resection control. No resection was performed after this final acquisition. As such, the

images in this database have been acquired by expert users to guide the actual resection of the tumors, and not primarily for research purposes.

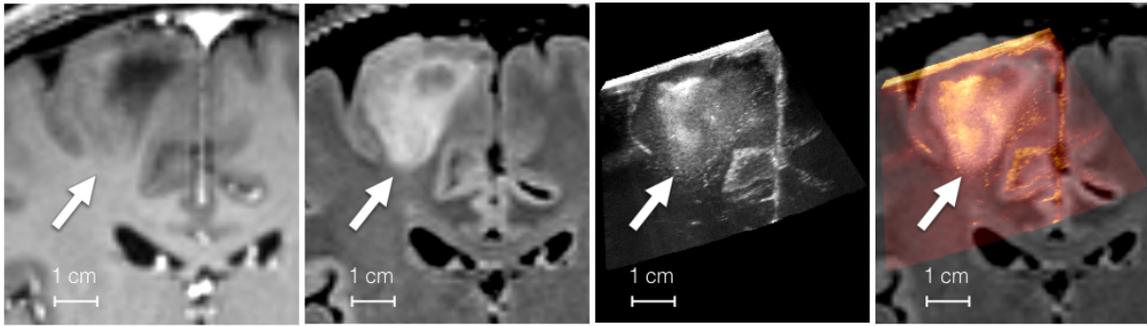
215



220 **FIG. 1.** Setup for intra-operative US acquisition in the operating room (top) and tracked
US image acquisition as shown in the surgical navigation system before, during and after
resection (bottom row from left to right, respectively). Note that the tumor appears bright
in the US image and the resection left a cavity in the tissue.

225 **II.D. Low-grade glioma characteristics on MRI and US**

On T1-weighted MRIs, low-grade gliomas appear dark in the image, and calcifications may appear
as foci of high T1 signals while T2 FLAIR images show the contrast between infiltrating tumor
(bright signal) margins and normal brain tissues. A comparison of tumor characteristics between
MRIs and US obtained from Patient #12 is shown in *Fig.2*. Because the contrast between the LGGs
230 and normal tissues are better in FLAIR images, they are used more actively than T1w MRIs during
surgery.

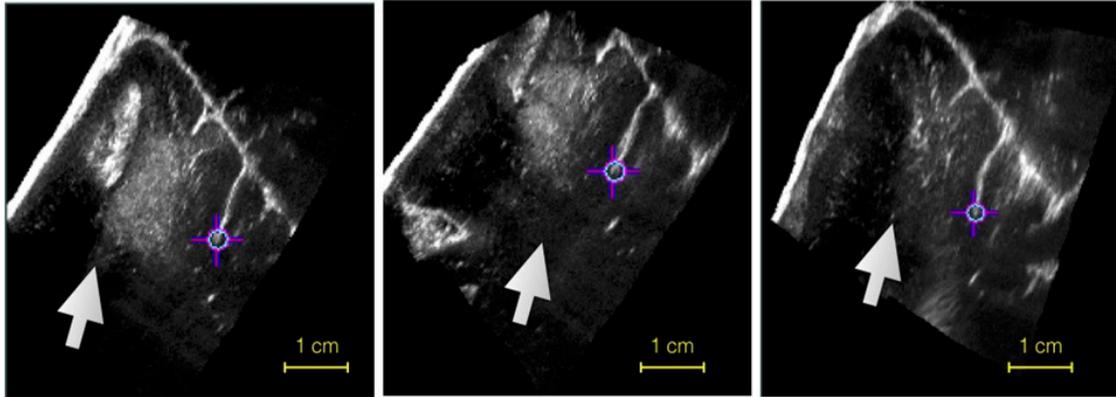


235 **FIG. 2.** Demonstration of pre-operative MRIs and intra-operative US of Patient #12 showing the characteristics of brain tumors. From left to right: T1w MRI, T2 FLAIR MRI, ultrasound image that was aligned to the MRIs, and overlay of T2 FLAIR MRI and the corresponding ultrasound image. The white arrows point to the brain tumor.

240 **II.E. Intra- and inter-modality landmarks for validation**
 In general, two types of homologous landmarks are provided in our database: US vs. US and MRI vs. US. We used the software named ‘register’ included in the MINC toolkit (<http://bic-mni.github.io>) to visualize the volumetric MRI and US data and create the homologous landmarks
 245 for the RESECT database. No image registration was performed using the software.

For the first type, two sets of landmarks were identified. First, coherent landmarks were selected across 3D US volumes acquired at the stages of before, during and after tumor resection, and for each patient, 10~17 unique landmarks were identified. We refer to this set of landmarks as Set 1.
 250 In Set 1, the total landmark numbers are usually limited by the available image features in US volumes after resection because of tissue removal. To enable more landmarks for inter-modality registration validation, we then tagged additional landmarks between US volumes before and during resection, and thus increased the number of landmarks to 16~34 per patient. This is referred to as Set 2. For both sets, landmarks chosen in pre-resection US were used as references to select
 255 those corresponding ones in other US volumes. Here, eligible landmarks include deep grooves and corners of sulci, convex points of gyri, and vanishing points of sulci. As the available image features differ with respect to the tumor location and the resection volume, we have obtained the landmarks of the first type for 17 patients. A demonstration of Landmark Set 1 for the three stages

of tumor resection is shown in *Fig. 3*. The details of the landmarks are included in Table I, where
260 the number of landmarks, and mean initial Euclidean distance between landmark pairs are
specified.



265

FIG. 3. Demonstration of a corresponding landmark in US volumes before, during and
after US resection (not co-registered). From left to right: US volumes before, during and
after resection. The landmark location is marked with a cursor and the white arrows point
270 to the tumor, which appear hyper-intense in US images, and was removed across the three
stages.

275

280

285

Table I. Details of intra-modality landmarks (Set 1 & 2) for each patient, with the information for Landmark Set 2 are shown in **bold fonts**. The number of landmarks and mean initial Euclidean distances between landmark pairs are shown, and the range (min ~ max) of the distances is shown in parenthesis after the mean value.

Patient ID	# of landmarks US vs.US	Mean initial distance (range) – in mm before vs. during	Mean initial distance (range)– in mm before vs. after
1	13 / 34	2.31 (1.49~3.29) / 2.32 (1.49~3.29)	5.80 (3.62~7.22)
2	10 / 16	3.31 (1.90~5.19) / 3.10 (1.79~5.19)	3.65 (1.71~6.72)
3	11 / 17	2.09 (1.44~3.02) / 1.93 (0.67~3.02)	2.91 (1.53~4.30)
4	12 / 19	3.67 (3.03~4.76) / 4.00 (3.03~5.22)	2.22 (1.25~2.94)
6	11 / 21	4.98 (2.60~7.18) / 5.19 (2.60~7.18)	2.12 (0.75~3.82)
7	18 / 22	4.51 (0.94~8.16) / 4.69 (0.94~8.16)	3.62 (1.19~5.93)
12	11 / 24	2.99 (1.74~4.81) / 3.39 (1.74~4.81)	3.97 (2.58~6.35)
14	17 / 22	0.72 (0.42~1.59) / 0.71 (0.42~1.59)	0.63 (0.17~1.76)
15	15 / 21	1.99 (0.85~2.84) / 2.04 (0.85~2.84)	1.63 (0.62~2.69)
16	17 / 19	3.24 (1.22~4.53) / 3.19 (1.22~4.53)	3.13 (0.82~5.41)
17	11 / 17	6.46 (4.65~8.07) / 6.32 (4.65~8.07)	5.71 (4.25~8.03)
18	13 / 23	4.47 (1.55~7.01) / 5.06 (1.55~7.44)	5.29 (2.94~9.26)
19	13 / 21	2.44 (1.44~3.40) / 2.06 (0.42~3.40)	2.05 (0.43~3.24)
21	9 / 18	5.28 (4.73~5.60) / 5.10 (3.37~5.94)	3.35 (2.34~5.64)
24	14 / 21	1.82 (1.16~2.65) / 1.76 (1.16~2.65)	2.61 (1.96~3.41)
25	12 / 20	3.63 (2.28~5.02) / 3.60 (2.19~5.02)	7.61 (6.40~10.25)
27	12 / 16	4.90 (3.61~7.05) / 4.93 (3.61~7.01)	3.98 (3.09~4.82)
mean±sd	12.9±2.6 / 20.0±4.8	3.46±1.50 / 3.49±1.55	3.55±1.76

295

For the second type, there are two sets of pair-wise landmarks: pre-operative MRI vs. pre-resection US volume (Set 3), and pre-operative MRI vs. post-resection US volume (Set 4). Here, we used the T2 FLAIR MRI to identify the landmarks since the contrast of the tumor against the normal tissue is better in T2 FLAIR compared to the T1w MRI. The landmarks selected within the MRI are used as references to tag the corresponding points in the US images. Unlike the US vs. US cases, where corresponding landmarks were chosen across resection stages, the MRI vs. US landmarks' locations may differ between the two sets for each patient. Eligible landmarks include the edge of the tumor (MRI vs. pre-resection US only), deep grooves of sulci, corners of sulci, convex points of gyri, and lateral ventricle horns. Again, due to the availability of anatomical features in the US volumes, we have identified landmarks of Set 3 for 22 patients (15~16

305

landmarks per patient), and landmarks of Set 4 for 20 patients (9~16 landmarks per patient). As image features of MRI and US are very different, corresponding landmarks are more difficult to identify than intra-modality cases. As tumor resection progresses (i.e., in Set 4), more and more reliable anatomical landmarks were lost. We did not provide the landmark set for the patients who did not have sufficient number of landmarks. A demonstration of Landmark Set 3 and Set 4 is shown in *Fig. 4*. The number of landmarks, and the mean initial Euclidean distance between landmark pairs are included in Table II.

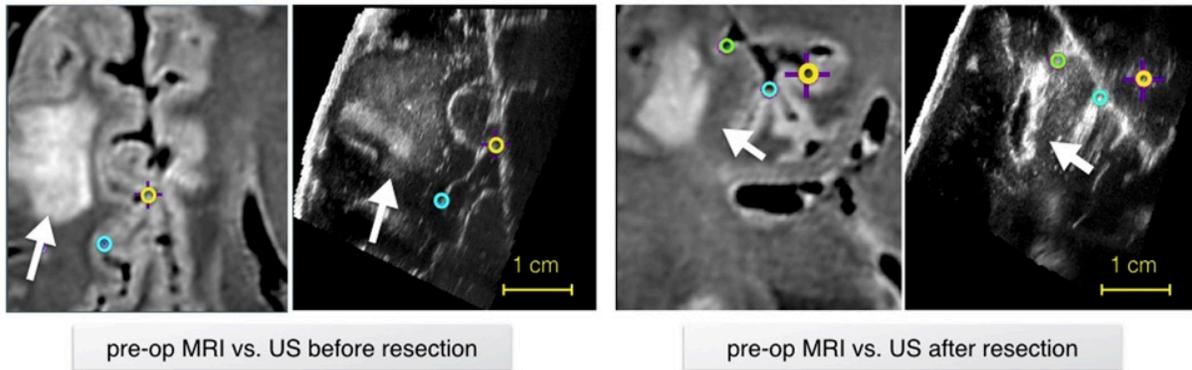


FIG. 4. Demonstration of a corresponding landmarks in Landmark Set 3 (pre-operative MRI vs. US before resection) and Set 4 (pre-operative MRI vs. US after resection). In each pair, the MRI and the US volume are shown on the left and right, respectively. The corresponding landmarks are marked as circles of the same colors, and the white arrows point to the tumor. Note that for the US image after resection, the tumor was removed and a cavity was left as shown in the image.

325

Table II. Details of inter-modality landmarks (Set 3 & 4) for each patient. The number of landmarks and mean initial Euclidean distances between landmark pairs are shown, and the range (min ~ max) of the distances is shown in parenthesis after the mean value.

Patient ID	# of landmarks MRI vs. before US	Mean initial distance (range) in mm MRI vs. before US	# of landmarks MRI vs. after US	Mean initial distance (range) in mm MRI vs. after US
1	15	1.82 (0.56~3.84)	12	3.56 (1.07~9.65)
2	15	5.68 (3.43~8.99)	14	2.15 (0.66~5.78)
3	15	9.58 (8.57~10.34)	15	7.44 (5.10~10.44)
4	15	2.99 (1.61~4.55)	14	3.97 (1.78~5.14)
5	15	12.02 (10.08~14.18)	NA	NA
6	15	3.27 (2.27~4.26)	9	2.97 (1.18~6.33)
7	15	1.82 (0.22~3.63)	15	4.86 (3.06~6.98)
8	15	2.63 (1.00~4.15)	11	1.74 (1.13~2.91)
11	NA	NA	15	3.02 (1.43~7.30)
12	16	19.68 (18.53~21.30)	15	19.31 (17.42~22.69)
13	15	4.57 (2.73~7.52)	NA	NA
14	15	3.03 (1.99~4.43)	16	2.91 (1.50~5.15)
15	15	3.21 (1.15~5.90)	15	4.94 (2.62~7.94)
16	15	3.39 (1.68~4.47)	14	4.02 (2.01~6.99)
17	16	6.39 (4.46~7.83)	16	7.39 (5.58~9.69)
18	16	3.56 (1.44~5.47)	14	8.91 (7.86~10.53)
19	16	3.28 (1.30~5.42)	15	2.80 (1.36~5.22)
21	16	4.55 (3.44~6.17)	15	5.60 (3.34~6.92)
23	15	7.01 (5.26~8.26)	15	7.84 (3.43~9.51)
24	16	1.10 (0.45~2.04)	15	2.46 (0.41~4.29)
25	15	10.06 (7.10~15.12)	12	15.89 (13.95~20.81)
26	16	2.83 (1.60~4.40)	NA	NA
27	16	5.76 (4.84~7.14)	14	6.54 (5.25~8.32)
mean±sd	15.4±0.5	5.37±4.27	14.0±1.8	5.92±4.54

II.F. Evaluation of landmarks

To ensure the quality of the landmarks, they were repeatedly tagged by two experienced raters in human anatomy and medical imaging (authors YX and MF as Rater 1 and 2, respectively), and the inter- and intra-rater variabilities were assessed. The detailed procedures are as follows. First, Rater 1 defined the landmarks in the pre-resection US and pre-operative MRIs as the reference landmarks. Then, Rater 1 and Rater 2 proceeded to locate and mark the corresponding landmarks independently within other US volumes. While keeping the reference landmarks unchanged, for each patient, the corresponding landmarks were tagged twice by both raters, and a 1~2 weeks

interval was ensured between the repetitions. Lastly, the final landmarks in the RESECT database are provided as the averaged results of two trials of landmark marked by both raters (four 3D points for each landmark).

345 To evaluate the intra- and inter-rater variability, we used the mean Euclidean distance between two sets of corresponding landmark points for each patient. More specifically, for intra-rater variability, we computed the metric between two trials of landmark picking for each rater; for inter-rater variability, the average of two trials for each rater was first obtained, and used to compute the metric between two raters. The intra- and inter-rater variability evaluations are presented in Table III. As for intra-modality landmarks between before and during resection (“before US vs. during US”), since Landmark Set 2 contains the same and more landmarks than Landmark Set 1, we used Set 2 for evaluation. The intra-rater variability of Rater 1 is higher than that of Rater 2 for the cases of “Before vs. during US”, “MRI vs. before US”, and “MRI vs. after US” by two-tailed two-sample t-tests ($p < 0.05$). However, the mean values of the distances are all 355 below 0.5mm, which is half a voxel of most MRI data.

Table III. Inter- and Intra- rater evaluations with mean Euclidean distance between landmark sets, and the results are shown as mean±standard deviation. Note that Landmark Set 2 was used to perform the evaluation for the case of “before US vs. during US”.

Type	Intra-rater Rater 1	Intra-rater Rater 2	Inter-rater R1 vs. R2
Before US vs. during US	0.39±0.11 mm	0.31±0.07 mm	0.27±0.05 mm
Before US vs. after US	0.37±0.09 mm	0.34±0.10 mm	0.27±0.11 mm
MRI vs. before US	0.47±0.10 mm	0.33±0.06 mm	0.33±0.08 mm
MRI vs. after US	0.46±0.11 mm	0.34±0.09 mm	0.30±0.07 mm

360

III. Data format and access

For each patient, the T1w MRI is registered to the T2-FLAIR MRI with rigid registration, and for anonymization, facial features within the MRIs and personal information in the MRI image headers have been removed. Except the procedures mentioned, no additional image processing steps, such as image intensity inhomogeneity correction and standardization, were performed on the MR images. All MRIs and 3D ultrasound volumes were distributed in MINC2 and NIFTI 365 formats. The matching landmark pairs between images (US vs. US & MRI vs. US) were recorded

using MNI tag files. While NIFTI format images can be viewed with many medical image visualization packages, such as 3DSlicer, ITKSNAP (www.itksnap.org), and FSL view (<http://fsl.fmrib.ox.ac.uk/fsl/fslview>), the MINC format images can be opened with MINC toolkit (<http://bic-mni.github.io>). Lastly, the complete set of imaging data and landmarks are made freely available to the public as the RESECT database at <https://archive.norstore.no> (search for keyword “RESECT”, dataset doi: 10.11582/2016.00003).

IV. Potential impact

375 We have curated a collection of MR and intra-operative US images using more concurrent imaging technology from 23 low-grade gliomas patients, and provided the images with homologous landmarks for validating and comparing image registration algorithms. More specifically, landmark Set 1 & 2 can provide intra-modality registration validation for US images, and thus enable continuous tracking of soft tissue shift and resection to offer accurate clinical information during surgeries. Landmark Set 3 & 4 can be used to validate inter-modality registration algorithms between MRI and US images. On one hand, landmark Set 3 can be used to study the impact of craniotomy on the initial tissue shift, which often sets the tone for rest of the intervention. On the other hand, by comparing pre-operative MRIs and post-resection US, landmark Set 4 can be used to test nonlinear registration algorithms that can directly map pre-surgical plan to the deformed 385 tissue in order to verify if the resection is complete.

Besides image registration, the imaging data can also be used to develop other image processing methods, such as denoising and segmentation. Furthermore, as medical image visualization²⁵ is also an important component of modern image-guided interventions, the dataset can be used to retrospectively investigate different visualization strategies to allow a better understanding of the anatomy for planning and performing tumor resection. In the future, we will further enrich the existing database with expert tumor segmentation and post-operative data, and include more 390 patients.

The open access to real clinical databases with expert annotations is crucial for the development, evaluation, and comparison of different image processing techniques. While online databases of brain images for studying normal ageing and neuro-degenerative diseases are becoming more and more mature, those to help investigate neuro-surgical procedures are still rare. We hope that the 395

RESECT database will serve as a common ground for the technical and clinical communities to promote the development of image registration algorithms and other image processing methods to advance the surgical treatment of brain cancer.

400

Acknowledgements

This project was partly funded by NSERC Discovery Grant (RGPIN-2015-04136) and the Norwegian National Advisory Unit for Ultrasound and image-guided therapy.

405 *The authors have no relevant conflicts of interest to disclose.*

References

- 410 ¹ E.C. Holland, "Progenitor cells and glioma formation," *Curr Opin Neurol* **14**, 683-688 (2001).
- ² D.N. Louis, H. Ohgaki, O.D. Wiestler, W.K. Cavenee, P.C. Burger, A. Jouvet, B.W. Scheithauer, P. Kleihues, "The 2007 WHO classification of tumours of the central nervous system," *Acta Neuropathol* **114**, 97-109 (2007).
- 415 ³ T.A. Dolecek, J.M. Propp, N.E. Stroup, C. Kruchko, "CBTRUS statistical report: primary brain and central nervous system tumors diagnosed in the United States in 2005-2009," *Neuro Oncol* **14 Suppl 5**, v1-49 (2012).
- ⁴ D.A. Schomas, N.N. Laack, R.D. Rao, F.B. Meyer, E.G. Shaw, B.P. O'Neill, C. Giannini, P.D. Brown, "Intracranial low-grade gliomas in adults: 30-year experience with long-term follow-up at Mayo Clinic," *Neuro Oncol* **11**, 437-445 (2009).
- 420 ⁵ A.S. Jakola, K.S. Myrnel, R. Kloster, S.H. Torp, S. Lindal, G. Unsgard, O. Solheim, "Comparison of a strategy favoring early surgical resection vs a strategy favoring watchful waiting in low-grade gliomas," *JAMA* **308**, 1881-1888 (2012).
- ⁶ R.D. Bucholz, D.D. Yeh, J. Trobaugh, L.L. McDurmont, C.D. Sturm, C. Baumann, J.M. Henderson, A. Levy, P. Kessman, "The correction of stereotactic inaccuracy caused by
- 425

brain shift using an intraoperative ultrasound device," *Lect Notes Comput Sc* **1205**, 459-466 (1997).

7 I.J. Gerard, M. Kersten-Oertel, K. Petrecca, D. Sirhan, J.A. Hall, D.L. Collins, "Brain shift in neuronavigation of brain tumors: A review," *Med Image Anal* **35**, 403-420 (2016).

430 8 C. Nimsky, O. Ganslandt, S. Cerny, P. Hastreiter, G. Greiner, R. Fahlbusch, "Quantification of, visualization of, and compensation for brain shift using intraoperative magnetic resonance imaging," *Neurosurgery* **47**, 1070-1079; discussion 1079-1080 (2000).

9 I.A. Rasmussen, F. Lindseth, O.M. Rygh, E.M. Berntsen, T. Selbekk, J. Xu, T.A.N. Hernes, E. Harg, A. Haberg, G. Unsgaard, "Functional neuronavigation combined with intra-
435 operative 3D ultrasound: Initial experiences during surgical resections close to eloquent brain areas and future directions in automatic brain shift compensation of preoperative data," *Acta Neurochir* **149**, 365-378 (2007).

10 D.L. Hill, P.G. Batchelor, M. Holden, D.J. Hawkes, "Medical image registration," *Phys Med Biol* **46**, R1-45 (2001).

440 11 A. Sotiras, C. Davatzikos, N. Paragios, "Deformable medical image registration: a survey," *IEEE Trans Med Imaging* **32**, 1153-1190 (2013).

12 J.A. Schnabel, M.P. Heinrich, B.W. Papiez, S.J. Brady, "Advances and challenges in deformable image registration: From image fusion to complex motion modelling," *Med Image Anal* **33**, 145-148 (2016).

445 13 J.M. Fitzpatrick, J.B. West, C.R. Maurer, Jr., "Predicting error in rigid-body point-based registration," *IEEE Trans Med Imaging* **17**, 694-702 (1998).

14 D.W. Shattuck, M. Mirza, V. Adisetiyo, C. Hojatkashani, G. Salamon, K.L. Narr, R.A. Poldrack, R.M. Bilder, A.W. Toga, "Construction of a 3D probabilistic atlas of human cortical structures," *Neuroimage* **39**, 1064-1080 (2008).

450 15 S.G. Mueller, M.W. Weiner, L.J. Thal, R.C. Petersen, C. Jack, W. Jagust, J.Q. Trojanowski, A.W. Toga, L. Beckett, "The Alzheimer's disease neuroimaging initiative," *Neuroimaging Clin N Am* **15**, 869-877, xi-xii (2005).

16 L. Mercier, R.F. Del Maestro, K. Petrecca, D. Araujo, C. Haegelen, D.L. Collins, "Online database of clinical MR and ultrasound images of brain tumors," *Med Phys* **39**, 3253-3261
455 (2012).

- 17 E. Ferrante, V. Fecamp, N. Paragios, "Slice-to-volume deformable registration: efficient one-shot consensus between plane selection and in-plane deformation," *Int J Comput Assist Radiol Surg* **10**, 791-800 (2015).
- 18 H. Rivaz, S.J. Chen, D.L. Collins, "Automatic deformable MR-ultrasound registration for
460 image-guided neurosurgery," *IEEE Trans Med Imaging* **34**, 366-380 (2015).
- 19 H. Rivaz, D.L. Collins, "Near real-time robust non-rigid registration of volumetric
ultrasound images for neurosurgery," *Ultrasound Med Biol* **41**, 574-587 (2015).
- 20 B. Fuerst, W. Wein, M. Muller, N. Navab, "Automatic ultrasound-MRI registration for
neurosurgery using the 2D and 3D LC(2) Metric," *Med Image Anal* **18**, 1312-1319 (2014).
- 465 21 W. Wein, A. Ladikos, B. Fuerst, A. Shah, K. Sharma, N. Navab, "Global registration of
ultrasound to MRI using the LC2 metric for enabling neurosurgical guidance," *Med Image
Comput Comput Assist Interv* **16**, 34-41 (2013).
- 22 A. Gronningsaeter, A. Kleven, S. Ommedal, T.E. Aarseth, T. Lie, F. Lindseth, T. Lango,
G. Unsgard, "SonoWand, an ultrasound-based neuronavigation system," *Neurosurgery* **47**,
470 1373-1379; discussion 1379-1380 (2000).
- 23 N. Gordillo, E. Montseny, P. Sobrevilla, "State of the art survey on MRI brain tumor
segmentation," *Magn Reson Imaging* **31**, 1426-1438 (2013).
- 24 J. Mohan, V. Krishnaveni, Y.H. Guo, "A survey on the magnetic resonance image
denoising methods," *Biomed Signal Proces* **9**, 56-69 (2014).
- 475 25 M. Kersten-Oertel, P. Jannin, D.L. Collins, "DVV: a taxonomy for mixed reality
visualization in image guided surgery," *IEEE Trans Vis Comput Graph* **18**, 332-352 (2012).
- 26 A. Fedorov, R. Beichel, J. Kalpathy-Cramer, J. Finet, J.C. Fillion-Robin, S. Pujol, C. Bauer,
D. Jennings, F. Fennessy, M. Sonka, J. Buatti, S. Aylward, J.V. Miller, S. Pieper, R. Kikinis,
"3D Slicer as an image computing platform for the Quantitative Imaging Network,"
480 *Magnetic Resonance Imaging* **30**, 1323-1341 (2012).
- 27 C. Askeland, O.V. Solberg, J.B. Bakeng, I. Reinertsen, G.A. Tangen, E.F. Hofstad, D.H.
Iversen, C. Vapenstad, T. Selbekk, T. Lango, T.A. Hernes, H. Olav Leira, G. Unsgard, F.
Lindseth, "CustusX: an open-source research platform for image-guided therapy," *Int J
Comput Assist Radiol Surg* **11**, 505-519 (2016).
- 485 28 S. Drouin, A. Kochanowska, M. Kersten-Oertel, I.J. Gerard, R. Zelmann, D. De Nigris, S.
Beriault, T. Arbel, D. Sirhan, A.F. Sadikot, J.A. Hall, D.S. Sinclair, K. Petrecca, R.F.

DelMaestro, D.L. Collins, "IBIS: an OR ready open-source platform for image-guided neurosurgery," Int J Comput Assist Radiol Surg2016).

490