bioRxiv preprint doi: https://doi.org/10.1101/2021.12.03.471134; this version posted December 5, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

HippUnfold: Automated hippocampal unfolding, morphometry, and subfield segmentation

Jordan DeKraker[1,2], Roy AM Haast[1,3], Mohamed D Yousif[1], Bradley Karat[1], Stefan Köhler[2,4], Ali R Khan[1,2,5,6]

[1]Robarts Research Institute, Schulich School of Medicine and Dentistry, University of Western Ontario, Canada
[2]Brain and Mind Institute, University of Western Ontario, Canada
[3]Aix-Marseille University, CNRS, CRMBM, UMR 7339, Marseille, France
[4]Dept Psychology, University of Western Ontario, Canada
[5]School of Biomedical Engineering, University of Western Ontario, Canada
[6]Dept Medical Biophysics, University of Western Ontario, Canada

Corresponding author: Jordan DeKraker (jordan.dekraker@mail.mcgill.ca), Ali Khan (alik@robarts.ca)

Key words: hippocampal subfields, MRI, morphology, deep learning, neuroimaging software

Abstract

The archicortical hippocampus differs, like the neocortex, in its folding patterns between individuals. Here, we present an automated and robust BIDS-App, HippUnfold, for defining and indexing subject-specific hippocampal folding in MRI, analogous to popular tools used in neocortical reconstruction. This is critical for inter-individual alignment, with topology as the basis for homology. This topological framework enables qualitatively new analyses of morphological and laminar structure in the hippocampus or hippocampal subfields, and is critical for the advancement of neuroimaging analyses at a meso- or micro-scale. HippUnfold uses state-of-the-art deep learning combined with previously developed topological constraints on hippocampal tissue. It is designed to work with commonly employed sub-millimetric MRI acquisitions, with extensibility to microscopic resolutions as well. In this paper we illustrate the power of HippUnfold in feature extraction, and its construct validity compared to several extant hippocampal subfield analysis methods.

Introduction

Most neurological or psychiatric diseases with widespread effects on the brain show strong and early impact on the hippocampus (e.g. [1]). This highly plastic grey matter (GM) structure is also critical in the fast formation of episodic and spatial memories (e.g. [2]). Examination of this structure with non-invasive neuroimaging, such as MRI, provides great promise for furthering our understanding, diagnosis, and subtyping of these diseases and cognitive processes in the hippocampus and its component subfields [3].

In current neuroimaging analyses the hippocampus is typically modelled as a subcortical volume, but it is actually made up of a folded archicortical mantle, or ‘ribbon’ [4]. Representing

bioRxiv preprint doi: https://doi.org/10.1101/2021.12.03.471134; this version posted December 5, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

the hippocampus as such can be leveraged to enable qualitatively new analyses, such as registration, despite inter-individual differences in gyrification or folding structure, through topological alignment. Additionally, representation as a ribbon allows the hippocampus to be factorized into surface area and thickness, which can be further subdivided for laminar analyses. These methods are thus critical in advancing MRI research from the macroscopic scale to the subfield, cortical column, and laminar scales. Similar approaches have already yielded a paradigm shift in neocortical analysis methods [5,6].

Denoting the hippocampal archicortex or ribbon is challenging because it is thin (0.5-2mm), its folding pattern varies considerably between individuals [7,8], and this folding may even continue to change from early development through adulthood [9]. We present here a set of tools to overcome these challenges using a highly sensitive and generalizable “U-Net’ deep learning architecture [10], combined with previous work that enforces topological constraints on hippocampal tissue [11].

In previous work [11], we developed a method to computationally unfold the hippocampus along its geodesic anterior-posterior (AP) and proximal-distal (PD, i.e., proximal to the neocortex, with the dentate gyrus being most distal) axes. We demonstrated for the first time several qualitative properties using in vivo MRI, such as the contiguity of all subfields along the curvature of the hippocampal head (anterior) and tail (posterior), previously described only in histology. This pioneering work relied heavily on detailed manual tissue segmentations including the high-myelinated stratum radiatum, lacunosum, and moleculaire (SRLM), a commonly used landmark that separates hippocampal folds along the inward ‘curl’ of the hippocampus. In this work we also considered curvature and digitations along the AP axis of the hippocampus, most prominently occurring in the hippocampal head [4,7,8,11]. Each of these features are highly variable between individuals, making them difficult to capture using automated volumetric atlas-based methods and time-consuming to detect manually.

The current work automates the detailed tissue segmentation required for hippocampal unfolding using a state-of-the-art ‘U-Net’ deep convolutional neural network [10]. In particular, we aimed to capture morphological variability between hippocampi which are not seen using existing automated methods which employ either a single atlas or multi-atlas fusion (eg. [12–14]). U-Net architectures have been shown to be generalizable and sensitive to anatomical variations in many medical image processing tasks [15], making them ideal to overcome this challenge.

Estimating hippocampal subfield boundaries in MRI is challenging since their histological hallmarks are not directly available in MRI due to lower spatial resolution and lack of appropriate contrasts, which is an ongoing hurdle in neuroimaging [16,17]. However, post-mortem studies show that the subfields are topologically constrained according to their differentiation from a common flat cortical mantle [4]. Thus a folded representation of hippocampal tissue provides a powerful intermediate between a raw MRI and subfield labels [18], analogous to the reconstruction of a 3D neocortical surface. This surface can then be parcellated into subregions without topological breaks [5], overcoming many limitations of current subfield segmentation methods [17]. Here, we apply surface-based subfield boundary definitions obtained via manual segmentation of BigBrain 3D histology [19] which was additionally supported by a data-driven parcellation [20]. We additionally demonstrate how labels used in the popular Freesurfer [21]

bioRxiv preprint doi: https://doi.org/10.1101/2021.12.03.471134; this version posted December 5, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

and Automatic Segmentation of Hippocampal Subfields (ASHS) [12] software packages can be applied under our topologically-constrained framework.

Altogether, we combine novel U-Net tissue classification, previously developed hippocampal unfolding [11], and topologically-constrained subfield labelling [20] together into a single pipeline which we refer to as 'HippUnfold' hereinafter. We designed this pipeline to employ FAIR principles (findability, accessibility, interoperability, reusability) with support across a wide range of use-cases centered around sub-millimetric MRI.

Results

Data

HippUnfold was designed and trained with the Human Connectome Project (HCP) 1200 young adult subject data release (HCP-YA) [22], and additionally tested on the HCP Aging dataset (HCP-A) [23], and anisotropic (or thick-slice) 7T data (7T-TSE) from [24] which is considered optimal by many hippocampal subfield researchers [17]. These data are summarized briefly in Table 1, and for more details see Online Methods.

Table 1. MRI datasets used in training, evaluation, and comparison to extant methods. Methods employed include those proposed here (HippUnfold), the same processing but with manual segmentation (similar to previous work [20]) (manual unfold), Freesurfer v7.1 [21], and an atlas of manual segmentations [24] used in ASHS [12].

Name	**Modalities**	**Resolution**	**Sample size (L+R)**	**Methods employed**
HCP-YA	T1w, T2w	0.7x0.7x0.7mm	n=590 (training)	HippUnfold Manual unfold
			n=148 (testing)	HippUnfold Manual unfold Freesurfer (v7.1)
HCP-A	T1w, T2w	0.8x0.8x0.8mm	n=1312	HippUnfold
7T-TSE	T2w	0.4x0.4x1.0mm	n=70	HippUnfold Manual subfields (ASHS atlas)

HippUnfold aligns and visualizes data on folded or unfolded surfaces

HippUnfold is presented here as a fully-automated pipeline with outputs including hippocampal tissue and subfield segmentations, geodesic Laplace coordinates spanning over hippocampal GM voxels, and inner, midthickness and outer hippocampal surfaces. These surfaces have corresponding vertices, providing an implicit topological registration between individuals.

The overall pipeline for HippUnfold is illustrated briefly in Figure 1. A comprehensive breakdown of each step is provided in the online Methods.

bioRxiv preprint doi: https://doi.org/10.1101/2021.12.03.471134; this version posted December 5, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

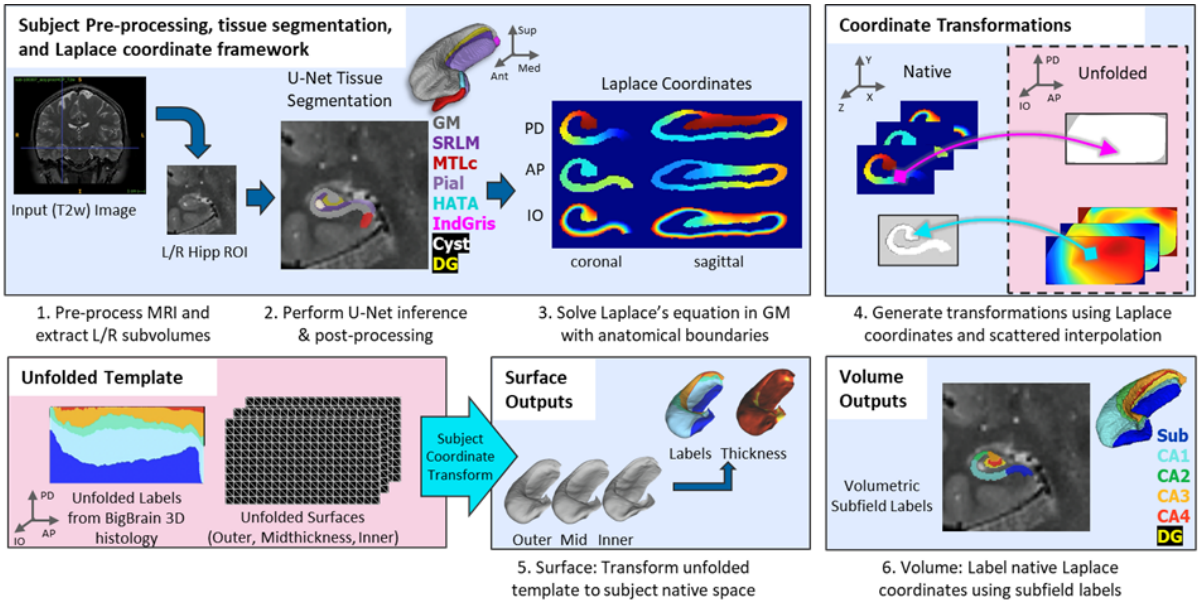


Figure 1. Overview of HippUnfold pipeline. First, input MRI images are preprocessed and cropped around the left and right hippocampi. Second, a U-Net neural network architecture (nnUNet [10]) is used to segment hippocampal grey matter (GM), the high-myelinated stratum radiatum, lacunosum, and moleculare (SRLM), and structures surrounding the hippocampus. Segmentations are post-processed via template shape injection. Third, Laplace's equation is solved across the anterior-posterior (AP), proximal-distal (PD) and inner-outer (IO) extent of hippocampal GM, making up a geodesic coordinate framework. Fourth, scattered interpolants are used to determine equivalent coordinates between native Cartesian space and unfolded space. Fifth, unfolded surfaces with template subfield labels [20] are transformed to subjects' native folded hippocampal configurations. Morphological features (eg. thickness) are extracted using Connectome Workbench [25] on these folded native space surfaces. Sixth, volumetric subfields are generated by filling the voxels between inner and outer surfaces with the corresponding subfield labels. Additional details on this pipeline can be found in the online Methods.

In addition to subfield segmentation, HippUnfold extracts morphological features and can be used to sample quantitative MRI data along a midthickness surface to minimize partial voluming with surrounding structures. This is visualized across n=148 test subjects on an unfolded surface and group-averaged folded surface in Figure 2. Note that the group averaging takes place on a surface and so does not break individual subjects' topologies. Quantitative MRI features examined here include T1w/T2w ratio as a proxy measure for intracortical myelin [26], mean diffusivity, and fractional anisotropy [27,28].

bioRxiv preprint doi: https://doi.org/10.1101/2021.12.03.471134; this version posted December 5, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

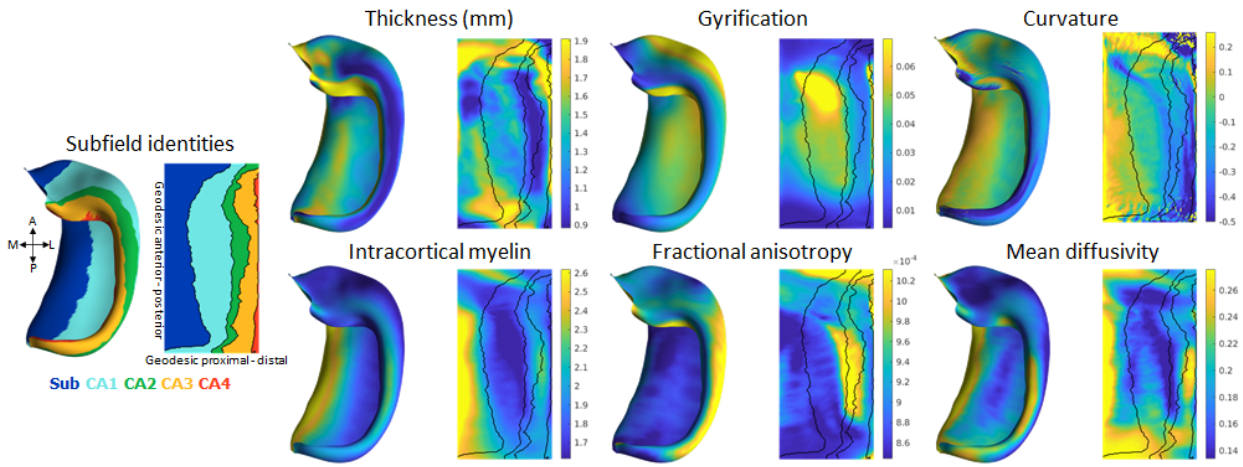


Figure 2. Average hippocampal folded and unfolded surfaces showing subfields, morphometric and quantitative MRI measures from the HCP-YA test dataset. The same topologically defined subfields were applied in unfolded space to all subjects (left), which are also overlaid on quantitative MRI plots (black lines). Note that many morphological and quantitative MRI measures show clear distinctions across subfield boundaries.

Clear differences in morphological and quantitative MRI features can be seen across the hippocampus, particularly across subfields as defined here from a histologically-derived unfolded reference atlas [20]. This highlights the advantages of the present method. These folded and unfolded representations of hippocampal characteristics are broadly in line with previous work examining differences in such morphological and quantitative MRI features across hippocampal subfields or along the hippocampal AP extent (eg. [29,30]). However, in previous work these features differed between predefined subfields on average, but did not necessarily follow subfield contours as seen here. Some advantages of the current pipeline that likely contribute to this clarity include i) the detail of the hippocampal GM segmentation, ii) sampling along a midthickness surface to minimize partial voluming with surrounding structures, and iii) the fact that subjects are topologically aligned leading to less blurring of features after group-averaging.

Extant methods do not respect the topological continuity of hippocampal subfields

Several automatic methods for labelling hippocampal subfields in MRI exist, of which Freesurfer [21] (FS, v7.1) and Automatic Segmentation of Hippocampal Subfields [12] (ASHS) are among the most widely adopted. These methods rely on volumetric registrations between a target hippocampus and a reference or atlas. Specifically, ASHS makes use of multi-atlas registration, wherein multiple gold standard manual hippocampal subfield segmentations are registered to a target sample. Typically the multi-atlas consists of roughly a dozen samples which are then fused together to generate a reliable yet oftentimes smooth or simplified final product. FS uses a combination of voxel-wise classification and, bijectively, volumetric registration between a target hippocampus and a probabilistic reference atlas, which is

bioRxiv preprint doi: https://doi.org/10.1101/2021.12.03.471134; this version posted December 5, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

generated via combined in vivo MRI and 9.4T ex vivo hippocampal subfield segmentations [21]. When hippocampi take on different folding configurations, such registrations can become ill-posed. HippUnfold overcomes these limitations in two ways: with extensive training (in this case n=590), U-Net can capture detailed inter-individual differences in folding and, secondly, our unfolding technique ensures that subfield labelling is topologically constrained [18].

We applied Freesurfer's (v7.1) hippocampal subfields pipeline as well as ASHS using a recent manual subfield multi-atlas [24] to the HCP-YA test set. We then compared resulting subfield segmentations to those generated via HippUnfold in native and unfolded space, which is shown in Figure 3 in one representative subject. For comparison, we additionally mapped FS and ASHS subfield boundaries in folded and unfolded space.

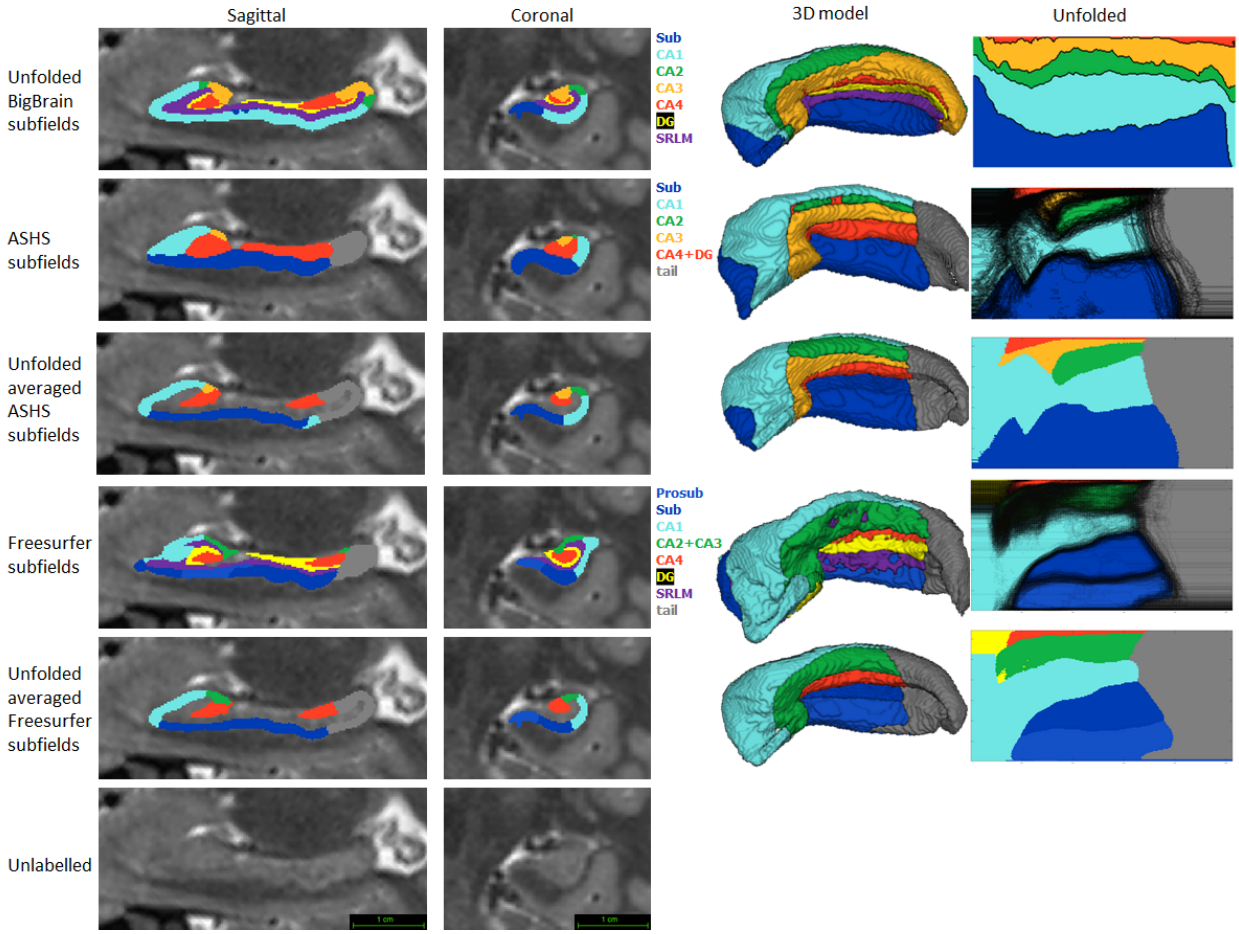


Figure 3. Comparison of HippUnfold, ASHS, and Freesurfer subfield segmentations in native and unfolded space. Sagittal and coronal slices and 3D models are shown for one representative subject. Note that for HippUnfold hippocampal subfields are the same for all individuals in unfolded space, but for ASHS and FS we mapped all subjects' subfield boundaries which are shown in the black lines in column 4 rows 2 and 4. We then took the mode subfield label from ASHS and FS in unfolded space and projected it back to native space, which is shown in rows 3 and 5.

Both ASHS and FS showed subfield discontinuities in unfolded space in at least some subjects, and FS even showed discontinuities in the group-averaged unfolded subfields. That is,

bioRxiv preprint doi: https://doi.org/10.1101/2021.12.03.471134; this version posted December 5, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

some pieces of a given label were separated from the rest of that label. ASHS does not include an SRLM label and the SRLM produced by FS was not consistently aligned with that used in unfolding. Thus, subfields sometimes erroneously crossed the SRLM, breaking topology and explaining why discontinuities were sometimes observed in unfolded space. Ordering of labels was also not consistent in ASHS and FS. For example, sometimes CA1 would border not only CA2 but also CA3, CA4, and/or DG. Additionally, neither ASHS nor FS extends all subfields to the full anterior and posterior extent of the hippocampus. Instead, both methods simplify most of the anterior hippocampus as being CA1 and opt not to label subfields in the posterior hippocampus at all. These qualities are not in line with the anatomical ground truth shown in both classic and contemporary ex-vivo histological studies [4,8], which were indeed well captured by HippUnfold. FS also over-labelled hippocampal tissue, which can be seen reaching laterally into the ventricles in the coronal view. Similar errors have been documented for FS in other recent work [31,32].

Trained U-Net performance is similar to manual segmentation

From the HCP-YA dataset, a set of 738 (left and right from 369 subjects) gold standard hippocampal tissue (that is, hippocampal GM and surrounding structures) segmentations were generated according to the manual protocol defined in [20]. Automated tissue segmentation was performed using nnUNet, a recent and highly generalizable implementation of a U-Net architecture [10] wrapped into a Snakemake workflow [DOI]. This software was trained on 80% (n=590) of the gold standard segmentation data described above, with the remaining 20% (n=148) making up a test set. Left and right hippocampi from the same participant were never split across training and testing sets due to their high symmetry. Note that all input images were preprocessed, resampled, and cropped (see Figure 1 and Online Methods) prior to training. Within the training set, 5-fold cross-validation was performed as implemented in the nnUNet code. Training took place on an NVIDIA T4 Turing GPU over 72 hours. This process was carried out using either T1w or T2w input data with the same training/testing data split. All default nnUNet data augmentation and hyperparameters were used.

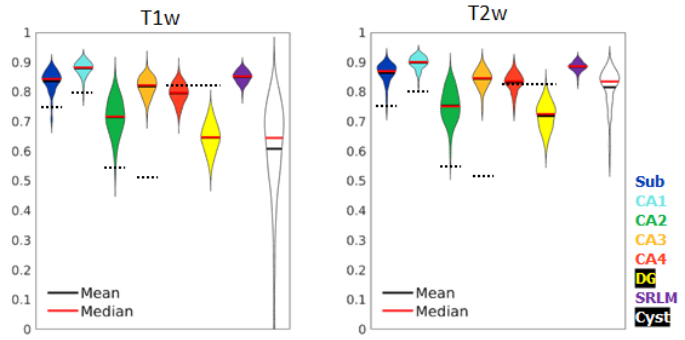


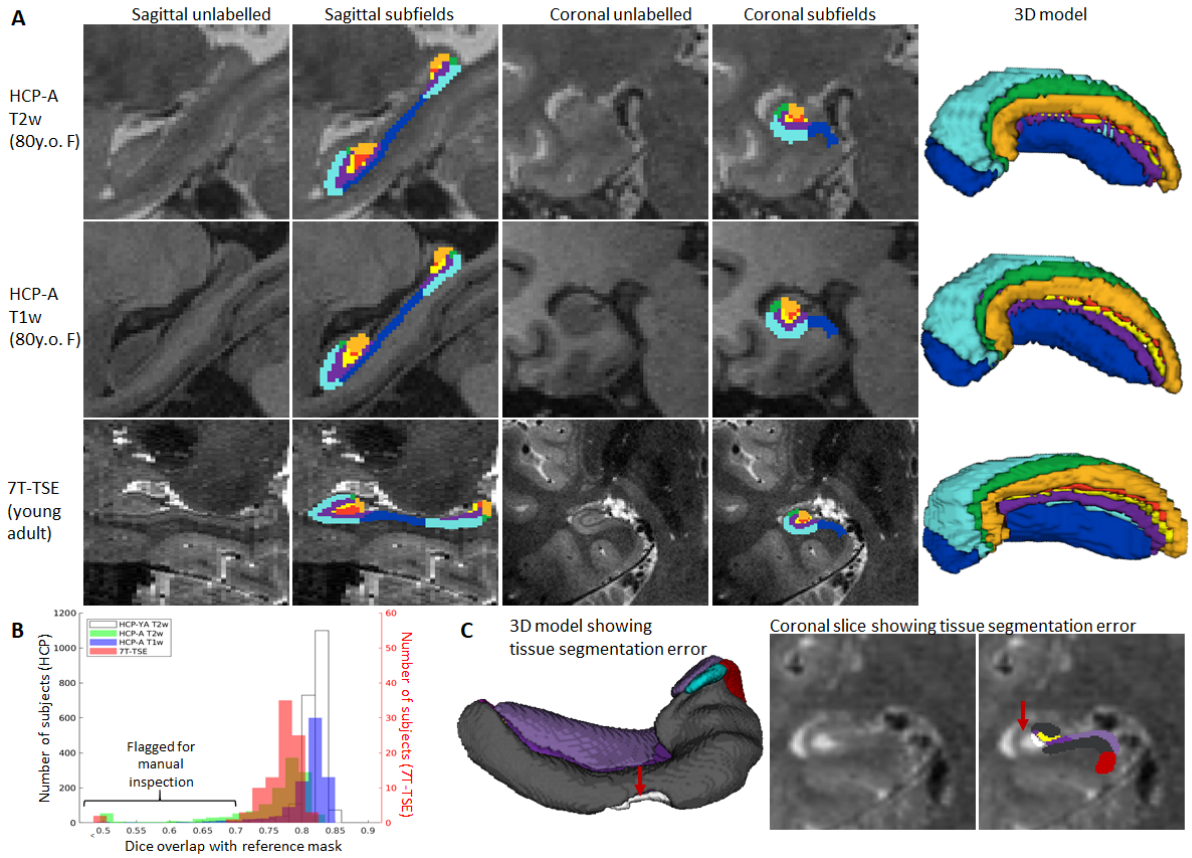
Figure 4. Test set performance in Dice overlaps between HippUnfold and manually unfolded subfields. All values are compared to ground truth manually defined tissues followed by unfolded subfield definition (manual unfold). Two models were trained in parallel using the same labels but different input MRI data modalities consisting of T1w or T2w data. Dotted black lines indicate corresponding values from [12], who include SRLM in all labels and combine CA4 and DG into one label.

bioRxiv preprint doi: https://doi.org/10.1101/2021.12.03.471134; this version posted December 5, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Dice overlap depends heavily on the size of the label in question, being lower for smaller labels. Typically a score of >0.7 is considered good, and many fully manual protocols show dice scores of >0.8 for the larger subfields like CA1 or the subiculum, and 0.6-0.8 for smaller subfields like CA2 or CA3 (see [17] for overview). Within the HCP-YA test set, performance was similar or better than most fully manual protocols for T1w and T2w data. Performance on T1w images was only marginally poorer than T2w images which typically better show the SRLM and are popular in manual subfield segmentation protocols [17].

Generalizability to unseen datasets and populations

We aimed to determine whether our pipeline would generalize to unseen datasets with different acquisition protocols and sample populations. Hippocampal morphometry, integrity, and subfields are often of interest in disease states where atrophy or other structural abnormalities are observed [1,33–35]. For this reason, we examined the HCP-A datasets in which we anticipated cases of severe atrophy would be present in some older subjects. Figure 5A shows results from one representative individual (an 80 y.o. female with signs of age-related atrophy but good scan quality). Another common use-case for hippocampal subfield segmentation is on anisotropic T2w data which is considered optimal for performing manual segmentation in most protocols [17], but may impose challenges for our method due to the difference in resolution. We thus applied HippUnfold to 7T-TSE data and also illustrate one representative subfield segmentation result in Figure 5A.



bioRxiv preprint doi: https://doi.org/10.1101/2021.12.03.471134; this version posted December 5, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Figure 5. Examination of HippUnfold performance on additional datasets HCP-A (T1w and T2w) and anisotropic 7T-TSE data. A) Sample subjects' HippUnfold subfield segmentation in native resolution. The first two rows come from the same subjects but using different input data modalities. B) Subjects flagged for Quality Assurance from each dataset based on Dice overlap with a reference mask approximated via deformable registration. C) Failed subject example illustrating missed tissue (red arrows) at the nnUNet pipeline stage.

Gold standard manual segmentations under the protocol used for subsequent unfolding were not available in the generalization datasets. Manually inspecting results from hundreds of subjects is time consuming. We thus streamlined this process by flagging potential segmentation errors by examining Dice overlap with a more conventional segmentation approach: deformable registration. For all datasets described above, we applied deformable fast B-spline registration [36] to the corresponding T1w or T2w template. Tissue segmentation results (generated at the nnUNet stage) were then propagated to template space and overlap with standard template hippocampal masks were examined, which is shown in Figure 5B. Any subject with a Dice overlap score of less than 0.7 was flagged and manually inspected for quality assurance. This made up 34/2126 (1.6%) samples in the HCP-YA T2w set (including training and testing subsets), 188/1312 (14.3%) samples from the HCP-A T2w set, 37/1312 (2.8%) samples from the HCP-A T1w set, and 3/92 (3.3%) samples from the 7T-TSE set. Closer inspection revealed that the vast majority of flagged cases were due to missed tissue in the nnUNet segmentation, an example of which is shown in Figure 5C. It is interesting to note that the most flagged cases were seen in the HCP-A T2w dataset even though T2w is a popular acquisition protocol for hippocampal subfield segmentation [17], and showed the best performance within the HCP-YA test set (Figure 4). This was likely not due to the age of subjects since few of the HCP-A T1w were flagged as possible errors, but instead may have been due to T2w scan quality, which was observed to be poor in some subjects, causing poor definition of the outer hippocampal boundaries. We recommend that future users carefully inspect results from any flagged subjects, and cases with errors can be either discarded or manually corrected. We cannot determine whether HippUnfold will work as intended on all new datasets, but within the generalization datasets examined here, results were excellent. Some work has already demonstrated it is possible to synthesize or convert between MRI modalities [37], which could be used to alleviate the dependency on any single MR contrast.

FAIR principles in development

We designed this pipeline to employ FAIR principles (findability, accessibility, interoperability, reusability). As such, we have made use of several tools, conventions, and data standards to make HippUnfold extensible and easy to use.

The default file input-output structure of the HippUnfold command line interface was built in compliance with the Brain Imaging Data Standards (BIDS) [38] Applications (BIDS-Apps) guidelines [39], and easily findable amongst the list of available BIDS Apps[1]. This is achieved via Snakebids, a tool designed to interface between BIDS datasets and Snakemake [40]. All aspects of HippUnfold use Snakemake [41], a workflow management system based on Python which is reproducible, scalable, and seamlessly combines shell commands, Python code, and

[1] https://bids-apps.neuroimaging.io/apps/

bioRxiv preprint doi: https://doi.org/10.1101/2021.12.03.471134; this version posted December 5, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

external dependencies in a human-readable workflow. There is no need to install these dependencies, which are containerized within the Singularity or Docker versions of HippUnfold.

Altogether, this means that in a single line this pipeline can be applied intelligently to any BIDS-complaint dataset containing a whole-brain T1w image and a T2w image (whole-brain or limited field of view) without having to specify further details. Typical runtimes on a standard desktop are 1 hour per subject, but this is further parallelized for faster processing when multiple subjects and added compute resources (or cloud computing) are available. Additional flags can be used to extend functionality to many other use-cases, including T1w only, T2w only, diffusion-weighted imaging, cases where a manual tissue segmentation is already available, or ex-vivo tissue samples.

Outputs of HippUnfold follow the standards for BIDS derivatives, and include preprocessed input images, volumetric subfield segmentations, inner, midthickness, and outer hippocampal surfaces11, vertex-wise morphometric measures of thickness, curvature, and gyrification, and a brief quality control (QC) report. All surface-based outputs are combined into a Connectome Workbench [42] specification file for straightforward visualization in alignment with HCP neocortical reconstructions. Outputs can be specified to include images in the original T1w space or in the resampled, cropped space that processing is performed in.

All code, code history, documentation, and support are offered online[2].

Discussion

One of the most powerful features of HippUnfold is its ability to provide topological alignment between subjects despite differences in folding (or digitation) structure. This is a critical element of mainstream neocortical analysis methods that until now had not been carried out at scale in the archicortex, or hippocampus. The power of this form of topological alignment is evident when mapping morphological or quantitative features across the hippocampus in a large population, which we demonstrate in Figure 2.

We compare HippUnfold to other commonly used tools for hippocampal analysis, Freesurfer v7.1 (FS) and Automated Segmentation of Hippocampal Subfields (ASHS) (Figure 3). Both of these methods rely on smooth deformation of single or multi-atlas references, meaning they cannot easily be generalized to drastically different hippocampal folding patterns which are often seen in the hippocampal head and tail. Both of these methods showed unfolded subfield patterns that were inconsistent with ground truth histology literature, including breaks in subfield topology, simplifications like the exclusion of the hippocampal tail, or inconsistent ordering of subfields. This highlights some of the advantages of HippUnfold, which did not suffer from these issues.

Several factors make surface-based methods difficult to implement in the hippocampus, including its small size and the difficulty of distinguishing the hippocampal sulcus or SRLM laminae that separate hippocampal folds. Here we have overcome these issues using a highly generalizable and sensitive neural network ‘U-Net’ architecture, combined with our previously developed topological unfolding framework. Together, these methods achieved similar or better Dice overlap scores than what is typically seen between two manual raters on all subfields. We

[2] https://github.com/khanlab/hippunfold

bioRxiv preprint doi: https://doi.org/10.1101/2021.12.03.471134; this version posted December 5, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

tested performance on new datasets ('generalization' datasets with different characteristics than the HCP training set) and saw good performance in nearly all cases. Specifically, we tested other common imaging protocols including different sample age groups (HCP-A) and thick-slice 7T TSE acquisitions often used in targeted hippocampal subfield imaging [17]. Though errors rates were low, we do show how and why such errors sometimes occur, highlighting the importance that future users examine the brief quality control reports included for each subject. Thus, while HippUnfold is demonstrated to work well with all datasets examined here, we expect the widespread adoption of higher-resolution acquisition techniques will further improve feasibility at other research institutes.

One important limitation of our method is that HippUnfold did not consistently show clear digitation in the hippocampal head, body, and tail which was sometimes seen in manual segmentation in the training set and in other work (see online Methods). This reflects a lack of detail that should ideally be captured by this pipeline, and affects downstream processing. That is, an erroneously smoothed hippocampi will appear thicker and have a smaller surface area compared to those that shows the full extent of digitations. This smaller surface area also results in each subfield boundary being proportionally shifted. Future work could improve this pipeline by training and testing with higher-resolution data where digitations can more clearly be distinguished both in labelmaps and in the underlying images.

The current work has applications beyond subfield imaging, enabling new investigations of the hippocampus on a columnar and laminar scale. For example, rather than performing ROI-based analyses, statistics can be performed on a per-vertex basis for vertices generated at different depths. This is in line with state-of-the-art neocortical analysis methods, and opens up the possibility of more precise localization of hippocampal properties. Similarly, it is worth noting that the methods used here are not necessarily restricted to MRI, as we have used the same surface-based unfolding in combination with manual segmentation to characterize the hippocampus in 3D BigBrain histology [20].

Altogether, we show that the BIDS App 'HippUnfold' that we have developed in this work respects the different internal hippocampal folding configurations seen between individuals, can be applied flexibly to T1w or T2w data, sub-millimetric isotropic or thick-slice anisotropic data, and compares favourably to other popular methods including manual segmentation, ASHS, and Freesurfer. We believe this tool will open up many avenues for future work including examination of variability in hippocampal morphology which may show developmental trajectories or be linked to disease, or the examination of hippocampal properties perpendicular or tangential to its laminar organization with diffusion-weighted imaging. Finally, it is worth noting that the methods here stand to improve existing techniques like subfield ROI-based analyses, quantitative or functional MRI sampling, or other techniques by providing greater anatomical detail and, critically, topological alignment between subjects.

bioRxiv preprint doi: https://doi.org/10.1101/2021.12.03.471134; this version posted December 5, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

References

1. Thom M. Review: Hippocampal sclerosis in epilepsy: a neuropathology review. Neuropathol Appl Neurobiol. 2014;40: 520–543.

2. O'Keefe J, Nadel L, Regents Professor of Psychology Lynn Nadel. The Hippocampus as a Cognitive Map. Oxford University Press, USA; 1978.

3. Dill V, Franco AR, Pinho MS. Automated Methods for Hippocampus Segmentation: the Evolution and a Review of the State of the Art. Neuroinformatics. 2015. pp. 133–150. doi:10.1007/s12021-014-9243-4

4. Duvernoy HM. The Human Hippocampus. 1998. doi:10.1007/978-3-662-03628-0

5. Van Essen DC, Drury HA, Joshi S, Miller MI. Functional and structural mapping of human cerebral cortex: solutions are in the surfaces. Adv Neurol. 2000;84: 23–34.

6. Waehnert MD, Dinse J, Weiss M, Streicher MN, Waehnert P, Geyer S, et al. Anatomically motivated modeling of cortical laminae. Neuroimage. 2014;93 Pt 2: 210–220.

7. Chang C, Huang C, Zhou N, Li SX, Ver Hoef L, Gao Y. The bumps under the hippocampus. Hum Brain Mapp. 2018;39: 472–490.

8. Ding S-L, Van Hoesen GW. Organization and Detailed Parcellation of Human Hippocampal Head and Body Regions Based on a Combined Analysis of Cyto- and Chemoarchitecture. J Comp Neurol. 2015;523: 2233–2253.

9. Cai S, Yu X, Zhang Q, Huang C, Gao Y. Is hippocampus getting bumpier with age: a quantitative analysis of fine-scale dentational feature under the hippocampus on 552 healthy subjects. Medical Imaging 2019: Image Processing. 2019. doi:10.1117/12.2512701

10. Isensee F, Jaeger PF, Kohl SAA, Petersen J, Maier-Hein KH. nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation. Nat Methods. 2021;18: 203–211.

11. DeKraker J, Ferko KM, Lau JC, Köhler S, Khan AR. Unfolding the hippocampus: An intrinsic coordinate system for subfield segmentations and quantitative mapping. Neuroimage. 2018;167: 408–418.

12. Yushkevich PA, Pluta JB, Wang H, Xie L, Ding S-L, Gertje EC, et al. Automated volumetry and regional thickness analysis of hippocampal subfields and medial temporal cortical structures in mild cognitive impairment. Hum Brain Mapp. 2015;36: 258–287.

13. Chakravarty MM, Steadman P, van Eede MC, Calcott RD, Gu V, Shaw P, et al. Performing label-fusion-based segmentation using multiple automatically generated templates. Hum Brain Mapp. 2013;34: 2635–2654.

14. Pipitone J, Park MTM, Winterburn J, Lett TA, Lerch JP, Pruessner JC, et al. Multi-atlas segmentation of the whole hippocampus and subfields using multiple automatically generated templates. Neuroimage. 2014;101: 494–512.

15. Du G, Cao X, Liang J, Chen X, Zhan Y. Medical Image Segmentation based on U-Net: A

bioRxiv preprint doi: https://doi.org/10.1101/2021.12.03.471134; this version posted December 5, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Review. Journal of Imaging Science and Technology. 2020. pp. 20508–20501. doi:10.2352/j.imagingsci.technol.2020.64.2.020508

16. Wisse LEM, Daugherty AM, Olsen RK, Berron D, Carr VA, Stark CEL, et al. A harmonized segmentation protocol for hippocampal and parahippocampal subregions: Why do we need one and what are the key goals? Hippocampus. 2017;27: 3–11.

17. Yushkevich PA, Amaral RSC, Augustinack JC, Bender AR, Bernstein JD, Boccardi M, et al. Quantitative comparison of 21 protocols for labeling hippocampal subfields and parahippocampal subregions in in vivo MRI: towards a harmonized segmentation protocol. Neuroimage. 2015;111: 526–541.

18. DeKraker J, Köhler S, Khan AR. Surface-based hippocampal subfield segmentation. Trends in Neurosciences. 2021. doi:10.1016/j.tins.2021.06.005

19. Amunts K, Lepage C, Borgeat L, Mohlberg H, Dickscheid T, Rousseau M-É, et al. BigBrain: an ultrahigh-resolution 3D human brain model. Science. 2013;340: 1472–1475.

20. DeKraker J, Lau JC, Ferko KM, Khan AR, Köhler S. Hippocampal subfields revealed through unfolding and unsupervised clustering of laminar and morphological features in 3D BigBrain. Neuroimage. 2020;206: 116328.

21. Iglesias JE, Augustinack JC, Nguyen K, Player CM, Player A, Wright M, et al. A computational atlas of the hippocampal formation using ex vivo, ultra-high resolution MRI: Application to adaptive segmentation of in vivo MRI. Neuroimage. 2015;115: 117–137.

22. Van Essen DC, Smith SM, Barch DM, Behrens TEJ, Yacoub E, Ugurbil K, et al. The WU-Minn Human Connectome Project: an overview. Neuroimage. 2013;80: 62–79.

23. Bookheimer SY, Salat DH, Terpstra M, Ances BM, Barch DM, Buckner RL, et al. The Lifespan Human Connectome Project in Aging: An overview. Neuroimage. 2019;185: 335–348.

24. Berron D, Vieweg P, Hochkeppler A, Pluta JB, Ding S-L, Maass A, et al. A protocol for manual segmentation of medial temporal lobe subregions in 7 Tesla MRI. Neuroimage Clin. 2017;15: 466–482.

25. Glasser MF, Sotiropoulos SN, Wilson JA, Coalson TS, Fischl B, Andersson JL, et al. The minimal preprocessing pipelines for the Human Connectome Project. Neuroimage. 2013;80: 105–124.

26. Ganzetti M, Wenderoth N, Mantini D. Whole brain myelin mapping using T1- and T2-weighted MR imaging data. Front Hum Neurosci. 2014;8: 671.

27. Hernández M, Guerrero GD, Cecilia JM, García JM, Inuggi A, Jbabdi S, et al. Accelerating fibre orientation estimation from diffusion weighted magnetic resonance imaging using GPUs. PLoS One. 2013;8: e61892.

28. Sotiropoulos SN, Hernández-Fernández M, Vu AT, Andersson JL, Moeller S, Yacoub E, et al. Fusion in diffusion MRI for improved fibre orientation estimation: An application to the 3T and 7T data of the Human Connectome Project. Neuroimage. 2016;134: 396–409.

29. Wael RV de, de Wael RV, Larivière S, Caldairou B, Hong S-J, Margulies DS, et al.

bioRxiv preprint doi: https://doi.org/10.1101/2021.12.03.471134; this version posted December 5, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Anatomical and microstructural determinants of hippocampal subfield functional connectome embedding. Proceedings of the National Academy of Sciences. 2018. pp. 10154–10159. doi:10.1073/pnas.1803667115

30. Crombe A, Planche V, Raffard G, Bourel J, Dubourdieu N, Panatier A, et al. Deciphering the microstructure of hippocampal subfields with in vivo DTI and NODDI: Applications to experimental multiple sclerosis. Neuroimage. 2018;172: 357–368.

31. Wisse LEM, Biessels GJ, Geerlings MI. A Critical Appraisal of the Hippocampal Subfield Segmentation Package in FreeSurfer. Front Aging Neurosci. 2014;6: 261.

32. Haast RAM, Lau JC, Ivanov D, Menon RS, Uludağ K, Khan AR. Effects of MP2RAGE B sensitivity on inter-site T reproducibility and hippocampal morphometry at 7T. Neuroimage. 2021;224: 117373.

33. Haukvik UK, Tamnes CK, Söderman E, Agartz I. Neuroimaging hippocampal subfields in schizophrenia and bipolar disorder: A systematic review and meta-analysis. J Psychiatr Res. 2018;104: 217–226.

34. Steve TA, Jirsch JD, Gross DW. Quantification of subfield pathology in hippocampal sclerosis: a systematic review and meta-analysis. Epilepsy Res. 2014;108: 1279–1285.

35. Carr VA, Bernstein JD, Favila SE, Rutt BK, Kerchner GA, Wagner AD. Individual differences in associative memory among older adults explained by hippocampal subfield structure and function. Proc Natl Acad Sci U S A. 2017;114: 12075–12080.

36. Modat M, Ridgway GR, Taylor ZA, Lehmann M, Barnes J, Hawkes DJ, et al. Fast free-form deformation using graphics processing units. Comput Methods Programs Biomed. 2010;98: 278–284.

37. Iglesias JE, Billot B, Balbastre Y, Tabari A, Conklin J, Gilberto González R, et al. Joint super-resolution and synthesis of 1 mm isotropic MP-RAGE volumes from clinical MRI exams with scans of different orientation, resolution and contrast. NeuroImage. 2021. p. 118206. doi:10.1016/j.neuroimage.2021.118206

38. Gorgolewski KJ, Auer T, Calhoun VD, Craddock RC, Das S, Duff EP, et al. The brain imaging data structure, a format for organizing and describing outputs of neuroimaging experiments. Sci Data. 2016;3: 160044.

39. Gorgolewski KJ, Alfaro-Almagro F, Auer T, Bellec P, Capotă M, Chakravarty MM, et al. BIDS apps: Improving ease of use, accessibility, and reproducibility of neuroimaging data analysis methods. PLoS Comput Biol. 2017;13: e1005209.

40. Khan A, Haast R. Snakebids - BIDS integration into snakemake workflows. Zenodo; 2021. doi:10.5281/ZENODO.4488249

41. Mölder F, Jablonski KP, Letcher B, Hall MB, Tomkins-Tinch CH, Sochat V, et al. Sustainable data analysis with Snakemake. F1000Res. 2021;10: 33.

42. Marcus DS, Harms MP, Snyder AZ, Jenkinson M, Wilson JA, Glasser MF, et al. Human Connectome Project informatics: quality control, database services, and data visualization. Neuroimage. 2013;80: 202–219.

bioRxiv preprint doi: https://doi.org/10.1101/2021.12.03.471134; this version posted December 5, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.