RESEARCH ARTICLE

WILEY

# Fully automated delineation of the optic radiation for surgical planning using clinically feasible sequences

Lee B. Reid[1] | Eloy Martínez-Heras[2] | Jose V. Manjón[3] | Rosalind L. Jeffree[4,5] | Hamish Alexander[4] | Julie Trinder[1] | Elisabeth Solana[2] | Sara Llufriu[2] | Stephen Rose[1] | Marita Prior[4] | Jurgen Fripp[1]

[1]The Australian e-Health Research Centre, CSIRO, Brisbane, Queensland, Australia

[2]Center of Neuroimmunology, Laboratory of Advanced Imaging in Neuroimmunological Diseases (ImaginEM), Hospital Clinic Barcelona, Institut d'Investigacions Biomediques August Pi i Sunyer (IDIBAPS) and Universitat de Barcelona, Barcelona, Spain

[3]Instituto de Aplicaciones de las Tecnologías de la Información y de las Comunicaciones Avanzadas (ITACA), Universitat Politècnica de València, Valencia, Spain

[4]Royal Brisbane and Women's Hospital, Metro North, Queensland, Australia

[5]School of Clinical Medicine, University of Queensland, Herston, Queensland, Australia

**Correspondence**
Lee B. Reid, The Australian e-Health Research Centre, CSIRO, Surgical, Treatment and Rehabilitation Service – STARS, Level 7, 296 Herston Road, Herston, QLD, 4029, Australia.
Email: lee.reid1@uqconnect.edu.au

## Abstract

Quadrantanopia caused by inadvertent severing of Meyer's Loop of the optic radiation is a well-recognised complication of temporal lobectomy for conditions such as epilepsy. Dissection studies indicate that the anterior extent of Meyer's Loop varies considerably between individuals. Quantifying this for individual patients is thus an important step to improve the safety profile of temporal lobectomies. Previous attempts to delineate Meyer's Loop using diffusion MRI tractography have had difficulty estimating its full anterior extent, required manual ROI placement, and/or relied on advanced diffusion sequences that cannot be acquired routinely in most clinics. Here we present CONSULT: a pipeline that can delineate the optic radiation from raw DICOM data in a completely automated way via a combination of robust pre-processing, segmentation, and alignment stages, plus simple improvements that bolster the efficiency and reliability of standard tractography. We tested CONSULT on 696 scans of predominantly healthy participants (539 unique brains), including both advanced acquisitions and simpler acquisitions that could be acquired in clinically acceptable timeframes. Delineations completed without error in 99.4% of the scans. The distance between Meyer's Loop and the temporal pole closely matched both averages and ranges reported in dissection studies for all tested sequences. Median scan-rescan error of this distance was 1 mm. When tested on two participants with considerable pathology, delineations were successful and realistic. Through this, we demonstrate not only how to identify Meyer's Loop with clinically feasible sequences, but also that this can be achieved without fundamental changes to tractography algorithms or complex post-processing methods.

KEYWORDS
diffusion magnetic resonance imaging, epilepsy, Meyer's Loop, optic radiation, temporal lobectomy, tractography

# 1 | INTRODUCTION

The axons of the visual pathway emanate from the retina, partially bifurcate at the optic chiasm, pass on either side of the midbrain, and synapse in the lateral geniculate nucleus (LGN). Subsequent axons emanating from the LGN form the posterior, central and anterior bundles, referred to collectively as the optic radiation, which synapse in the primary visual cortex (Peuskens et al., 2004), and potentially V2 and V3 (Alvarez et al., 2015). The posterior bundle travels almost directly posterior from the LGN, whilst the central bundle first travels more laterally, before following a similar path. The anterior bundle, which carries information from the superior visual quadrant, travels anteriorly, inferiorly, and laterally as it fans (Peuskens et al., 2004). It then curves around the lateral ventricle to travel posteriorly to the visual cortex. The entire optic radiation is critical to normal vision, but the anterior bend of the anterior bundle, often referred to as 'Meyer's Loop', has received ongoing focus in neurosurgical science for two related reasons. Firstly, unlike other parts of the optic radiation, it is frequently severed during anterior temporal lobectomies and selective amygdalohippocampectomies, which are common treatments for epilepsy, resulting in permanent quadrantanopia (Winston, 2013). Such partial blindness can result in a reduced quality of life and independence, particularly by removing patients' legal eligibility to drive (Pathak-Ray et al., 2002; Winston, 2013). Secondly, radiologists can use standard structural images to localise vision-critical gyri in the occipital lobe and often even posterior aspects of the optic radiation itself (which appears as a band; Figure S1). By contrast, the majority of Meyer's Loop is not identifiable on such images.

The primarily investigated morphological measurement for the optic radiation has been the distance between the most anterior point of Meyer's Loop and the temporal pole (ML-TP distance; Winston, 2013). This reflects the fact that temporal lobe resection commonly removes tissue from this zone, and this metric can inform the maximally appropriate posterior extent of tissue excision. Six influential cadaver dissection studies have suggested that this distance varies considerably between patients, ranging between 15 and 34 mm (Choi et al., 2006; Chowdhury & Khan, 2010; Ebeling & Reulen, 1988; Párraga et al., 2012; Peuskens et al., 2004; Rubino et al., 2005). The reported average distances have been 25 mm ($n = 20$; Rubino et al., 2005), 26 mm ($n = 11$; Chowdhury & Khan, 2010), 27 mm ($n = 25$; Ebeling & Reulen, 1988), 28.4 mm ($n = 20$; Párraga et al., 2012), and 31.4 mm ($n = 10$; Choi et al., 2006), giving a weighted mean of 27.2 mm. It is important to remember that it is difficult to delineate the precise anatomical boundaries via the Klinger dissection technique used in these studies, meaning that some of these estimates could be slightly overestimated (Ebeling & Reulen, 1988). Regardless, the range of distances reported implies that measurement of an individual surgical candidate's ML-TP distance is more prudent than relying on population averages.

To achieve maximal clinical utility, a tool that delineates the optic radiation needs to be non-invasive, accurate, reliable, and not require costly resources such as manual input, training, or special hardware. Given its non-invasive nature, diffusion MRI tractography is an obvious option, and several groups have delineated the optic radiation using this technique (Alvarez et al., 2015; Benjamin et al., 2012; Chamberland et al., 2017, 2018; Clatworthy et al., 2010; Dayan, Kreutzer, et al., 2015; Dayan, Munoz, et al., 2015; Dreessen de Gervai et al., 2014; James et al., 2015; Lilja et al., 2014; Lim et al., 2015; Martínez-Heras et al., 2015; Meesters et al., 2017; Nilsson et al., 2007; Sherbondy et al., 2008; Tax et al., 2014; Yamamoto et al., 2005; Yang et al., 2019; Yogarajah et al., 2009). Notably, Chamberland et al. (2018) presented an imaging method that produced arguably the most realistic range of ML-TP distances to date. The success of this method, however, relied on manual ROI placement and a multishell sequence with b-values up to 5,000 s/mm$^2$, potentially impeding widespread clinical adoption. Such clinical-accessibility issues have proven difficult to overcome: to the authors' knowledge, all published methods to date have required manual input (e.g., region-of-interest [ROI] placement and adjustment, skull stripping, or other preprocessing) from trained technicians, and/or advanced MR sequences that rely on MRI hardware or multiband packages that are often unavailable clinically.

Beyond clinical accessibility, perhaps the largest challenge has proven to be achieving reasonable accuracy and reliability in optic radiation delineations. As noted by Chamberland, almost all published works reporting ML-TP distances, regardless as to their automaticity or acquisitions, have tended to at least somewhat underestimate the anterior extent of Meyer's Loop (i.e., overestimate ML-TP distances; Chamberland et al., 2017, 2018a). For example, Kammen et al. (2016) described a sophisticated pipeline that automated all but image preprocessing, alignment, and skull-stripping steps. When applied to diffusion images that would take over an hour to acquire on a typical hospital scanner, this produced visibly convincing tractography but a mean ML-TP distance around 4 mm greater than most dissection studies. Although there have been some promising feasibility studies (Lim et al., 2015), ongoing difficulties have raised the question as to whether well-established tractogram generating algorithms are simply incapable of reliably delineating the optic radiation's complex morphology. A number of groups have improved optic radiation delineations through fundamental changes in the way tractography is interpreted, performed, or filtered after generation (Chamberland et al., 2017; Kammen et al., 2016; Meesters et al., 2017; Sherbondy et al., 2008; Tax et al., 2014). However, it is not settled as to whether such purpose-designed algorithms are fundamentally required for optic radiation delineation, or whether standard tractography could be capable of this task if confounding issues were resolved, such as being able to reliably delineate the typically invisible boundaries of the LGN (Tax et al., 2014).

Herein, we describe an image processing pipeline central to our Connectivity Based Neurosurgical Planning Toolkit (CONSULT), which automatically processes images from their raw form and produces a binary map of the optic radiation suitable for surgical planning. For the clinical reasons discussed earlier, the focus of this pipeline, and its assessment here, is an accurate estimation of the ML-TP distance. CONSULT's primary strength is that it is highly clinically accessible: it processes data from a range of sequences in their raw (DICOM) form,

without requiring any human input, and produces ML-TP distances closely in line with historical dissection studies. We describe its stages and report ML-TP distances and reproducibility for a large sample of predominantly neurologically normal participants. Through this, we demonstrate that adequate estimates of ML-TP distance for surgical planning can be achieved without manual input or specialised sequences that require specialised scanner capabilities and impractical scan times.

## 2 | METHODS

We processed several datasets using the CONSULT pipeline, measuring the ML-TP distance in 539 unique brains. Within this pipeline are a small number of changes made to MRtrix3 (Tournier et al., 2019) that are designed to improve the efficiency of streamline generation but do not alter the underlying principles of the probabilistic tractography algorithm itself. We now describe the data processed, the changes made to MRtrix3, then the pipeline itself.

### 2.1 | Data processed

Table 1 details the scans processed, grouped into eight datasets. The three largest datasets were participants from the Human Connectome Project (HCP; Glasser et al., 2013; Van Essen et al., 2012), two were sourced from the MASSIVE dataset, and the remaining three were data acquired by our team on scanners in clinical facilities. Approval was granted for data collection and acquisition from the relevant ethics committees in Metro North Australia, Hospital Clíníc de Barcelona, Washington University, University Medical Center Utrecht, and CSIRO. All participants gave written informed consent. Scan times may be longer than those listed below when using clinical scanners with lower gradient strengths. That said, listed scan times are for full brain coverage; approximately half this many slices is enough to image the optic radiation.

To assess the ML-TP distance in many participants, we processed the first 500 participants from the HCP Young Adults 1,200 release (99% aged 22–35 years, 40% male). This dataset is named within as 'HCP Multishell-90' (HCP-M90). Due to the large number of scans, pre-processed images were used due to reduce the computational cost to an achievable level.

To assess the reproducibility of CONSULT pipeline on high quality *raw* data, the 'HCP Multishell 60 Repeat' (HCP-M60) dataset was used. This consisted of 40 participants (28% male, aged 22–35 years) from the Young Adults Release that were acquired at two time points. We processed both time points from raw diffusion data, but discarded all but 60 volumes per shell prior to pre-processing (see Table 1). Directions were removed by selecting sub-shells that each had an approximately evenly distributed set of directions using a freely available script (https://bitbucket.csiro.au/projects/CONSULT/repos/dwi_tools/). An equivalent sequence for a Siemens Prisma scanner equipped with multiband would carry an acquisition time of approximately 38 min.

The 'HCP Single Shell' (HCP-S) dataset was processed to assess the reproducibility and performance of the CONSULT pipeline with a lower quality dataset. This dataset consisted of the same participant data as the HCP-M60 dataset, but prior to pre-processing, we removed all but $2 \times b = 0$ volumes and $60 \times b = 3,000$ s/mm$^2$ volumes using the aforementioned script and downsampled this raw diffusion data to 2 mm isotropic. An equivalent sequence can be acquired in ~12 min on a Siemens Prisma scanner without multiband (~6 min with multiband factor of 2).

We also sought to determine CONSULT pipeline's performance on raw data acquired at clinical sites in timeframes that are more clinically acceptable. The 'Hospital-A' dataset consisted of three-shell images from 19 neurologically healthy participants aged 21–45 years (mean: 33.7 years), and is regularly acquired at Hospital Clínic, Barcelona. Scan time for diffusion images and the associated field map was 20 min using multiband. The Hospital-B dataset consisted of three-shell images of five adult neurosurgical candidates (two with temporal lobe epilepsy, two with low grade gliomas, and one with an arteriovenous malformation) acquired at the Royal Brisbane and Women's Hospital campus, using a sequence that is regularly acquired at several Australian hospitals as part of clinical research (TA: 11.5 min with multiband, 23 min without). Results for Hospital-B were assessed qualitatively as genuine deviations from normal morphology could be expected in several cases. Finally, the Hospital-C dataset consisted of a low-resolution two-shell scan of two healthy adults acquired on a Siemens Biograph mMR scanner located on the Royal Brisbane and Women's Hospital Campus (TA: ~17 min without multiband).

Finally, we sought to determine the influence of adding a very high *b*-value shell, given one previously published suggestion that $b = 3,000$ s/mm$^2$ may be insufficient to identify the Meyer's Loop (Chamberland et al., 2018b). For this, we utilised the MASSIVE dataset: a multishell acquisition of a single neurologically-healthy woman using a Philips Achieva scanner (Froeling et al., 2017). MASSIVE's acquisition was complex and so the reader is referred to the original manuscript for further detail. As MASSIVE was acquired over multiple sessions, the signal-drift corrected data was used. As with the HCP datasets, we discarded some directions in each shell before pre-processing so that the input data reflected an acquisition with a more-achievable acquisition time if used in a clinical or scientific setting (see Table 1). Specifically, we split these acquisitions into five unique diffusion series. These series formed the 'MASSIVE-3k' dataset (1,000, 2,000, and 3,000 s/mm$^2$ shells only) and 'MASSIVE-4k' dataset (1,000, 2000, 3000, and 4,000 s/mm$^2$). On a Siemen's Prisma scanner assuming a GRAPPA factor of 2, equivalent acquisitions for the MASSIVE-3k and MASSIVE-4k datasets have respective acquisition times of ~8 and ~16 min using multiband, or 16 and ~35 min if without multiband.

### 2.2 | Computationally efficient tractography

Seeding naively from near the LGN can convincingly generate the anterior bundle but only if left running for a very long time, as

**TABLE 1** Input image properties for the datasets used

| Name | Purpose | Scanner | Participant count | Time points | Coil channels | Diffusion | | Unique directions per shell | | | | | Direction repeats | Distortion correction | Estimated TA with/out multiband (min) | T1 MPRAGE | |
| | | | | | | Data used | Resolution | $b = 0$ | $b = 1,000$ | $b = 2000$ | $b = 3,000$ | $b = 4,000$ | | | | Data used | Resolution |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| HCP-M90 | ML-TP against prior dissection studies in very high-quality data | Connectome Skyra | 500 | 1 | 32 | Pre-proc. | 1.25 mm | 18 | 90 | 90 | 90 | 0 | Included | Rev-P | >60/>60 | Pre-proc. | 1 mm |
| HCP-M60 | ML-TP reproducibility in high quality data | Connectome Skyra | 40 | 2 | 32 | Raw | 1.25 mm | 12 | 60 | 60 | 60 | 0 | Removed | Rev-P | 38/>60 | Pre-proc. | 1 mm |
| HCP-S | ML-TP reproducibility and population norms in lower quality data | Connectome Skyra | 40 | 2 | 32 | Raw | 2 mm | 2 | 0 | 0 | 60 | 0 | Removed | Rev-P | 6/12 | Pre-proc. | 1 mm |
| Hospital-A | ML-TP against prior dissection studies in high quality raw data | Siemens Prisma (Barcelona) | 19 | 1 | 64 | Raw | 1.5 mm | 5 | 30 | 60 | 90 | 0 | Not acquired | Fieldmap | 20/40 | Raw | 0.9 mm |
| Hospital-B | Qualitative assessment in pre-surgical patients collected in a more clinically acceptable timeframe | Siemens Prisma (Brisbane) | 5 | 1 | 64 | Raw | 2 mm | 12 | 20 | 32 | 60 | 0 | Not acquired | Rev-P | 11.5/23 | Raw | 1 mm |
| Hospital-C | Demonstration using hardware and sequences without multiband | Siemens Biograph mMR (PET/MRI) | 2 | 1 | 8 | Raw | 2.5 mm | 12 | 20 | 0 | 60 | 0 | Not acquired | Rev-P | 8.5†/17† | Raw | 1 mm |
| MASSIVE-4K | Impact of additional higher b-value shell | Philips Achieva | 1 | 5 | 8 | Drift corrected | 2.5 mm | 12 | 20 | 32 | 60 | 80 | Not acquired | Rev-P | 16/35 | Raw | 1 mm |
| MASSIVE-3K | Impact of additional higher b-value shell | Philips Achieva | 1 | 5 | 8 | Drift corrected | 2.5 mm | 12 | 20 | 32 | 60 | 0 | Not acquired | Rev-P | 8/16 | Raw | 1 mm |

*Note:* Changes were made to original data to produce the 'HCP Multishell 60' (HCP-M60), 'HCP Single Shell' (HCP-S), and MASSIVE datasets (see text). 'Preprocessed' indicates that pre-processed images were used rather than raw images. Direction Repeats indicates whether any measurement repeats of each diffusion-weighted gradient direction were included in the data. Shell *b*-values are in s/mm$^2$. As not all scanners were available to the study authors, estimated TA is listed for a Siemens Prisma scanner with/without multiband, excepting Hospital-C (†) whose values are in s/mm$^2$. TAs can be further reduced by acquiring only partial brain coverage. The HCP Connectome Skyra is a modified Siemens Skyra scanner with differing performance to the commercially available Skyra model.

Abbreviations: HCP, Human Connectome Project; HCP-M, Human Connectome Project multishell; ML-TP, Meyer's-loop to Temporal Pole distance; Pre-proc, minimally pre-processed data; Rev-P, reverse-phase encoding; TA, acquisition time.

streamlines are far more likely to follow the central and posterior bundles, or aberrant paths, than successfully complete Meyer's Loop. Although constraints may eliminate erroneous streamlines, tractography's processing time remains prohibitive because such constraints rarely prevent most computational time being spent generating streamlines that are ultimately rejected, falsely giving the impression that delineating such complex anatomy is beyond the capabilities of existing tractography algorithms. We have forked MRtrix3 (Tournier et al., 2019) and added three features that can improve efficiency but do not alter the nature of tractography itself. For brevity, detailed reasoning for each of the following features can be found in Supporting Information.

Firstly, we propose a method by which seeding directions can be set in a voxel-wise manner from a unipolar dixel image (I.E. where each voxel contains a histogram of unipolar-direction probabilities) and converts pre-existing tractography to such images. This theoretically allows selective use of voxelwise seeding directions that are expected to be fruitful in generating a tractogram, including in a way previously demonstrated by mesh-based seeding (Reid et al., 2016).

Secondly, we have added an option to tckgen that ensures that streamlines pass through inclusion regions in the correct order; this feature has also been pulled into the official distribution of MRtrix.

Anatomically constrained tractography, an established method, can rewind streamlines when they terminate unsatisfactorily, allowing a re-attempt at completion (Smith et al., 2012). We have enabled 'back-tracking' within MRtrix for more general use by allowing this to run outside of anatomically constrained tractography, and requiring rewinding to explicitly defined ROIs that can be placed before problematic branch points. Here, we conservatively set the maximum number of back-track attempts at two per streamline.

A robust performance benchmark is outside the intended scope of this work but anecdotally we have found that the combination of these features typically reduces CONSULT's tractography wall-clock time by around 5×, with a range of approximately 2–20× for an individual, depending on the brain's anatomical particulars, the computing hardware, and the compilation options used. Source code for this forked version of MRtrix can be found at https://bitbucket.csiro.au/projects/CONSULT/repos/mrtrix3/.

## 2.3 | Structural processing

### 2.3.1 | Template generation

Before processing, we created population probability images for labelling of the temporal lobe and V1/V2 region by affine registering 138 HCP 'Young Adult' participants to MNI space (Fonov et al., 2009). Freesurfer-defined labels from these datasets were transformed into MNI space, and their mean taken to represent population probabilities for the temporal lobe and V1/V2 zones. Of these 138 scans, 54 participants also belonged to the HCP-M90 dataset. No participants within HCP-M60, HCP-S, Hospital, or MASSIVE datasets contributed to such atlas creation.

### 2.3.2 | Basic processing

For all Hospital and MASSIVE datasets, raw T1 scans were N4 bias corrected (Tustison et al., 2010) and skull stripped using HD-BET (Isensee et al., 2019). Global Approximate Block Matching was used to denoise and reduce any motion artefacts, as well as segment these images into grey matter, white matter, and cerebrospinal fluid (Reid et al., 2018; Reid & Pagnozzi, 2018). A non-linear transform from MNI space was also calculated using ANTs SyN (Schwarz et al., 2014) using a multi-atlas method.

Due to the unusual bias-correction requirements for HCP structural data (Glasser et al., 2013), minimally pre-processed T1 data were used for all HCP datasets. To reduce computational overhead these scans were resampled to 1 mm isotropic. These data were then processed identically to the Hospital and MASSIVE datasets, excepting that bias-correction was not (re-)performed. No masks, registrations, or parcellations provided with the HCP data were used for either the structural or diffusion data, except for generation of population templates (see above).

## 2.4 | Diffusion processing

### 2.4.1 | Pre-processing

The HCP-M90 dataset was used in its minimally pre-processed form to ensure an achievable computational cost given the large numbers of scans processed. This pre-processing included correction for $b0$ intensity inhomogeneities, EPI distortion, eddy currents, head motion, gradient non-linearities, as well as reorientation and resampling to 1.25 mm isotropic (Glasser et al., 2013).

Data from all datasets except HCP-M90 were pre-processed by CONSULT from their raw form as follows. Images were denoised using MRtrix's dwidenoise, and motion-affected volumes were removed using a previously described method (Pannek et al., 2012). If a fieldmap was detected (here, only the Hospital-A dataset), a brainmask was calculated from the magnitude phase image using bet2 (Jenkinson et al., 2005) as this is more robust to signal dropout in the temporal lobe, and a fieldmap calculated using fsl_prepare_fieldmap. Alternatively, for reverse-phase encoded scans (all but the Hospital-A dataset), an initial brainmask was calculated using MRTrix's dwi2mask, but dilated to ensure full temporal-lobe coverage, and spatial distortions were calculated using FSL's topup (http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/TOPUP). For all datasets motion, distortion, and eddy-current correction were performed through eddy_cuda 8.0. Intensity inhomogeneities were also corrected using N4 bias correction, the field being estimated from the eddy-corrected mean b0 image. For reverse-phase encoded scans, a final brainmask was then recalculated using dwi2mask and dilation operations.

### 2.4.2 | Processing

After pre-processing, all datasets were processed identically, unless otherwise stated. Tissue response functions were calculated using the

Dhollander algorithm (Dhollander et al., 2016). Fibre orientation dispersion maps were calculated for white matter, grey matter, and cerebrospinal fluid using multishell multitissue contrained spherical deconvolution (multishell data) or Single-Shell 3-Tissue constrained spherical deconvolution (HCP-S dataset; https://3Tissue.github.io). Fractional anisotropy was calculated using MRtrix3, upsampled by a factor of 2 using sync interpolation, and cropped to the final brainmask.

Several ROIs, defined in MNI space, were needed for tractography (see below). To transform these into native diffusion space, a transform from diffusion to MNI space was calculated: the fractional anisotropy image was rigid registered to the participant's T1 image using ANTs and combined with the previously-calculated non-rigid (SyN) transform from T1 space to MNI space.

## 2.5 | Posterior bundle

To aid identification of the anterior bundle, the posterior bundle, and a partial central bundle, were delineated by tracking unidirectionally from the optic tract to V1/V2. Streamlines were generated until Tractogram Bootstrapping (Reid et al., 2020) stability criteria were met (min Dice, 0.95; reliability, 0.95; resolution equal to diffusion spatial resolution; $b_{int}$, $0.001 \times n$; minimum streamline count, 10,000). The V1/V2 region was calculated by transforming the previously described V1/V2 template from MNI into native diffusion space, multiplying by individual participant's grey-matter probability image taken from the FOD maps, thresholding, and binarising. Exclusion masks included the grey-matter segmentation (excepting in primary visual areas) and a mask defined in MNI space similar to that described by Martínez-Heras et al. (2015).

We seeded from the optic tract near the midbrain, which was identified using a convolutional neural network, as we found seeding from the optic chiasm (Kammen et al., 2016) was sometimes impaired by MR artefacts. Details of this network are supplied in Supporting Information. In brief, the T1 and white matter fibre orientation dispersion images were transformed into MNI space and cropped to a stack of six 1 mm-thick axial patches in the centre of the brain. These were provided to this network, which had a UNet-style architecture (Ronneberger et al., 2015) and had been trained with datasets for which the optic tract had been identified using tractography seeded from the optic chiasm. The output was a binary mask of the optic tract as it entered the cerebral hemispheres, which could then be padded and transformed back into native participant space (Figure S3).

## 2.6 | Central and anterior bundles

The remaining partial-central and anterior bundles were delineated by tracking from the LGN to the visual cortex. This was performed in a two-pass procedure. In both instances, exclusion ROIs were those previously described.

### 2.6.1 | First pass: identifying an efficient seed point

The first pass was designed solely to identify a seeding position and direction near the LGN, for efficient delineation of the anterior bundle during the second pass. A seeding region was calculated by binarising tractography of the posterior bundle (Figure 1a,b), multiplying by an ROI slightly anterior to the LGN (transformed from MNI space; Figure 2, top) and dilating once (Figure 1c). This mask was slightly anterior to the true LGN because directly seeding from the LGN can
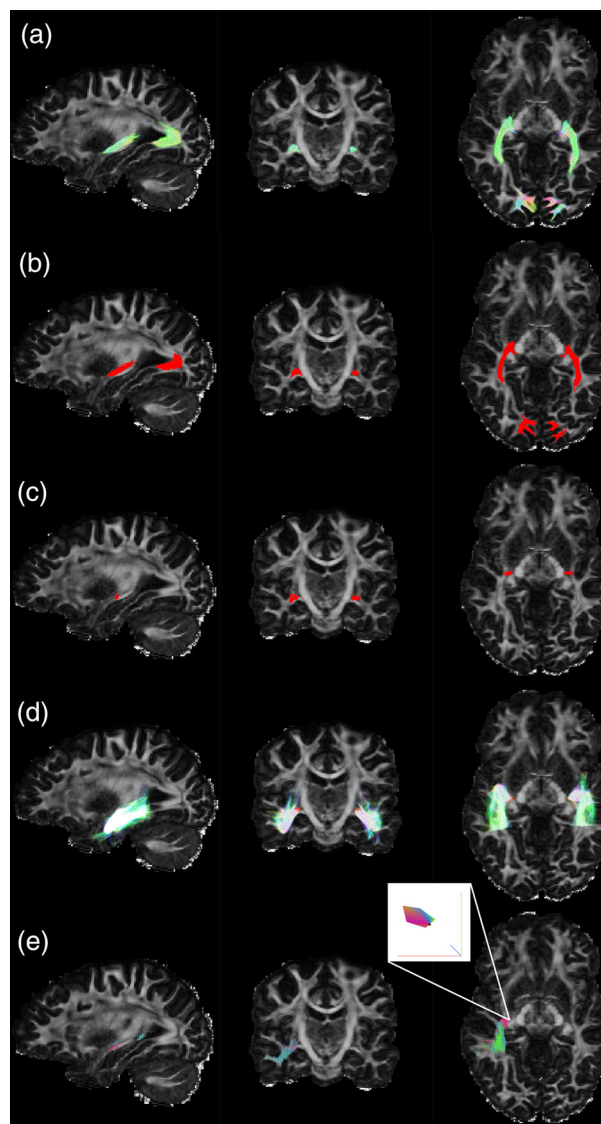


**FIGURE 1** Steps related to the first pass generation of streamlines describing anterior and middle bundles. Tractography of the posterior bundle (a) was converted into a binary ROI (b), then masked to an ROI slightly anterior to the LGN (c). Tractography was then performed (d) and converted into a map of streamline directions (e). The inset in e shows a zoomed view of a single voxel containing a histogram of streamline directions (unipolar dixel). This dixel map was used to improve efficiency of the seeding operation for the second pass of tractography (see Section 3 for examples). Note that tractography is displayed cropped to the shown slices
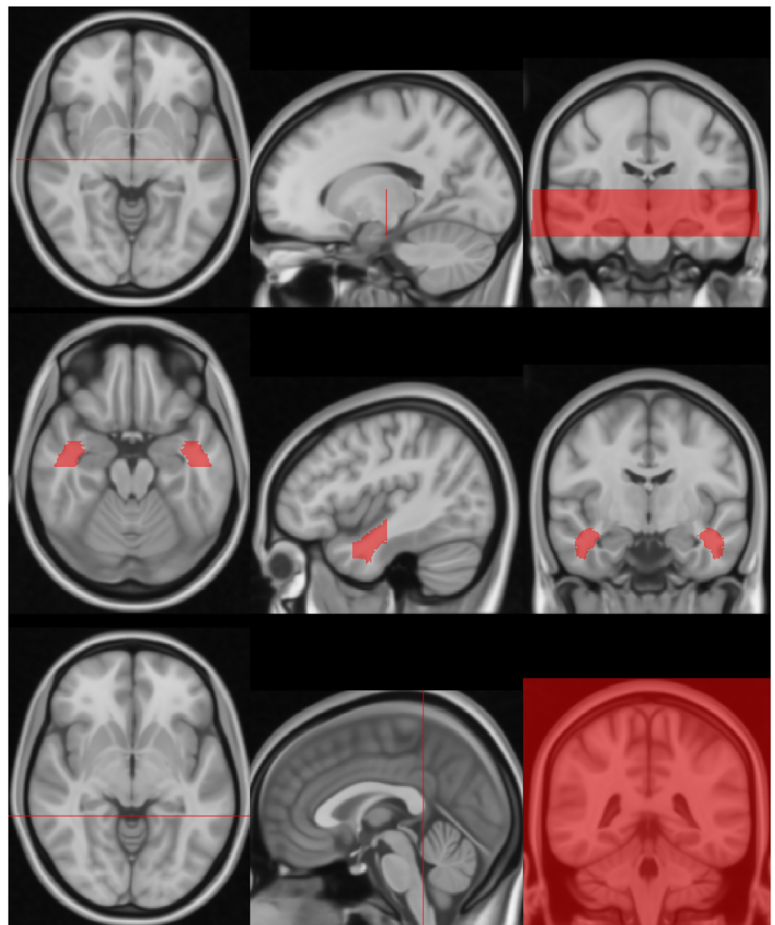
meaningfully slow streamline generation. Tractography was then seeded from this ROI (Figure 1d), the streamline count being set automatically by Tractogram Bootstrapping (Reid et al., 2020) for a reproducible map of streamline directions (I.E. unipolar-dixel trackmap; default settings used and minimum streamline count of 1,000). Three ordered inclusion regions were used: the seed ROI, the temporal lobe posterior and medial to the expected tip of the anterior bundle (Figure 2, middle), and a pre-occipital plane (Figure 2, bottom). To improve efficiency, the second inclusion region was used as a back-tracking ROI. This back-tracking ROI allowed streamlines that entered the second inclusion region but were ultimately rejected to be rewound to approximately 44 mm posterior to the temporal lobe and reattempt completion (up to twice per streamline).

### 2.6.2 | Second pass: tracking the full bundle

The tractogram generated by the first pass was converted into a unipolar dixel image in which each voxel represented a histogram of seeding directions that were successful in delineating Meyer's Loop (Figure 1e). The first-pass tractogram was then discarded. Streamline generation was seeded from the same ROI as the first

pass, but with initial seeding directions sampled from this dixel map. Four ordered inclusion ROIs were used: the same three as the first pass, followed by the visual cortex ROI previously described. Streamlines were generated until Tractogram Bootstrapping (Reid et al., 2020) stability criteria were met for a binary track map (min Dice, 0.95; reliability, 0.95; spatial resolution double that of the diffusion MR; $b_{int}$, 0.001 × $n$; minimum streamline count, 10,000). The final tractogram was converted into a track density image, thresholded at 0.001 × number of streamlines generated and binarised, in line with the Tractogram Bootstrapping criteria.

### 2.7 | Morphological measurement

To enable fair comparisons between datasets in terms of brain volume, binarised tractograms were warped into MNI space (Fonov et al., 2009) for morphological measurements. We defined the 'normalised ML-TP distance' (nML-TP) as the anterior distance between the most anterior voxel of the temporal lobe ($y = 25$ mm) and the most anterior voxel of the anterior bundle, in MNI space.
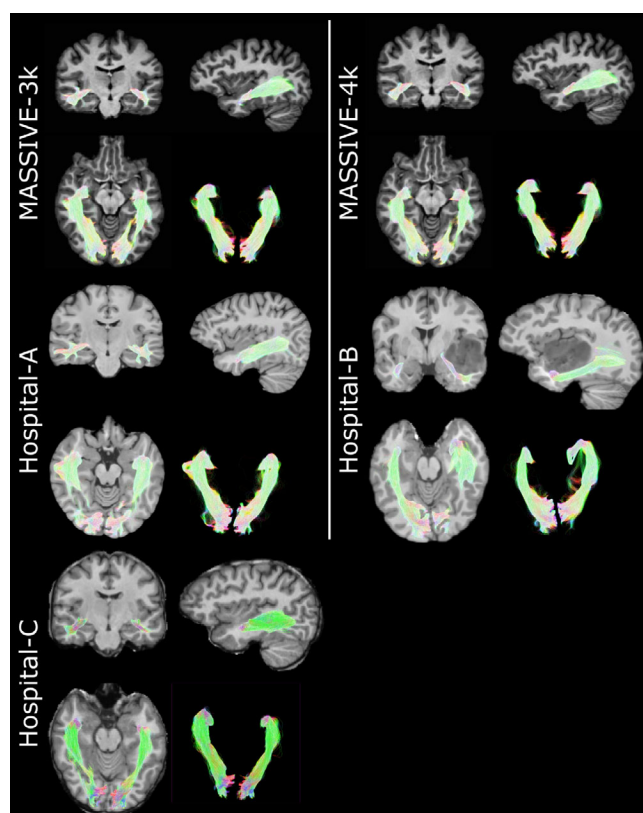
# 3 | RESULTS

## 3.1 | Success rates and normative statistics from the HCP

Example tractography, prior to filtering through binarization, is shown in Figures 3 and 4; mean images of the final results are shown in Figures 5 and 6. Typical per-scan processing times for 2 mm multishell datasets were 3–4 hr on a 20 × 2.6 GHz core GPU-enabled machine, the majority of which was spent on diffusion pre-processing and registration steps. As data were processed on a network, drive better performance can be expected on a dedicated machine. Data processing failed in 3 of 696 datasets (Table 2). Two failures were in HCP-M90 datasets: one due to a failed registration between the T1

and fractional anisotropy image, and another due to failed skull stripping in a participant displaying presumed cutis verticis gyrata. The third failure was an HCP-S dataset for which data processing completed but the nML-TP distance for one hemisphere was implausible (39 mm); this dataset was still included in all statistical analyses. Qualitatively, tractography aligned well with expected anatomy. The middle and posterior sections of the optic radiation aligned well with the corresponding band visible on T1 images, though its likely lateral extent was often slightly overestimated for acquisitions with spatial resolution >1.5 mm (Figure S4). Normalised ML-TP distances are summarised in Table 2. The HCP-M90 dataset showed a median of 27 mm and range (14–33 mm) in line with historical dissection studies (Choi et al., 2006). Despite this wide range, the interquartile range was tight (25–28 mm) and 91% of measurements were within 4 mm
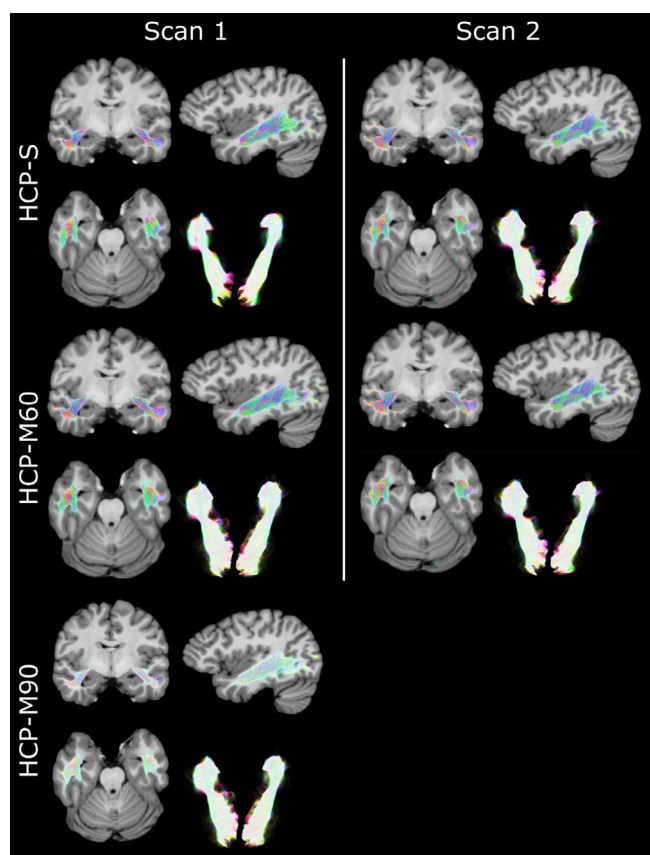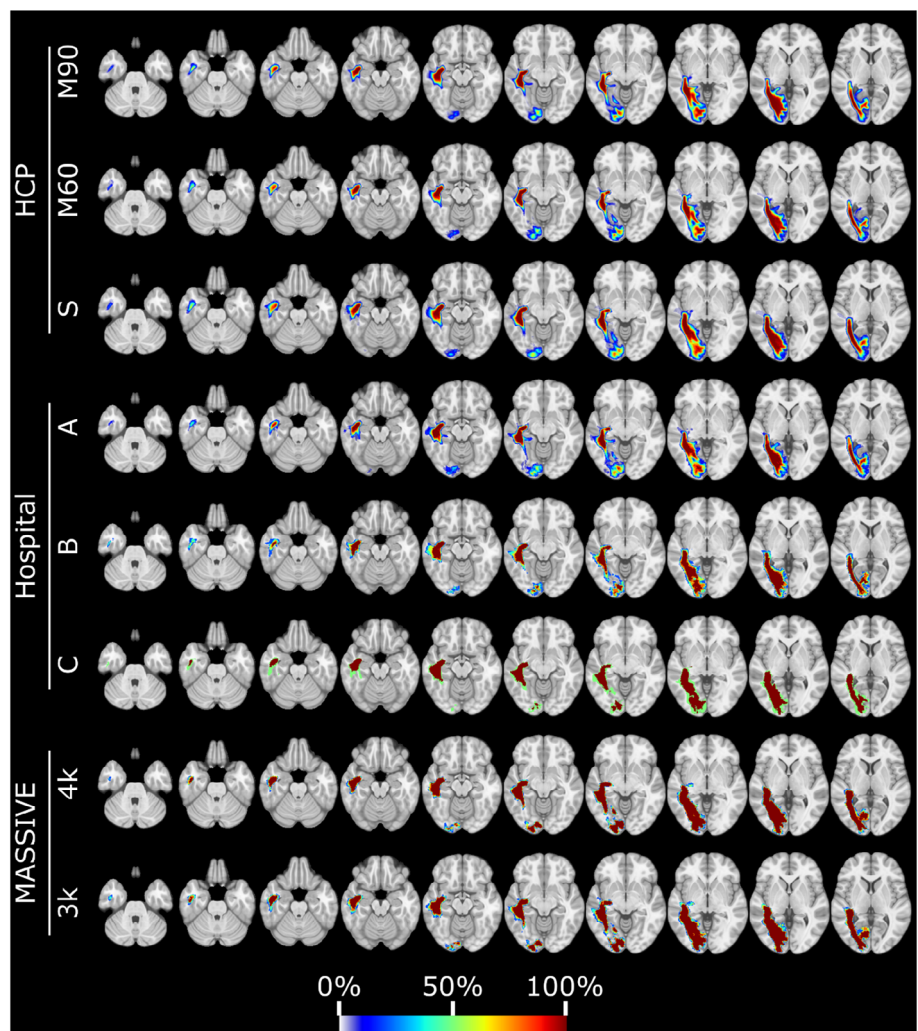


**FIGURE 3** Example Meyer's loop tractography for a single HCP participant in radiological orientation. Tractograms shown here include aberrant streamlines which are largely filtered out in the final rasterised versions that are produced for clinical consumption. This fact and the awkward orientation of the loop relative to standard image orientations can make it difficult to perceive the 3D shape of the delineated structure from singular 2D views, and so we encourage readers to also review Figures 5 and 6. Tractograms overlaid on the T1 image are cropped to 10 mm slabs. The sagittal view shows the left hemisphere. Tractograms shown without an image show all streamlines from an axial view. All images in Figures 3 and 4 have the same degree of transparency applied, which can appear as a whiter tractogram in images with higher numbers of streamlines



**FIGURE 4** Example Meyer's loop tractography for four individual participants in different datasets, in radiological orientation. The MASSIVE datasets show the same participant. Tractograms shown here include aberrant streamlines, which are largely filtered out in the final rasterised versions that are produced for clinical consumption. This fact and the awkward orientation of the loop relative to standard image orientations can make it difficult to perceive the 3D shape of the delineated structure from singular 2D views, and so we encourage readers to also review Figures 5 and 6. Tractograms overlaid on the T1 image are cropped to 10 mm slabs. The sagittal view shows the left hemisphere. Tractograms shown without an image show all streamlines from an axial view. All images in Figures 3 and 4 have the same degree of transparency applied, which can appear as a whiter tractogram in images with higher numbers of streamlines

of the median. Median nML-TP distances did not differ between men and women.

Meyer's Loop typically extended marginally more anterior on the left side than the right (median nML-TP distance difference 1 mm; $p < .001$, one-sample Wilcoxon signed-rank test; Figure 8, middle). Such asymmetry did not differ in left-handed versus right-handed participants (Edinburgh Handedness Inventory score, $< -0.2$ vs. $> 0.2$; medians both 1 mm).

## 3.2 | Reproducibility and dataset comparison

Healthy datasets showed visibly similar results to one another (Figures 5 and 6), as well as numerically similar medians and interquartile ranges for nML-TP distance (Table 2; Figure 8, top). When restricted to identical participants ($n = 28$), median nML-TP distances matched for HCP-M60 and HCP-M90 (both 26 mm; $p = .48$, Wilcoxon Signed Rank Test), but HCP-S showed a greater median nML-TP distance than HCP-M90 (28 vs. 26 mm; $p < .001$, Wilcoxon Signed Rank statistic). Hospital-A's median nML-TP distance was marginally greater than that of HCP-M90 (28 vs. 27 mm; $p < .001$, Mann–

Whitney $U$ Test). For the two Hospital-B scans showing gliomas, CONSULT was able to delineate plausible axonal pathways affected by nearby oedema (Figure 7) and produce nML-TP distances within the normal range for the contralesional hemisphere. Similarly, delineations were numerically and visually normal for both hemispheres of the remaining Hospital-B patients, two healthy participants in Hospital-C and all delineations of the single participant from MASSIVE-3k and MASSIVE-4k. Despite producing similar anterior extents, lower resolution sequences typically resulted in physically-broader delineations than higher resolution sequences (Table 2).

For the HCP-S and HCP-M60 datasets, we processed both scans of each participant that were available. Median absolute scan-rescan error in nML-TP measurements were 1 mm for both datasets (Figure 8, bottom; $p = .15$ Wilcoxon Signed Rank Test). For MASSIVE datasets, we processed five diffusion series of the same participant. The addition of a $b = 4,000$ s/mm$^2$ shell did not alter the mean nML-TP measurement for this participant (27 mm MASSIVE-3k vs. 27 mm MASSIVE-4k; $p = .48$ paired $t$-test). While MASSIVE-4k displayed a marginally tighter range of measurements than MASSIVE-3k, this was not statistically significant ($p = .08$, Levene's Test).
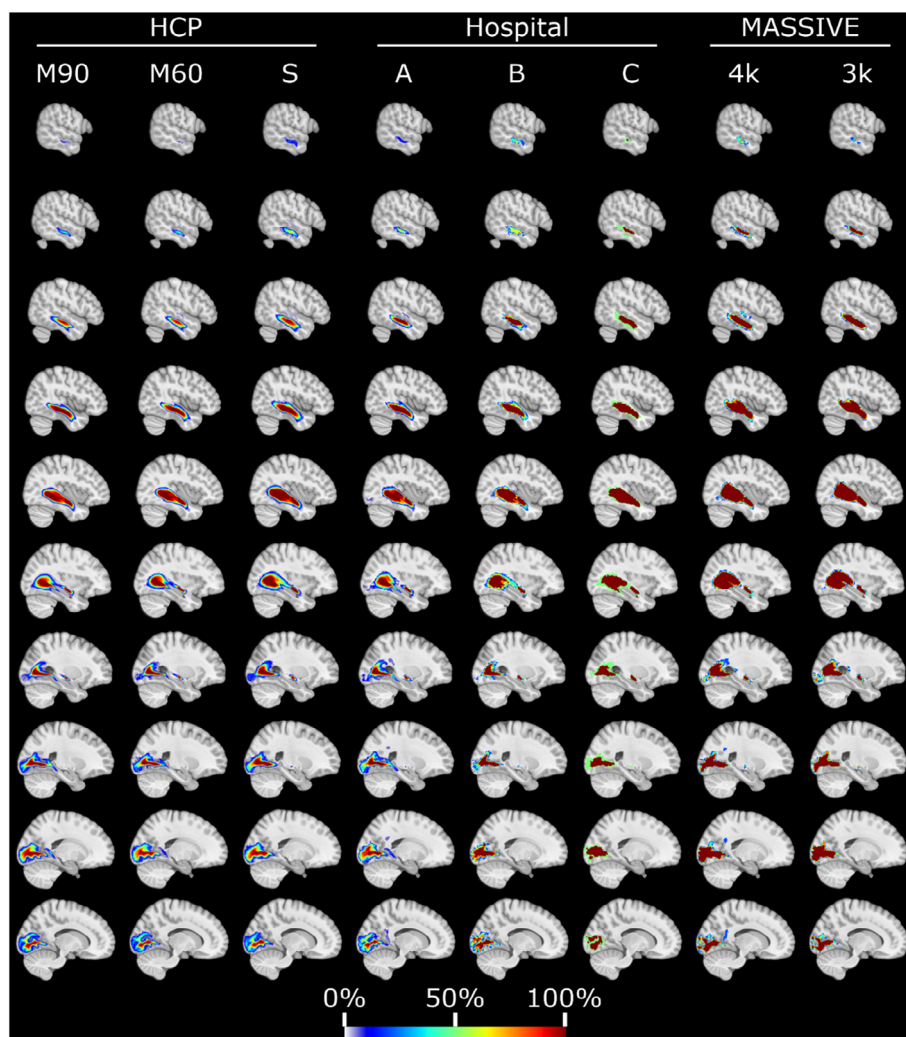
**FIGURE 6** The right anterior bundle for all participants, pooled by dataset into columns, warped into MNI-152 space. Colour indicates the percentage of participants with binarised tractography in that location. Mean morphology is largely the same between all datasets. Hospital-C's simple colouration is because this dataset only comprised two participants. Simple colouration for MASSIVE datasets indicates high agreement between delineations of multiple scans of the same participant

## 4 | DISCUSSION

We have demonstrated a pipeline primarily designed to delineate Meyer's Loop, which is often inadvertently severed during brain surgery. CONSULT distinguishes itself from previously published works particularly in its ability to process raw data in a fully automated way. Here, CONSULT achieved nML-TP distances for a range of acquisitions (medians of 26–28 mm) very similar to average distances reported in four of five dissections studies reporting this distance (25–28.4 mm; Chowdhury & Khan, 2010; Ebeling & Reulen, 1988; Párraga et al., 2012; Rubino et al., 2005) and similar range (14–33 mm, plus one failure at 39 mm) to all five of such dissection studies (15–34 mm; Choi et al., 2006; Chowdhury & Khan, 2010; Ebeling & Reulen, 1988; Párraga et al., 2012; Rubino et al., 2005). By comparison, most previously described MRI pipelines have reported average ML-TP distances larger than those reported in most dissection studies (Chamberland et al., 2018a), implying that their clinical use may underestimate the risk of partial blindness associated with temporal lobectomy. For HCP-M60 and HCP-S, median scan-rescan of reproducibility of nML-TP distances was 1 mm, though in ∼25% of cases was greater than 2 mm, implying that it is prudent to artificially enlarge these

delineations before providing them for clinical use. Our informal testing (not presented here) has suggested that the non-deterministic registration steps account for much of this variability, suggesting that CONSULT's reproducibility may benefit from alternative registration algorithms or utilisation of structural-like images that are guaranteed to be aligned to diffusion space (Beaumont et al., 2021).

### 4.1 | Clinical accessibility of CONSULT

To achieve maximal clinical utility, an imaging tool that delineates the optic radiation needs to minimise its dependency on manual input, training, and special sequences/hardware. Manual input can be a high barrier to clinical adoption as it is time consuming, requires training, and interrupts existing clinical workflows. CONSULT is not the first pipeline published with automated components; most recently-described pipelines contain some degree of automation. For example, Kammen et al. (2016) and Clatworthy et al. (2010) both automated ROI placement, but also both required that supplied diffusion and structural images were already preprocessed, skull-stripped, and aligned to one another: steps that

**TABLE 2** Normalized Meyer's loop to temporal pole distances (nML-TP) and data failure rates for each dataset

| Dataset | Hemisphere | Resolution | No datasets processed | Processing failed | nML-TP (mm) Percentile | | | | | | | Median anterior + central bundle volume (ml) |
| | | | | | Min | 5% | 25% | 50% | 75% | 95% | Max | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| HCP-M90 | Both | 1.25 | 500 | 3 (0.6%) | 14.0 | 21.0 | 25.0 | 27.0 | 28.0 | 30.0 | 33.0 | 21.0 |
| HCP-M60 | Both | 1.25 | 40 | 0 (0%) | 19.0 | 21.0 | 25.0 | 26.0 | 28.0 | 29.0 | 31.0 | 20.0 |
| Hospital-A | Both | 1.5 | 19 | 0 (0%) | 24.0 | 25.0 | 27.0 | 28.0 | 29.0 | 30.0 | 30.0 | 24.1 |
| Hospital-B | Ipsilesional | 2 | 5 | 0 (0%) | 22.0 | 22.6 | 25.0 | 27.0 | 28.0 | 29.6 | 30.0 | 30.0 |
| | Contralesional | 2 | 5 | 0 (0%) | 25.0 | 25.2 | 26.0 | 27.0 | 28.0 | 28.8 | 29.0 | 30.2 |
| Hospital-C | Both | 2.5 | 2 | 0 (0%) | 26.0 | N/A | 26.8 | 27.5 | 28.5 | N/A | 30.0 | 32.7 |
| MASSIVE-4k | Both | 2.5 | 5 | 0 (0%) | 26.0 | 26.0 | 26.0 | 27.0 | 27.0 | 28.0 | 28.0 | 37.0 |
| MASSIVE-3k | Both | 2.5 | 5 | 0 (0%) | 25.0 | 25.0 | 26.0 | 27.0 | 28.0 | 29.0 | 29.0 | 38.1 |
| HCP-S | Both | 2 | 40 | 1 (2.3%) | 20.0 | 23.0 | 26.0 | 28.0 | 29.0 | 31.0 | 39.0 | 25.0 |

*Note:* The 50% percentile refers to the median value. 'Both' hemispheres indicates that both hemispheres' statistics have been pooled. HCP-M60 and HCP-S datasets show values for the first timepoint only. Spatial resolution is displayed in mm and volume in ml.
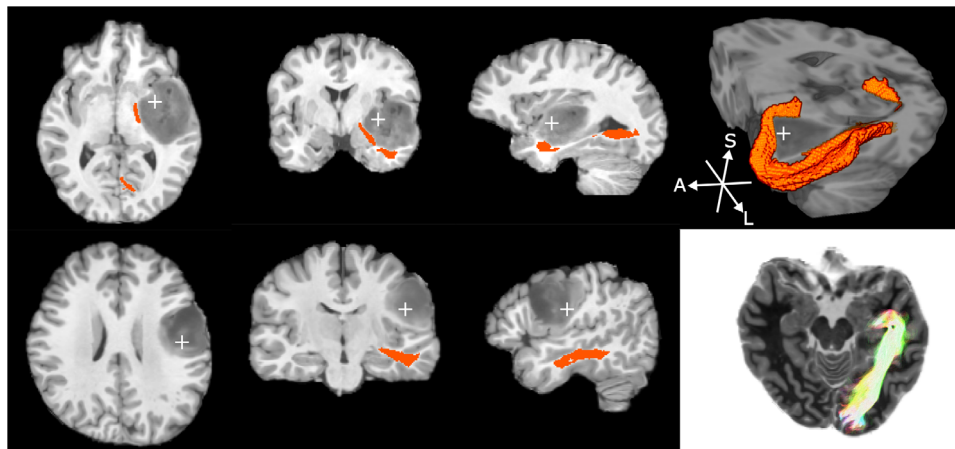


**FIGURE 7** Pathology for two participants with gliomas in the Hospital-B dataset, shown on T1 MPRAGE scans. Colour indicates the delineated anterior bundle on the ipsilesional side. *Top row*: a neurosurgical candidate with pathology in the temporal lobe. The 3D cut-away image (right) shows the delineated Meyer's loop circling the region of oedema. *Bottom row*: a neurosurgical candidate with pathology to the frontal lobe causing some geometric distortion of the temporal lobe. The far-right image on this row is an axial slice with streamlines from all slices visible, and image intensities inverted to ease visualisation

often require manual adjustment and quality control to ensure adequate performance. Other works have presumably automated several of these steps, but still required manual ROI placement. By contrast, CONSULT can process raw images without need for human input at any stage of the process and, as discussed below, has demonstrated an ability to process data acquired with a range of parameters and hardware.

Presentation is also important: visualising streamlines or streamline density images requires meaningful experience because streamline density does not strictly represent tissue density (2010). As discussed in a previous publication (Reid et al., 2020), CONSULT's conversion of streamlines into a binary map for visualisation may be beneficial for clinical decision making because it allows the boundaries to be defined mathematically rather than intuitively. This approach

also allows results to be converted into DICOM format for viewing in standard clinical PACS-enabled viewers, minimising disruption to existing clinical workflows. Commendably, some earlier works converted tractography into binary maps (Clatworthy et al., 2010; Martínez-Heras et al., 2015), though tractography bootstrapping (applied here) carries the advantage of producing maps with better-quantified reproducibility (Reid et al., 2020).

## 4.2 | iFOD2 and crossing fibres

Some previous works have suggested that there is no indication that standard tracking algorithms, such as iFOD2 (Tournier et al., 2010), know how to propagate through the fibre-crossing regions of the
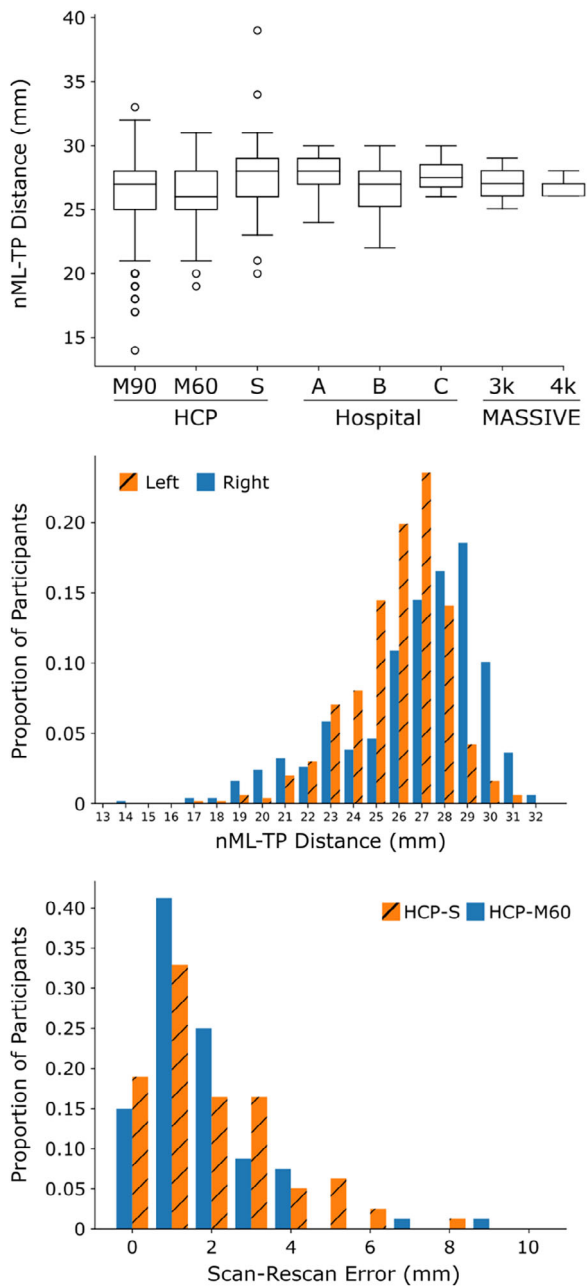
**FIGURE 8** *Top*: Normalised Meyer's loop to temporal pole (nML-TP) distances for all processed datasets. Measurements are pooled from both hemispheres. Sequence details can be found in Table 1. See text for information on statistical significance. *Middle*: nML-TP distances for the left and right hemispheres from the HCP-M90 dataset. *Bottom*: Absolute Scan-Rescan error for the HCP-S and HCP-M60. Note that more than 50% of scans showed error below the resolution of the raw images (2 mm for HCP-S, 1.25 mm for HCP-M60)

optic radiation in standard imaging (Chamberland et al., 2017). Similarly, others have suggested that reasonable delineations require specialised streamline-filtering algorithms or thresholding techniques tuned to specific sequence/hardware combinations (Clatworthy et al., 2010; Meesters et al., 2017; Tax et al., 2014). Although probably

beneficial, we found no evidence that such advances are absolutely required. Beyond a few efficiency boosting techniques, we were able to achieve plausible delineations using standard tractography algorithms, filtered by simply thresholding the resulting streamline density image without sequence-specific parameter tuning (Reid et al., 2020). For clarity, we would like to contrast our approach to a few alternative efficiency-boosting approaches seen in the literature. Rheault et al. (2019) proposed a generic means of constructing and utilising a tractography template to bias tractography to stay within the desired anatomy. Another method, 'MAGNET', was a purpose-built means of biasing streamline paths to follow Meyer's Loop (Chamberland et al., 2017). By contrast, CONSULT's approach was to use apply well-established methods (such as ROI placement) carefully and in a way that adapts to the patient's anatomy, then simply boosting efficiency by ensuring that tractography is seeded in the correct direction and eagerly rejecting or reattempting the generation of 'stray' streamlines. Theoretically, a potential advantage to this simpler method is that it makes fewer assumptions in unusual clinical situations, such as when pathology causes patients' brains to differ substantially from FOD-like templates (such as with our Hospital B dataset). Again, we do not doubt that other approaches may be useful and complementary in this space, but it is encouraging to our clinical partners there is the potential to improve patient care without needing to overturn well-established and widely available technologies.

## 4.3 | Sequence accessibility

Research and clinical groups have different levels of access to different scanning hardware. A primary concern surrounding accessibility of research-grade pipelines is whether acceptable quality images can be obtained in a time frame that is clinically acceptable. We believe that clinics with access to scanners with lower gradient field strengths will be capable of obtaining sufficient quality images in time frames they consider acceptable, though some optimisations may be required. The primary concerns that have previously been raised for the delineation of the optic radiation are requirements for very high *b*-values, high spatial resolution, multiband, and/or large numbers of gradient encoding directions.

*High b-value sequences*

Previously, Chamberland et al. (2018) demonstrated a semi-automated means of delineating Meyer's Loop, and found that its MLTP measurements were similar to dissection studies when provided with a very high quality ($\sim$1.2 mm isotropic, 60 direction @ $b = 1,200, 3,000, 5,000$ s/mm$^2$) imaging sequence, but not when provided with a lower quality sequence (2.4 mm isotropic, 30 directions @ $b = 1,200, 3,000$ s/mm$^2$). If such a high quality sequence was required for all analysis methods this would add significant barriers to achieving reliable delineations for most hospitals, particularly because shells with *b* values of 5,000 s/mm$^2$ require specialised hardware and/or long scan times. However, we were able to produce plausible delineations using a variety of acquisitions collected on five scanners,

without relying on this higher *b*-value shell, multishell data, or high spatial resolution. Critically, CONSULT behaved comparably when provided with the clinically feasible MASSIVE-3k dataset as when provided with the MASSIVE-4k dataset that contained the $b = 4{,}000$ s/mm$^2$ shell and 65% more unique directions. This does not equate exactly to use of a high-resolution $b = 5{,}000$ s/mm$^2$ shell, as used by Chamberland et al. (2018), and the present study was not designed to conclusively calculate the influence of certain gradient field strengths. Nevertheless, results here were very encouraging toward the goal of delineating the optic radiation using clinically-feasible sequences because MASSIVE sequences provided identical averages; if higher *b*-values were absolutely required, *some* level of difference would have been expected. It noteworthy that both the present work and Chamberland et al.'s work are in line with a previous publication (Tournier et al., 2013) demonstrating both that, for the spherical harmonics model, at least 45 directions at $b = 3{,}000$ s/mm$^2$ are required to adequately model fibre orientations, and that increasing *b*-values above 3,000 s/mm$^2$ provides little benefit to such a reconstruction.

### Number of shells and diffusion encoding directions

Multishell diffusion acquisitions can be obtained with virtually any MR scanning hardware and should not be seen as a barrier to implementation. We recommend that multishell, or higher spatial resolution, images are used where possible, partially because of the slight differences observed between our HCP-M90 and HCP-S results and partially because a higher degree of directional redundancy from longer acquisitions affords some protection against motion artefacts. Multishell images also have several advantages for scientific enquiry, such as when calculating more informative metrics than mean diffusivity and FA.

Despite our recommendation, our results indicate that high-resolution multishell images are clearly not a *requirement* to delimiting the optic radiation. Here, the anterior extent of Meyer's Loop, a variable but clinically-important measure for planning temporal lobectomies, differed only slightly between all of the tested acquisitions (Figure 8), including single-shell images (60 direction @ $b = 3{,}000$ s/mm$^2$), all producing averages and ranges that closely matched those of numerous dissection studies. If such images were insufficient, underestimation of the nML-TP distance would have been expected. It is noteworthy that all sequences here contained 60 or more directions in the $b = 3{,}000$ shell to enable tractography. While this means sequences were not the lower *b*-value sequences typically used in clinics for diffusion tensor imaging, many were clinically *feasible* sequences for many clinics in terms of the hardware and scan time required. Highlighting this, the Hospital datasets were acquired on hospital campuses, and the Hospital-C scans were acquired on a scanner with poorer MR capabilities than most clinical scanners.

Although outside the intended scope of the current work, it would be useful for future work to formally quantify the degree to which multishell data provide a benefit to pipelines such as CONSULT, particularly in the context of neurologically impaired patients, who are likely to move during scanning. Here, all multishell sequences contained at least the *b* values of 1,000 and 3,000 s/mm$^2$, and differences with single shelled data were slim. This suggests that a more optimal sequence could use fewer shells, different b-values, and/or fewer directions in the upper shells to reduce the scan time or improve spatial resolution. This could ultimately provide a much more optimal sequence for clinical contexts.

### Resolution, brain coverage, and SNR

We believe that common clinical scanners are capable of obtaining images of sufficient quality for CONSULT. A simple way to significantly improve signal-to-noise and decrease scan time is to lower spatial resolution. Here, three of our tested sequences were acquired at 2.5 mm resolution. This included the MASSIVE dataset that was acquired on a standard clinical scanner and is noted by its authors as having data quality of individual volumes 'comparable to data typically acquired in most clinical studies, which means that subset of the data … would closely resemble a clinical acquisition' (Froeling et al., 2017). These lower-resolution sequences estimated similar Meyer's Loop anterior extents to higher resolution sequences, boding well for routine delineation of the optic radiation using tractography in a pre-surgical setting. The lower resolution sequences also had broader delineations of the anterior loop than the higher resolution images, which can be expected as result of more limited boundary information. Of course, in a surgical context where safety is paramount, labelling the entire Meyer's Loop is more important than eliminating false-positive streamlines and so a mildly higher volume for these low-resolution sequences should not be considered a major drawback.

An alternative approach, where lower resolution is not desirable, is to lower brain coverage. Brain coverage can be lowered to about 50% while still retaining full coverage of the optic radiation. This would halve scan times and is unlikely to have strong repercussions for the reliability of any algorithm presented here. Formal testing of partial brain coverage in CONSULT or other pipelines would be a worthy future direction, given its drastic impact on scan times.

## 4.4 | Limitations

One potential criticism of our work is that, like many others, we focussed on achieving a reasonable and reliable delineation of Meyer's Loop, but focussed less on quantitatively validating the middle and posterior sections of the optic radiation. This focus was because our motivation was surgery-specific: unlike the optic radiation's posterior and middle sections, Meyer's Loop is more typically at risk from common neurosurgical operations that typically remove large amounts of the middle temporal lobe. The appropriate posterior extent of these incisions can be directly informed by the ML-TP distance, which requires specialised imaging techniques such as diffusion MRI. By contrast, as detailed in the Introduction, the medial and posterior sections of the optic radiation can be somewhat identified by more rudimentary means. Furthermore, demonstrating plausible Meyer's Loop delineations in predominantly healthy individuals is a reasonable validation for planning temporal lobectomies because patients with epilepsy are

the typical recipients of such treatments and characteristically have relatively normal appearing brains. Demonstrations in healthy controls for the posterior optic radiation, however, does not imply clinical success given that surgery near these areas is more likely to be for malformations that would cause a pipeline to fail.

A related limitation is that, like many tractography studies, we validated our tool by comparing its measurements to population norms, and by measuring scan-rescan reproducibility of some individual subjects. While these are important steps, dissection studies indicate that some patients genuinely differ from population averages by several millimetres, and so it is important that tools such as CONSULT can demonstrate that measurements for individual subjects are not just reliable and plausible, but also accurate for each individual. This is much more involved work, and is an ongoing project for our group, though some encouragement should come from the fact that both the average and range of morphological measurements found here closely matched several dissection studies, a challenge that the imaging community has only achieved a small number of times (Chamberland et al., 2017, 2018a).

An important limitation with the HCP-S dataset was that its signal-to-noise ratio (SNR) was unlikely to be a close match for clinical acquisitions. This is because HCP-S was generated by downsampling high-resolution volumes acquired on the Skyra Connectome Scanner. A clinical Skyra scanner can be expected to typically obtain images with SNR about one-third lower than the Connectome scanner at $b = 3,000$ s/mm$^2$ (Van Essen et al., 2012), but only for equivalent sequences. Scan-time optimisations are likely to affect this relationship, the most notable being that the HCP-S sequence used multiband, which reduces SNR of individual volumes and is often not available in clinical settings. Here, downsampling was also performed prior to CONSULT's pre-processing, meaning HCP-S images could not benefit properly from denoising, which assumes noise is voxelwise and can be expected to substantially improve SNR when applied to clinical MRIs (Cordero-Grande et al., 2019; Veraart et al., 2016). To counterbalance these uncertainties, we have provided results acquired on other scanners whose matches to common clinical systems have already been discussed. If in doubt, decreasing resolution from 2 to 2.5 mm would be a reasonable decision that would more than account for SNR differences between the Skyra Connectome and other clinical 3T scanner models.

## 4.5 | Conclusion

We have presented a clinically accessible, fully automated pipeline for delineation of the optic radiation. This pipeline plausibly estimated the anterior extents of Meyer's Loop for a range of hardware + sequence combinations, including both multishell and single-shell images that can be acquired in clinically acceptable time frames. We also demonstrated that, excepting some improvements in efficiency, such delineations can be achieved without resorting to complex streamline-selection criteria or fundamental changes to the way in which tractography is performed.

## DATA AVAILABILITY STATEMENT

Code availability: Several key components of the code are available on Bitbucket, where indicated in the manuscript. For reasons of patient safety and to ensure Australian Therapeutic and Goods Administration compliance, the remaining components of the CONSULT pipeline can only be made available through discussions with a CSIRO representative. Contact Jurgen Fripp (jurgen.fripp@csiro.au) for more information.

Data availability: Human Connectome Project data used in this study is available through the Human Connectome Project. To comply with ethics requirements, individual subject data has not been made available. The MASSIVE public dataset can be downloaded from massive-data.org. Study averages that appear as figures can be found as NIfTI files at https://identifiers.org/neurovault.collection:11045 and explored through your browser using NeuroVault''s online viewer (Reid et al., 2021).

## ORCID

Lee B. Reid ⬤ https://orcid.org/0000-0002-3585-8141
Eloy Martínez-Heras ⬤ https://orcid.org/0000-0001-9937-3162
Jose V. Manjón ⬤ https://orcid.org/0000-0001-6640-927X
Rosalind L. Jeffree ⬤ https://orcid.org/0000-0001-7663-2694
Hamish Alexander ⬤ https://orcid.org/0000-0001-5528-8214
Julie Trinder ⬤ https://orcid.org/0000-0002-9871-1796
Elisabeth Solana ⬤ https://orcid.org/0000-0001-7973-2439
Sara Llufriu ⬤ https://orcid.org/0000-0003-4273-9121

*Stephen Rose* https://orcid.org/0000-0003-1558-6297
*Marita Prior* https://orcid.org/0000-0001-7746-0097
*Jurgen Fripp* https://orcid.org/0000-0001-9705-0079

## REFERENCES

Alvarez, I., Schwarzkopf, D. S., & Clark, C. A. (2015). Extrastriate projections in human optic radiation revealed by fMRI-informed tractography. *Brain Structure and Function*, *220*(5), 2519–2532. https://doi.org/10.1007/s00429-014-0799-4

Beaumont, J., Gambarota, G., Prior, M., Fripp, J., & Reid, L. B. (2021). Avoiding data loss: Synthetic MRIs generated from diffusion imaging can replace corrupted structural acquisitions for freesurfer-seeded tractography. *BioRxiv*. https://doi.org/10.1101/2021.02.08.430215

Benjamin, C. F. A., Singh, J. M., Prabhu, S. P., & Warfield, S. K. (2012). Optimization of tractography of the optic radiations. *Human Brain Mapping*, *35*(2), 683–697. https://doi.org/10.1002/hbm.22204

Chamberland, M., Scherrer, B., Prabhu, S. P., Madsen, J., Fortin, D., Whittingstall, K., ... Warfield, S. K. (2017). Active delineation of Meyer's loop using oriented priors through MAGNEtic tractography (MAGNET). *Human Brain Mapping*, *38*(1), 509–527. https://doi.org/10.1002/hbm.23399

Chamberland, M., Tax, C. M. W., & Jones, D. K. (2018). Meyer's loop tractography for image-guided surgery depends on imaging protocol and hardware. *NeuroImage: Clinical*, *20*, 458–465. https://doi.org/10.1016/j.nicl.2018.08.021

Choi, C., Rubino, P. A., Fernandez-Miranda, J. C., Abe, H., & Rhoton, A. L. (2006). Meyer's loop and the optic radiations in the transsylvian approach to the mediobasal temporal lobe. *Neurosurgery*, *59*(4 Suppl 2), ONS235-6. discussion ONS235-6. https://doi.org/10.1227/01.NEU.0000223374.69144.81

Chowdhury, F. H., & Khan, A. H. (2010). Anterior & lateral extension of optic radiation & safety of amygdalohippocampectomy through middle temporal gyrus: A cadaveric study of 11 cerebral hemispheres. *Asian Journal of Neurosurgery*, *5*(1), 78–82.

Clatworthy, P. L., Williams, G. B., Acosta-Cabronero, J., Jones, S. P., Harding, S. G., Johansen-Berg, H., & Baron, J.-C. (2010). Probabilistic tractography of the optic radiations—An automated method and anatomical validation. *NeuroImage*, *49*(3), 2001–2012. https://doi.org/10.1016/j.neuroimage.2009.10.083

Cordero-Grande, L., Christiaens, D., Hutter, J., Price, A. N., & Hajnal, J. V. (2019). Complex diffusion-weighted image estimation via matrix recovery under general noise models. *NeuroImage*, *200*, 391–404. https://doi.org/10.1016/j.neuroimage.2019.06.039

Dayan, M., Kreutzer, S., & Clark, C. A. (2015). Tractography of the optic radiation: A repeatability and reproducibility study. *NMR in Biomedicine*, *28*(4), 423–431. https://doi.org/10.1002/nbm.3266

Dayan, M., Munoz, M., Jentschke, S., Chadwick, M. J., Cooper, J. M., Riney, K., ... Clark, C. A. (2015). Optic radiation structure and anatomy in the normally developing brain determined using diffusion MRI and tractography. *Brain Structure & Function*, *220*(1), 291–306. https://doi.org/10.1007/s00429-013-0655-y

Dhollander, T., Raffelt, D., & Connelly, A. (2016). *Unsupervised 3-tissue response function estimation from single-shell or multi-shell diffusion MR data without a co-registered T1 image*. ISMRM Workshop on Breaking the Barriers of Diffusion MRI, 5.

Dreessen de Gervai, P., Sboto-Frankenstein, U. N., Bolster, R. B., Thind, S., Gruwel, M. L. H., Smith, S. D., & Tomanek, B. (2014). Tractography of Meyer's loop asymmetries. *Epilepsy Research*, *108*(5), 872–882. https://doi.org/10.1016/j.eplepsyres.2014.03.006

Ebeling, U., & Reulen, H. J. (1988). Neurosurgical topography of the optic radiation in the temporal lobe. *Acta Neurochirurgica*, *92*(1–4), 29–36. https://doi.org/10.1007/bf01401969

Fonov, V., Evans, A., McKinstry, R., Almli, C., & Collins, D. (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

Froeling, M., Tax, C. M. W., Vos, S. B., Luijten, P. R., & Leemans, A. (2017). "MASSIVE" brain dataset: Multiple acquisitions for standardization of structural imaging validation and evaluation. *Magnetic Resonance in Medicine*, *77*(5), 1797–1809. https://doi.org/10.1002/mrm.26259

Glasser, M. F., Sotiropoulos, S. N., Wilson, J. A., Coalson, T. S., Fischl, B., Andersson, J. L., ... WU-Minn HCP Consortium. (2013). The minimal preprocessing pipelines for the Human Connectome Project. *NeuroImage*, *80*, 105–124. https://doi.org/10.1016/j.neuroimage.2013.04.127

Isensee, F., Schell, M., Pflueger, I., Brugnara, G., Bonekamp, D., Neuberger, U., ... Kickingereder, P. (2019). Automated brain extraction of multisequence MRI using artificial neural networks. *Human Brain Mapping*, *40*(17), 4952–4964. https://doi.org/10.1002/hbm.24750

James, J. S., Radhakrishnan, A., Thomas, B., Madhusoodanan, M., Kesavadas, C., Abraham, M., ... Vilanilam, G. (2015). Diffusion tensor imaging tractography of Meyer's loop in planning resective surgery for drug-resistant temporal lobe epilepsy. *Epilepsy Research*, *110*, 95–104. https://doi.org/10.1016/j.eplepsyres.2014.11.020

Jenkinson, M., Pechaud, M., & Smith, S. (2005). BET2—MR-based estimation of brain, skull and scalp surfaces. *Eleventh Annual Meeting of the Organization for Human Brain Mapping*, *17*(3), 167.

Kammen, A., Law, M., Tjan, B. S., Toga, A. W., & Shi, Y. (2016). Automated retinofugal visual pathway reconstruction with multi-shell HARDI and FOD-based analysis. *NeuroImage*, *125*, 767–779. https://doi.org/10.1016/j.neuroimage.2015.11.005

Lilja, Y., Ljungberg, M., Starck, G., Malmgren, K., Rydenhag, B., & Nilsson, D. T. (2014). Visualizing Meyer's loop: A comparison of deterministic and probabilistic tractography. *Epilepsy Research*, *108*(3), 481–490. https://doi.org/10.1016/j.eplepsyres.2014.01.017

Lim, J. C., Phal, P. M., Desmond, P. M., Nichols, A. D., Kokkinos, C., Danesh-Meyer, H. V., ... Moffat, B. A. (2015). Probabilistic MRI tractography of the optic radiation using constrained spherical deconvolution: A feasibility study. *PLoS One*, *10*(3), e0118948. https://doi.org/10.1371/journal.pone.0118948

Martínez-Heras, E., Varriano, F., Prčkovska, V., Laredo, C., Andorrà, M., Martínez-Lapiscina, E. H., ... Llufriu, S. (2015). Improved framework for tractography reconstruction of the optic radiation. *PLoS One*, *10*(9), e0137064. https://doi.org/10.1371/journal.pone.0137064

Meesters, S., Ossenblok, P., Wagner, L., Schijns, O., Boon, P., Florack, L., ... Duits, R. (2017). Stability metrics for optic radiation tractography: Towards damage prediction after resective surgery. *Journal of Neuroscience Methods*, *288*, 34–44. https://doi.org/10.1016/j.jneumeth.2017.05.029

Nilsson, D., Starck, G., Ljungberg, M., Ribbelin, S., Jönsson, L., Malmgren, K., & Rydenhag, B. (2007). Intersubject variability in the anterior extent of the optic radiation assessed by tractography. *Epilepsy Research*, *77*(1), 11–16. https://doi.org/10.1016/j.eplepsyres.2007.07.012

Pannek, K., Guzzetta, A., Colditz, P. B., & Rose, S. E. (2012). Diffusion MRI of the neonate brain: Acquisition, processing and analysis techniques. *Pediatric Radiology*, *42*(10), 1169–1182. https://doi.org/10.1007/s00247-012-2427-x

Párraga, R. G., Ribas, G. C., Welling, L. C., Alves, R. V., & De Oliveira, E. (2012). Microsurgical anatomy of the optic radiation and related fibers in 3-dimensional images. *Neurosurgery*, *71*(SUPPL. 1), 160–172. https://doi.org/10.1227/NEU.0b013e3182556fde

Pathak-Ray, V., Ray, A., Walters, R., & Hatfield, R. (2002). Detection of visual field defects in patients after anterior temporal lobectomy for mesial temporal sclerosis-establishing eligibility to drive. *Eye (London, England)*, *16*(6), 744–748. https://doi.org/10.1038/sj.eye.6700152

Peuskens, D., Van Loon, J., Van Calenbergh, F., Van Den Berg, R., Goffin, J., & Plets, C. (2004). Anatomy of the anterior temporal lobe and the frontotemporal region demonstrated by fiber dissection.

*Neurosurgery, 55*(5), 1174–1183. https://doi.org/10.1227/01.NEU.0000140843.62311.24

Reid, L. B., Cespedes, M. I., & Pannek, K. (2020). How many streamlines are required for reliable probabilistic tractography? Solutions for microstructural measurements and neurosurgical planning. *NeuroImage, 211*-(February), 116646. https://doi.org/10.1016/j.neuroimage.2020.116646

Reid, L. B., Cunnington, R., Boyd, R. N., & Rose, S. E. (2016). Surface-Based fMRI-Driven Diffusion Tractography in the Presence of Significant Brain Pathology: A Study Linking Structure and Function in Cerebral Palsy. *PLoS One, 11*(8), e0159540. https://doi.org/10.1371/journal.pone.0159540

Reid, L. B., Gillman, A., Pagnozzi, A. M., Manjón, J. V., & Fripp, J. (2018). MRI denoising and artefact removal using self-organizing maps for fast global block-matching. In W. Bai, G. Sanroma, G. Wu, B. C. Munsell, Y. Zhan, & P. Coupé (Eds.), *Lecture Notes in Computer Science* (Vol. 11075, pp. 20–27). Cham, Switzerland: Springer International Publishing. https://doi.org/10.1007/978-3-030-00500-9

Reid, L. B., & Pagnozzi, A. M. (2018). Rapid training data generation for tissue segmentation using global approximate block-matching with self-organizing maps. In D. Stoyanov, Z. Taylor, G. Carneiro, T. Syeda-Mahmood, A. Martel, L. Maier-Hein, et al. (Eds.), *Lecture Notes in Computer Science* (Vol. 11045, pp. 110–118). Cham, Switzerland: Springer International Publishing. https://doi.org/10.1007/978-3-319-67558-9

Reid, L. B., Martínez-Heras, E., Manjón, J. V., Jeffree, R. L., Alexander, H., Trinder, J., ... Fripp, J. (2021). Population optic radiation maps created by CONSULT. *NeuroVault*. Retrieved from https://identifiers.org/neurovault.collection:11045

Rheault, F., St-Onge, E., Sidhu, J., Maier-Hein, K., Tzourio-Mazoyer, N., Petit, L., & Descoteaux, M. (2019). Bundle-specific tractography with incorporated anatomical and orientational priors. *NeuroImage, 186*, 382–398. https://doi.org/10.1016/j.neuroimage.2018.11.018

Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional networks for biomedical image segmentation. In N. Navab, J. Hornegger, W. M. Wells, & A. F. Frangi (Eds.), *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2015* (pp. 234–241). Cham, Switzerland: Springer International Publishing.

Rubino, P. A., Rhoton, A. L., Tong, X., & De Oliveira, E. (2005). Three-dimensional relationships of the optic radiation. *Neurosurgery, 57*(4 Suppl), 219–227. https://doi.org/10.1227/01.NEU.0000176415.83417.16

Schwarz, C. G., Reid, R. I., Gunter, J. L., Senjem, M. L., Przybelski, S. A., Zuk, S. M., ... Alzheimer's Disease Neuroimaging Initiative. (2014). Improved DTI registration allows voxel-based analysis that outperforms tract-based spatial statistics. *NeuroImage, 94*(3), 65–78. https://doi.org/10.1016/j.neuroimage.2014.03.026

Sherbondy, A. J., Dougherty, R. F., Napel, S., & Wandell, B. A. (2008). Identifying the human optic radiation using diffusion imaging and fiber tractography. *Journal of Vision, 8*(10), 12.1–12.11. https://doi.org/10.1167/8.10.12

Smith, R. E., Tournier, J.-D., Calamante, F., & Connelly, A. (2012). Anatomically-constrained tractography: Improved diffusion MRI streamlines tractography through effective use of anatomical information. *NeuroImage, 62*(3), 1924–1938. https://doi.org/10.1016/j.neuroimage.2012.06.005

Tax, C. M. W., Duits, R., Vilanova, A., ter Haar Romeny, B. M., Hofman, P., Wagner, L., ... Ossenblok, P. (2014). Evaluating contextual processing in diffusion MRI: application to optic radiation reconstruction for epilepsy surgery. *PLoS One, 9*(7), e101524. https://doi.org/10.1371/journal.pone.0101524

Tournier, J.-D., Calamante, F., & Connelly, A. (2013). Determination of the appropriate b value and number of gradient directions for high-angular-resolution diffusion-weighted imaging. *NMR in Biomedicine, 26*(12), 1775–1786. https://doi.org/10.1002/nbm.3017

Tournier, J.-D., Calamante, F., & Connelly, A. (2010). *Improved probabilistic streamlines tractography by 2nd order integration over fibre orientation distributions*. Proceedings of the International Society for Magnetic Resonance in Medicine, 1670.

Tournier, J.-D., Smith, R., Raffelt, D., Tabbara, R., Dhollander, T., Pietsch, M., ... Connelly, A. (2019). MRtrix3: A fast, flexible and open software framework for medical image processing and visualisation. *NeuroImage, 202*, 116137. https://doi.org/10.1016/j.neuroimage.2019.116137

Tustison, N. J., Avants, B. B., Cook, P. A., Zheng, Y., Egan, A., Yushkevich, P. A., & Gee, J. C. (2010). N4ITK: improved N3 bias correction. *IEEE Transactions on Medical Imaging, 29*(6), 1310–1320. https://doi.org/10.1109/TMI.2010.2046908

Van Essen, D. C., Ugurbil, K., Auerbach, E., Barch, D., Behrens, T. E. J., Bucholz, R., ... WU-Minn HCP Consortium. (2012). The Human Connectome Project: A data acquisition perspective. *NeuroImage, 62*(4), 2222–2231. https://doi.org/10.1016/j.neuroimage.2012.02.018

Veraart, J., Novikov, D. S., Christiaens, D., Ades-aron, B., Sijbers, J., & Fieremans, E. (2016). Denoising of diffusion MRI using random matrix theory. *NeuroImage, 142*, 394–406. https://doi.org/10.1016/j.neuroimage.2016.08.016

Winston, G. P. (2013). Epilepsy surgery, vision, and driving: What has surgery taught us and could modern imaging reduce the risk of visual deficits? *Epilepsia, 54*(11), 1877–1888. https://doi.org/10.1111/epi.12372

Yamamoto, T., Yamada, K., Nishimura, T., & Kinoshita, S. (2005). Tractography to depict three layers of visual field trajectories to the calcarine gyri. *American Journal of Ophthalmology, 140*(5), 781–785. https://doi.org/10.1016/j.ajo.2005.05.018

Yang, J. Y.-M., Beare, R., Wu, M. H., Barton, S. M., Malpas, C. B., Yeh, C.-H., ... Seal, M. (2019). Optic radiation tractography in pediatric brain surgery applications: A reliability and agreement assessment of the tractography method. *Frontiers in Neuroscience. 13*, 1254. https://doi.org/10.3389/fnins.2019.01254

Yogarajah, M., Focke, N. K., Bonelli, S., Cercignani, M., Acheson, J., Parker, G. J. M., ... Duncan, J. S. (2009). Defining Meyer's loop-temporal lobe resections, visual field deficits and diffusion tensor tractography. *Brain: A Journal of Neurology, 132*(Pt 6), 1656–1668. https://doi.org/10.1093/brain/awp114

## SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.