

# **Detection, definition and application of anatomical landmarks in medical images**



**Shouhei HANAOKA**

Department of Electronic and Information Engineering,  
the Graduate School of Engineering of  
Tokyo University of Agriculture and Technology

This dissertation is submitted for the degree of  
*Doctor of Philosophy*

March 2017



## Acknowledgements

Here I would like to thank a lot of people who have supported me and my life in Tokyo University of Agriculture and Technology.

First of all, I would like to heartily thank Prof. Akinobu Shimizu who supervised my thesis. Without him, I just could make nothing.

I also thank all people in Shimizu laboratory, including Dr. Atsushi Saito, Mr. Yusuke Takada, Mr. Kohei Yamashita, Mr. Masashi Kishimoto, Mr. Taro Sunada and Mr. Tatsuya Ito for supporting me. It was a happy time for me in the labo.

I thank Prof. Kuni Ohtomo who allowed me to be a Ph.D. candidate with working under him. I also thank all of my colleagues in Department of Radiology, the University of Tokyo, including Prof. Naoto Hayashi, Prof. Takeharu Yoshikawa, Dr. Soichiro Miki, Dr. Yukihiro Nomura, Dr. Tomomi Takenaga and Dr. Takahiro Nakao. Particularly I appreciate Dr. Mitsutaka Nemoto who has devoted his effort to development of landmark detectors.

I thank Prof. Yoshitaka Masutani a lot. Without him, I could not be here.

Finally I say thank you to my brother and parents. I devote this thesis to my deceased father.



## Abstract

The theme of this thesis is detection, automatic definition and application of anatomical landmarks. Firstly, I propose a detection method for 197 anatomical landmarks in clinical CT datasets. Secondly, I propose a method to define anatomically meaningful landmarks automatically from a large number of CT datasets. Thirdly, a segmentation method for the vertebra and the pelvic bones in CT volumes is presented. Fourthly, I report an automatic detection method of cerebral aneurysms using a novel HoTPiG image feature set.

In Chapter 1, the definition and classification of anatomical landmarks are discussed. Then, basic methods for landmark detection and major applications of them are described.

In Chapter 2, I propose a method to detect 187 anatomical landmarks simultaneously. Firstly, candidate points for each landmark are detected by a corresponding detector. Each detector outputs 100 candidates for each landmark. Then, the final output is generated by solving a combinatorial optimization problem in which the algorithm chooses one candidate from each candidate list. This problem is formulated using landmark point distribution model and maximum a posteriori estimation, and is solved using Markov chain Monte Carlo method and simulated annealing. The proposed method was evaluated with 109 CT datasets and 96.5% of landmarks were successfully detected within 20 mm from the ground truth points.

In Chapter 3, I try to define anatomically meaningful landmarks automatically. A new triangular consistency criterion (TCC) is introduced to evaluate each point in the human body. TCC is calculated based on

inconsistency of registration results among three volumes. The proposed method was evaluated with 50 CT volumes and 48 new landmarks were successfully defined.

In Chapter 4, I propose a segmentation method for the vertebrae and the pelvis which utilizes my landmark detection system. I modified diffeomorphic demons algorithm so that it can utilize landmark position information as well as grayscale volume information. This landmark-guided diffeomorphic demons algorithm is coupled with the multiatlas method to segment the spinal and pelvic bone regions. The proposed method was evaluated with 50 whole torso CT datasets and showed a segmentation accuracy which is comparable to other state-of-the-art methods.

In Chapter 5, I proposed a novel graph structure-based image feature named HoTPiG. HoTPiG is a feature set which can be calculated at each foreground voxel in an arbitrary binary volume. HoTPiG is defined as a three-dimensional histogram of graph distances among three points in the graph structure. Using this feature set, an automatic detection system for cerebral aneurysms in MR angiography (MRA) images was developed. The proposed system was evaluated with 300 MRA datasets and showed 81.8% of sensitivity when the number of false positives was three per case.

In Chapter 6, I conclude this thesis.

# Table of Contents

Chapter 1 Introduction .....	1
1.1 Definition and categorization of anatomical landmarks...	3
1.2 Landmark detection and image features.....	11
1.3 Medical image analysis and anatomical landmarks .....	14
1.3.1 Registration .....	14
1.3.2 Segmentation.....	15
1.4 Purpose of this thesis .....	16
References .....	17
Chapter 2 Automatic detection of landmarks .....	19
2.1 Introduction.....	20
2.1.1 Two major problems .....	20
2.1.2 Two strategies .....	22
2.1.3 Previous works.....	23
2.1.4 Proposed approach .....	25
2.2. Methods .....	30
2.2.1 Single-landmark detector.....	32
2.2.2 General stochastic model of landmark detector.....	33
2.2.2.1 Single-candidate model .....	33
2.2.2.2 Multiple-candidate model .....	36
2.2.3 Landmark point distribution model.....	37
2.2.3.1 Coordinate-based L-PDM .....	38
2.2.3.2 Distance-based L-PDM .....	39
2.2.4 Maximum <i>a posteriori</i> estimation .....	40
2.2.5 Combinatory optimization with Gibbs sampling and simulated annealing.....	45
2.2.6 Proposed two-stage sampling.....	48
2.2.7 Evaluation .....	53
2.2 Results.....	58
2.3.1 Detection accuracy.....	58
2.3.2 Correlation between performance of each single- landmark detector and overall detection accuracy .....	65
2.3.3 Accuracy of vertebra detection and identification using	

Dataset A.....	67
2.3.4 Accuracy of vertebra detection and identification using Dataset B .....	70
2.3.5 Position estimation of landmarks outside the imaging range.....	73
2.3.6 Landmark position ambiguity analysis using interobserver errors among human experts .....	75
2.4 Discussion .....	76
2.5 Conclusion .....	81
Appendix 1. Calculation of the multiple candidate model.....	82
Appendix 2. Importance sampling of the artificial candidates	85
References .....	88
Chapter 3 Automatic definition of landmarks .....	93
3.1 Introduction.....	93
3.2 Methods .....	95
3.2.1 Datasets and registration.....	95
3.2.2 Triangular consistency criterion calculation .....	95
3.2.3 Automatic landmark definition.....	96
3.3 Results.....	99
3.4 Discussion .....	105
3.5 Conclusion.....	106
References .....	107
Chapter 4 Multiatlas-based segmentation of the vertebral and pelvic bones .....	109
4.1 Introduction.....	110
4.2 Methods.....	116
4.2.1 Landmark-guided diffeomorphic demons algorithm. 116	
4.2.1.1 Log-domain demons algorithm.....	116
4.2.1.2 Proposed landmark-guided demons algorithm ..	123
4.2.2 Multiatlas segmentation framework.....	126
4.2.2.1 Landmark detection system .....	129
4.2.2.2 Preprocessing .....	130
4.2.2.3 Atlas selection .....	130
4.2.2.4 Landmark-guided demons.....	131

4.2.2.5 Local histogram matching .....	132
4.2.2.6 Detailed demons algorithm .....	133
4.2.2.7 Label propagation and voting .....	133
4.2.3 Parameter optimization and evaluation .....	134
3. Result.....	136
4. Discussion.....	149
5. Conclusion.....	153
References .....	154
Chapter 5 HoTPiG: A novel geometrical feature for vessel morphometry and its application to cerebral aneurysm detection	161
5.1 Introduction.....	162
5.2 HoTPiG.....	165
5.3 Computer-assisted detection of aneurysms.....	168
5.3.1 Artery region extraction and HoTPiG feature calculation .....	168
5.3.2 Voxel-based classification by SVM.....	169
5.4 Experimental results .....	171
5.5 Discussion .....	173
5.6 Conclusion .....	174
References .....	174
Chapter 6 Conclusion.....	177
Appendix A: List of Publications .....	181
Papers used in this thesis:.....	181
International journal paper.....	181
International conference papers .....	182
Other publications:.....	182
International journal paper.....	182
International conference paper.....	185
Domestic conference papers:.....	190
Thesis (PhD of Medicine).....	192
Appendix B Anatomical landmark list.....	193



# Chapter 1 Introduction

From the discovery of X-ray by Röntgen in 1895, the research field of medical imaging has been rapidly developed. Today, computed tomography (CT) and magnetic resonance imaging (MRI) are indispensable diagnostic tools for physicians who must know what is going on in the bodies of patients. A single CT or MRI examination can generate huge data of the order of gigabytes. It is not an easy task for medical doctors to interpret such a huge amount of data in their daily work. That is why assistance by computer is desired in modern medical image interpretation. In a practical sense, the goal of medical image analysis is to help physicians to interpret medical images.

Physicians interpret images using their anatomical knowledge on the

human body. Likewise, when a certain medical image analysis process has to handle a particular organ, the computer program may have to know anatomical knowledge of the target organ, e.g. the position, pose, shape, appearance, and their statistical variations. In this sense, a methodology to migrate knowledge of physicians to computers is required.

Landmarks are one of the most primitive type of anatomical knowledge representation. In an anatomical sense, a landmark is a point with its own name. Owing to its primitiveness, landmarks can easily be handled by a computer application. On the other hand, landmark position information is frequently used to determine initial condition of more complicated medical image analysis processes such as segmentation and registration. That is why landmarks are important in medical image analysis. Automatic detection of anatomical landmarks and its application are the main issues of this thesis.

In the rest of this chapter, I firstly discuss the definition and categorization of landmarks. Secondly, classical detection methods for landmarks are described. Thirdly, the applications of anatomical landmark detection tasks are shown. Finally, the purpose and structure of this thesis are described.

## 1.1 Definition and categorization of anatomical landmarks

In the anatomical sense, “landmarks” are defined at points, curves, etc., with specific features that are commonly found in every individual with a certain correspondence in location and topology. In addition, unique names or labels are often given to these anatomical landmarks to distinguish them from each other. In the terminology for shape representation such as in statistical shape models (SSMs) [1], however, ‘landmark’ is often only used for points included in shape models even if they have no characteristic geometric and anatomical features.

In this thesis, I use the term “anatomical landmark” as landmarks which have characteristic anatomical features. It is expected to have a characteristic (but not always unique) appearance, can be given a unique name, and exists in most of human bodies.

An anatomical landmark can be defined by several ways, including topological, morphological, curvature-based, or manipulative manner. Sometimes it may be defined on a surface of an organ or a tissue, but sometimes it can be inside or even outside of the organ (e.g. the center of a hole of a bone). In the morphometric definition by Bookstein [2], landmarks including anatomical ones are classified into the following three

types depending on their degree of homology.

- Type 1. Discrete juxtapositions of tissues
- Type 2. Maxima of curvature or other local morphogenetic processes
- Type 3. Extremal points

Type 1 landmarks include branching points for tree structures, such as blood vessels, and the centers/centroids of sufficiently small structures such as vertebrate eyes. These appear to be the most easily defined and reproduced landmarks. For this reason, names or labels are given to most major anatomical structures. The points with maximum curvature on the profiles of structures are assigned as type 2 landmarks. For example, the tips of teeth and spinous processes of vertebra are categorized as this type. The third type of landmarks is constructed geometrically for the sake of convenience in measurement and includes the end-points of diameters and centroids, and so forth. Type 3 landmarks are often categorized as pseudo/semi-landmarks such as regularly sampled points on curves and surfaces used in shape models.

Masutani [3] proposed another categorization of anatomical landmarks in the viewpoint of automatic detection based on local image features. The categorization is as follows:

*Class 1.* Landmarks with salient feature, and uniquely detectable. In addition, there is no similar structure in other part of body. This

type of landmark is thought to be a key in medical image understanding.

*Class 2-a.* Ones with salient feature but there exist similar patterns around the landmark within a structure group. For instance, landmark at each spinous process of vertebra in the spinal column belongs to this category. Within such structures, those landmarks are virtually homologous and therefore are difficult to distinguish each other.

*Class 2-b.* There exist similar patterns in a few different structures. In other word, they are similar in image features but never homologous. For example, iliac crest has a ridge feature similar to ones at shoulder blade.

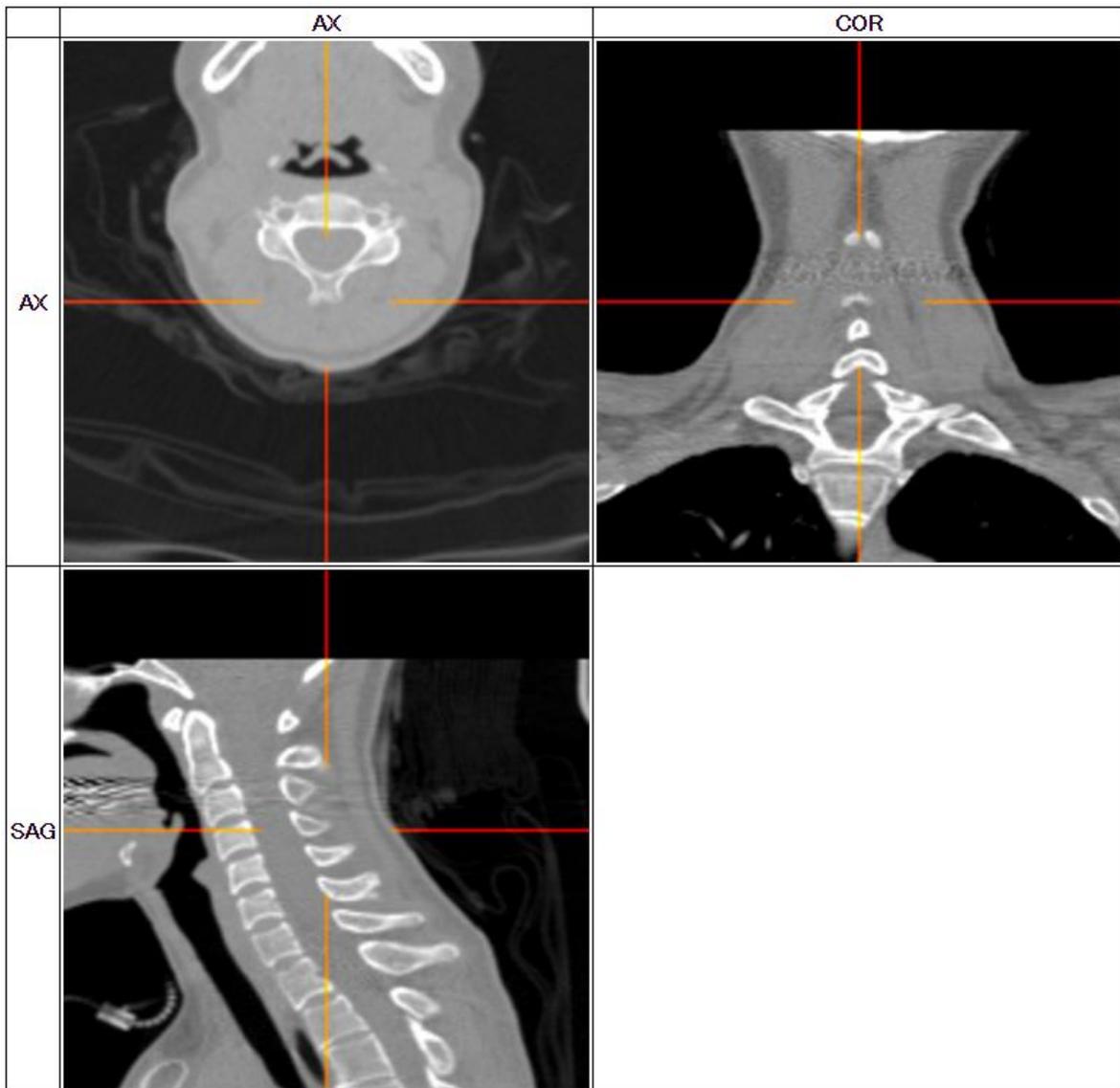
*Class 3.* Landmarks with few features, and therefore too many similar structures are found all over the data. This class includes points on flat part of liver surface. These are detected only via group-wise matching procedure such as registration. This kind of landmarks is sometimes called as semi-landmarks.

In Masutani's categorization, the class 1 landmarks are easiest to be detected and the class 3 landmarks are most difficult. Comparing class 2-a and 2-b, probably class 2-a landmarks are more difficult to detect. This is because when a series of class 2-a landmarks (e.g. processes of vertebrae) are detected, the identification of their numbers must be performed (e.g. as the

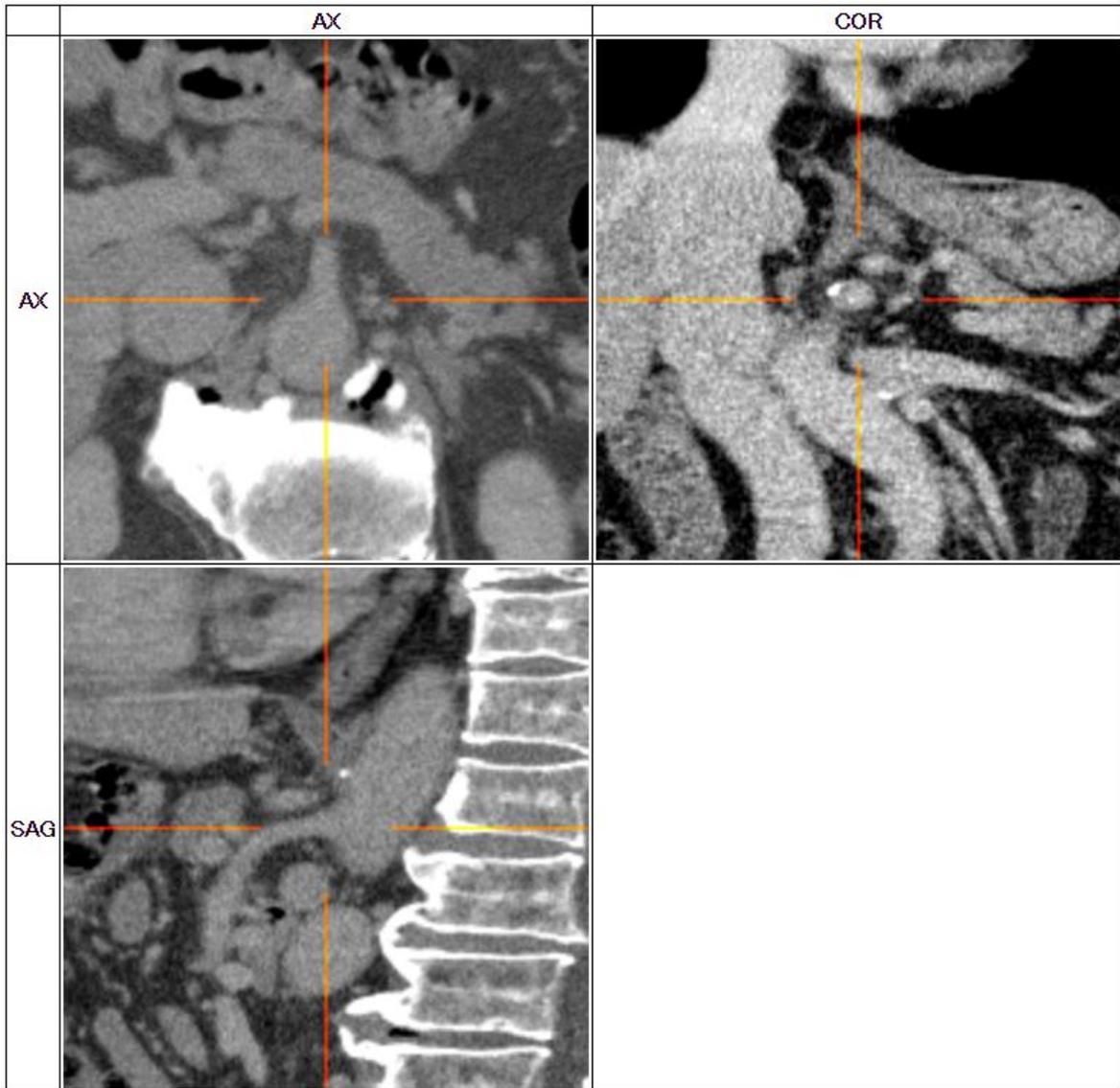
3rd cervical, 4th thoracic or 5th lumbar ones).



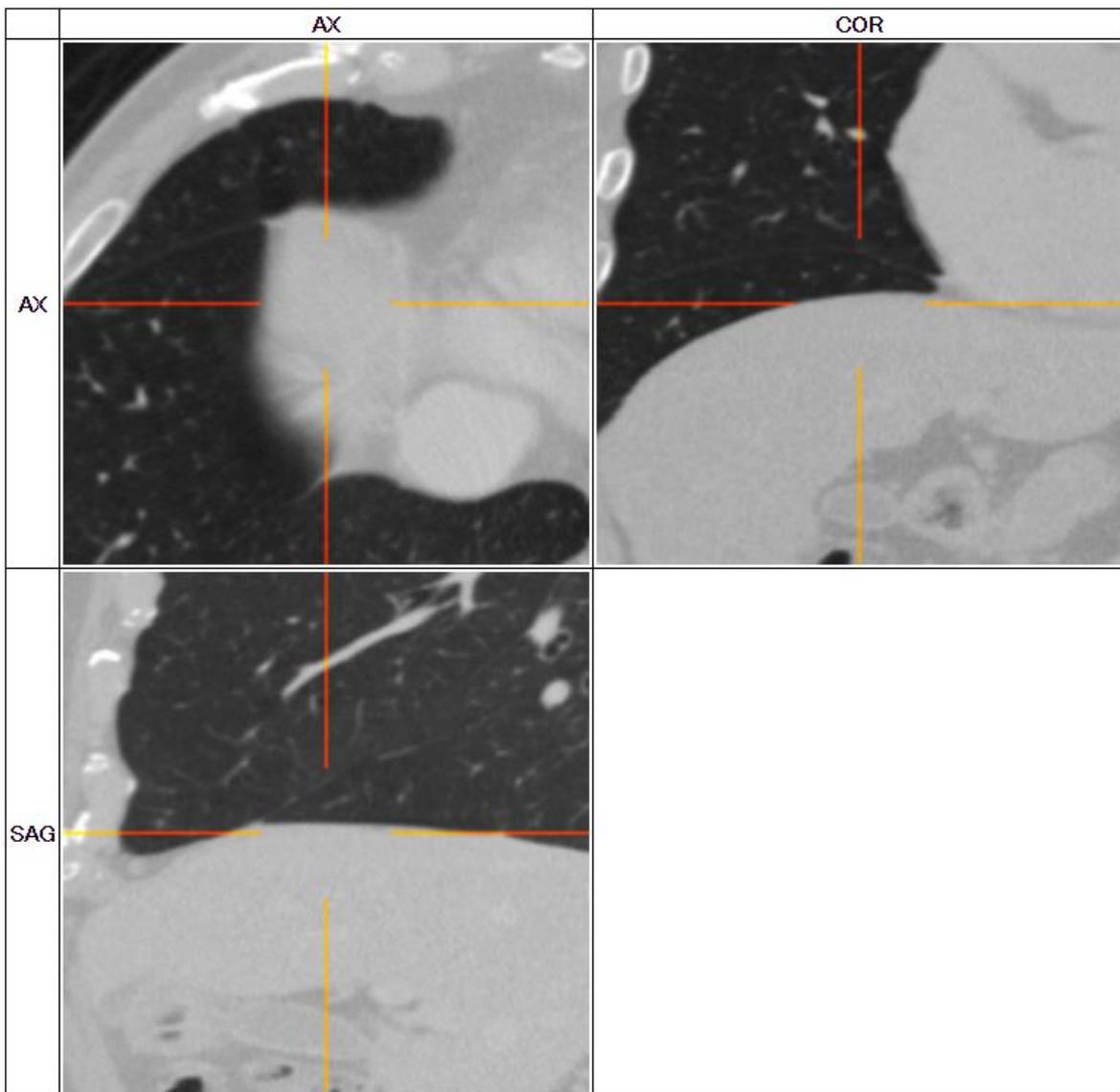
**Fig. 1.** Example of class 1 landmarks (the atlantoaxial joint). AX=axial, COR=coronal, SAG=sagittal cross-sections.



**Fig. 2.** Example of class 2-a landmarks (tip of the spinous process of the 4th cervical vertebra).



**Fig. 3.** Example of class 2-b landmarks (the root of the superior mesenteric artery).



**Fig. 4.** Example of class 3 landmarks (the superior margin of the liver surface).

Anatomical landmarks can also be categorized by their importance in academic or clinical applications. Note that clinically important landmarks are not necessarily categorized into class 1. An important example is landmarks defined on the vertebrae and the ribs. These are anatomically important class 2-a landmarks. It is also possible that an important landmark is class 3, or extremely, not visible in the given modality (e.g., CT or MRI). In other words, both Bookstein's and Masutani's classifications are modality-dependent.

All of Masutani's class 1, 2 and 3 landmarks are target of this thesis. Naturally, class 3 landmarks cannot be defined or detected unless a proper spatial context (for example, the positions of other landmarks) is given. This is because there are many points which have similar appearances to each class 3 landmark. On the other hand, the human body includes a large number of possible class 3 landmarks compared to class 1 or 2 landmarks. It can be possible to fill the human body with such class 3 landmarks. Such a large landmark set can be regarded as a kind of computational representation of anatomical knowledge and has a wide variety of applications as written below. Filling the human body with well-defined landmarks is an ultimate goal of my study.

Note that not all points in the human body can be a class 3 landmark; point-to-point correspondence among human subjects is a critical property of landmarks. Each landmark should be determined at one point in any

human body. A class 3 landmark can only be determined by a manipulative way; for example, the border point of the jejunum and the ileum can be determined as the 2:3 internally dividing point of the total small intestine. This can be a class 3 landmark because it can be determined with point-to-point correspondence among human subjects. On the other hand, the appendix can be used as a class 1 landmark. However, it is possible that the appendix is resected by a surgery. Generally speaking, a landmark can be lost or even duplicated due to anatomical variants or a surgery. A landmark detection system should also handle such a situation.

A lot of anatomically definable landmarks already have their own names. However, it is probable that there are nameless but anatomically definable landmarks which can be detected easily (e.g. being class 1) and thus useful in medical image analysis. In Chapter 3 I will describe a method to automatically determine such ‘nameless but useful’ landmarks.

## 1.2 Landmark detection and image features

The most classic and popular method for landmark detection is a sliding window method [4] [5]. In a sliding window method, a window with a certain size is placed on the given image. Then, the likelihood of existence of the target landmark at the center of the window is calculated, using the

intensity information of the pixels/voxels in the window. This likelihood calculation is repeated at every possible position, and the place with the largest likelihood is outputted as the estimated landmark position.

For calculating the likelihood, the simplest method is a template matching [6]. Here, a template is a typical partial image around the target landmark. In a template matching method, it is assumed that the partial images around the target landmark have a sufficiently small variation of intensity patterns among subjects. Under this assumption, a template matching method simply calculate the distance (e.g., sum of squared error) between the partial image of the window and the given template.

Obviously, template matching will not work well if the variation of the partial image around the landmark is large. To overcome this limitation, machine learning-based methods are frequently used. In a machine learning-based method, a large amount of training datasets are used to learn the intensity distribution of the partial images around the target landmark. When an unseen image is inputted, the likelihood of each window is estimated by the trained machine learning algorithm. Although image intensities of the given partial image can be directly used as the input of the machine learning, usually landmark-dedicated image features are extracted from the partial image, e.g., the curvature of the image contour at the center of the window. Other image features frequently used include Haar-like and DoG features [7]. In Chapter 5, I will introduce a novel HoTPiG image

feature set which can detect vessel bifurcations or aneurysmal objects.

Recently, a landmark detection method other than sliding windows has been reported [8]. It is based on regression, instead of likelihood estimation. When a window is given, a regression tree (a kind of machine learning method) estimates the dislocation between the center of the window and the landmark. This dislocation estimation is performed at every position in the image, and the final landmark position is estimated by voting.

The detection algorithm described above is about detecting one landmark. Detection of multiple landmarks is more complicated because spatial relationship among landmarks should be considered. One possible approach is a sequential approach in which landmarks are detected one by one [9]. However, it is not easy to detect many class 2 or 3 landmarks by the sequential approach. That is because, in the sequential approach, one detection miss can be crucial in the following process. Especially it is difficult to detect many class 2-a landmarks (i.e., landmarks on the spine and ribs) by the sequential approach when the imaging range does not include the entire spine. So forth, few study have been reported in which many class 2-a or 3 landmarks are simultaneously detected.

In Chapter 2 I will discuss how to detect multiple landmarks simultaneously. Each single landmark detection can fail due to absence of the target tissue (e.g. after a surgery), being out of imaging range, or insufficient sensitivity of the detector. I will build a stochastic detector

model which can handle these possibilities. Furthermore, I will introduce a landmark point distribution model (L-PDM) which is a stochastic model of spatial distribution of landmarks. Using these two models, I will formulate the multiple landmark detection problem as a maximum *a posteriori* problem.

## 1.3 Medical image analysis and anatomical landmarks

The automatic detection of anatomical landmark positions has a wide range of applications in medical image processing. For example, pre-detected landmark positions are frequently used in determining the initial condition of statistical shape models (SSMs) for the region segmentation of various organs [1] or the image registration of two human bodies [10]. In both segmentation and registration, landmark detection is frequently used to set the initial condition of the energy optimization problem.

### 1.3.1 Registration

Registration is a process to align two or more images spatially. Registration may be done between different modality images of the same patient or the same modality images of different patients. It is known that registration problem is an inverse problem which requires a sound regularization and the use of proper models [10]. Landmarks can be used

in registration by fitting corresponding landmark positions between two images. Deformation and warping of the positions other than landmarks can be calculated using interpolation. Otherwise, when registration problem is formulated as an energy optimization problem, landmarks can be used to set the initial condition of the problem solver. In Chapter 4, I will show a landmark-based registration algorithm in which landmark position fitting is integrated to the image fitting and deformation.

### 1.3.2 Segmentation

Segmentation is a process to determine the region of a target organ such as brain or a target tissue such as vessels. Segmentation can be done by either data-driven or model-based manner. A frequently used model-based method is a statistical shape model (SSM). Fitting a SSM to a given image is usually formularized as an energy optimization problem, so landmarks can be used to set the initial condition (i.e., position, pose and shape) of the SSM. Moreover, when SSM is represented as a point distribution model (PDM), the points can be landmarks themselves; otherwise, these points can be semilandmarks (i.e., Masutani's class 3 landmarks). In Chapter 2, I will describe a landmark point distribution model (L-PDM) which can represent landmark positions in the whole body. And In Chapter 4, I will show a registration-based segmentation method for the spine and the pelvis region in CT volumes.

## 1.4 Purpose of this thesis

The purpose of this thesis is to describe the methods for automatic detection of landmarks, automatic definition of landmarks, and application of detected landmarks to segmentation and registration problems.

In the rest of this thesis, I will describe:

- How to define a L-PDM and how to detect over 100 landmarks simultaneously (Chapter 2)
- How to automatically define anatomically meaningful landmarks from a given large dataset of CT volumes (Chapter 3)
- A landmark-guided image registration method and its application to multiatlas-based segmentation of the spinal and pelvic bones in CT volumes (Chapter 4)
- A novel graph-based image feature set named HoTPiG and its application to an automatic arterial aneurysm detection in magnetic resonance angiography images (Chapter 5)

Finally, I conclude the thesis in Chapter 6.

## References

1. Heimann T, Meinzer H-P (2009) Statistical shape models for 3D medical image segmentation: A review. *Medical image analysis* 13 (4):543-563
2. Bookstein FL (1997) *Morphometric Tools for Landmark Data: Geometry and Biology*. Cambridge University Press,
3. Masutani Y *Mathematical Foundations of Computational Anatomy*. In: *The First International Symposium on the Project Computational Anatomy*, Tokyo University of Agriculture and Technology, Tokyo, Japan, 2/7/2010 2010. Steering Committee of the Project Computational Anatomy, pp 45-51
4. Viola P, Jones M Rapid object detection using a boosted cascade of simple features. In: *Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on*, 2001. IEEE, pp I-511-I-518 vol. 511
5. Dalal N, Triggs B Histograms of oriented gradients for human detection. In: *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, 2005. IEEE, pp 886-893
6. Lewis JP Fast template matching. In: *Vision interface, 1995*. vol 120123. pp 15-19
7. Nemoto M, Masutani Y, Hanaoka S, Nomura Y, Yoshikawa T, Hayashi N, Yoshioka N, Ohtomo K (2011) A unified framework for concurrent detection

of anatomical landmarks for medical image understanding. In: SPIE Medical Imaging 2011. 7962, 03/14/2011 2011. pp 79623E-79623E-79613.

8. Criminisi A, Robertson D, Konukoglu E, Shotton J, Pathak S, White S, Siddiqui K (2013) Regression forests for efficient anatomy detection and localization in computed tomography scans. *Medical Image Analysis* 17 (8):1293-1303

9. Liu D, Zhou KS, Bernhardt D, Comaniciu D Search strategies for multiple landmark detection by submodular maximization. In: *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*, 2010. IEEE, pp 2831-2838

10. Fischer B, Modersitzki J (2008) Ill-posed medicine—an introduction to image registration. *Inverse Problems* 24 (3):034008

## Chapter 2 Automatic detection of landmarks

An automatic detection method for 197 anatomically defined landmarks in computed tomography (CT) volumes is presented. The proposed method can handle missed landmarks caused by detection failure, a limited imaging range and other problems using a novel combinatorial optimization framework with a two-stage sampling algorithm. After a list of candidates is generated by each landmark detector, the best combination of candidates is searched for by a combinatorial optimization algorithm using a landmark point distribution model (L-PDM) to provide prior knowledge. Optimization is performed by simulated annealing and iterative Gibbs sampling. Prior to each cycle of Gibbs sampling, another sampling algorithm is processed to estimate the spatial distribution of each target landmark, so that landmark positions without any correct detector-derived candidates can be estimated. The proposed method was evaluated using

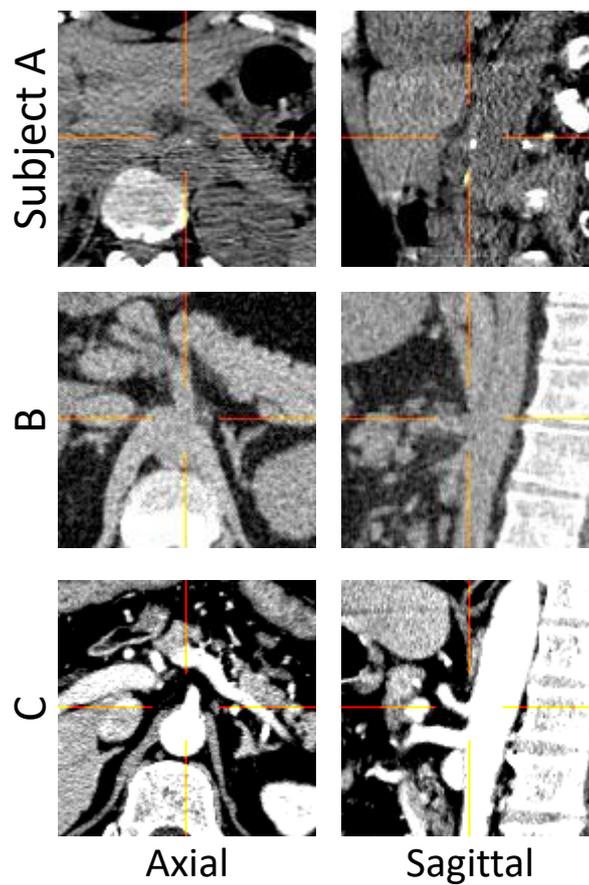
104 CT volumes with various imaging ranges. The overall average detection distance error was 6.6 mm, and 83.8, 93.2 and 96.5% of landmarks were detected within 10, 15 and 20 mm from the ground truth, respectively. The proposed method worked even when most of the landmarks were outside of the imaging range. The identification accuracy of the vertebral centroid was also evaluated using public datasets and the proposed method could identify 70% of vertebrae including severely diseased ones. From these results, the feasibility of my framework in detecting multiple landmarks in various CT datasets was validated.

## 2.1 Introduction

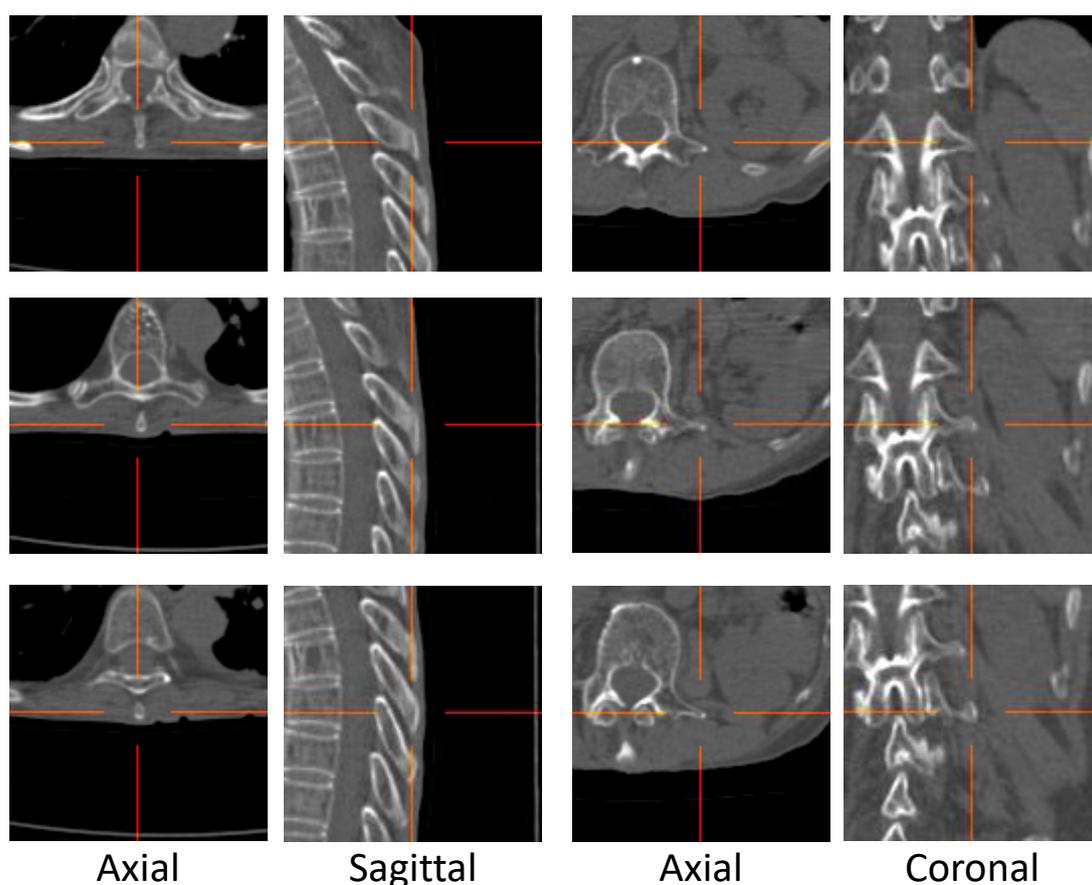
### 2.1.1 Two major problems

There are two major difficulties in detecting an anatomical landmark on the basis of its shape properties: (1) inter-individual variation and (2) insufficient intra-individual uniqueness of the shape. Firstly, landmark shape differences between individuals can be very large. A simple template-matching technique has been frequently used in practice under the assumption that the target landmark has sufficiently small inter-individual variations in its local appearance (i.e., the intensities of the voxels around the landmark point). However, many practically important anatomical landmarks in the human body, such as the tips of bone structures or vessel bifurcations, have large inter-individual variations in their appearance (Fig. 1). Furthermore, the human body includes many points whose local appearance is similar to the target landmark (e.g., the tip of another bony process or another blood vessel bifurcation). The most extreme example of this problem is the class of landmarks defined on repetitive body segment

structures, that is, the vertebrae and ribs (Fig. 2). To avoid these difficulties, it is necessary to utilize prior knowledge of the positions of landmarks in the human body. Since the positions of landmarks vary among individuals, a statistical model for the spatial distribution of the landmarks is required.



**Fig. 1.** Example of a landmark with large inter-individual variation (the root of the celiac artery).



**Fig. 2.** Examples of landmarks with repetitive shapes. (Left from top) Tips of the spinal processes of the 4th, 5th and 6th thoracic vertebrae. (Right from top) Tips of the transverse processes of the 1st, 2nd and 3rd lumbar vertebrae.

### 2.1.2 Two strategies

To overcome these difficulties, two strategies for detecting multiple landmarks can be considered, a sequential strategy and a simultaneous strategy. The former involves the detection of landmarks one by one. In this approach, the order of landmark detection is critical and a suitable order must be determined. For example, the order of landmark detection reliabilities can be used. One of the benefits of this approach is that each

detector can utilize the positions of landmarks that have already been detected. In this approach, however, a single detection failure may affect all of the subsequent landmark detections. Thus, the detection reliability for each landmark must be tolerably high. This becomes more problematic as the number of landmarks is increased. The latter strategy involves the detection of all landmarks independently. Typically, each detector is designed to detect only one landmark and outputs several candidate positions. After the candidates are enumerated for all landmarks, the best combination of candidates is searched for. This approach is advantageous when the detection reliability for each landmark is relatively low and not sufficient for the sequential approach. This is because a single detection error only has a limited affect compared with the sequential approach. One of the difficulties of this approach is how to solve the combinatorial optimization problem. Because each landmark may have multiple candidates, a large number of landmarks will lead to a combinatorial explosion. Another difficulty is how to handle landmarks outside of the imaging range in the combinatorial optimization process.

### 2.1.3 Previous works

Thus far, a number of automatic detection methods for anatomical landmarks in computed tomography (CT) images have been introduced (Table 1). First of all, two typical studies are introduced below.

A simultaneous and optimization-based method for detecting 22 landmarks in the body trunk was proposed by Potesil et al. [1]. They used a pictorial structure model in which the spatial distribution of landmarks is represented by a graph structure whose edges connect selected pairs of landmarks. The entire problem was formulated as an energy minimization

problem and solved by a belief propagation method. Recently, Criminisi et al. reported a method of determining bounding boxes for 26 organs including the lungs, kidneys and liver [2]. A single random regression forest that can detect all 26 organs simultaneously was trained and used in their method.

As a sequential strategy-based approach, Liu and Zhou [3,4] have reported a method for detecting 63 landmarks in the body trunk, in which the search ranges of landmarks yet to be detected are limited using the already detected landmark positions. The process begins with a certain landmark, namely, an “anchor landmark.” If the anchor landmark detection fails, another landmark is chosen as a new anchor. In this approach, prior knowledge of the spatial landmark distribution is used to localize the search space and reduce the number of false positive (FP) detection results. Recently, they improved their method [4] by providing an initial estimate of landmark positions by propagating the landmark positions of the most similar volume in their database found by a nearest-neighbor search.

In the context of spinal landmarks, vertebral landmark detection is equivalent to vertebral identification, in which each vertebra is localized and identified with its number (such as the 3rd cervical, 4th thoracic or 5th lumbar). In this research field, Klinder et al. first reported a whole-spine vertebral bone identification and segmentation method [5]. Recently, a couple of other recent studies have used Markov random field (MRF) optimization to solve combinatorial optimization problems involving vertebral disks [6] and whole-body landmarks [7]. Glocker et al. [8] utilized a hidden Markov model (HMM) to detect all 24 vertebral bone centroids in CT volumes with various imaging ranges. Their model includes a global scale factor as a hidden variable and is solved via dynamic linear programming. Recently, Glocker et al. reported another vertebral

centroid detection method using discriminative random classification forests [9] with better detection results. Kelm et al. [10] also reported a study in which iterative marginal space learning (MSL) was used to detect the position, orientation and scale of the intervertebral disks.

#### 2.1.4 Proposed approach

In this study, a framework for detecting 197 landmarks simultaneously from CT datasets is presented. Compared with the previous works introduced above, the main contributions and advantages of my framework are as follows:

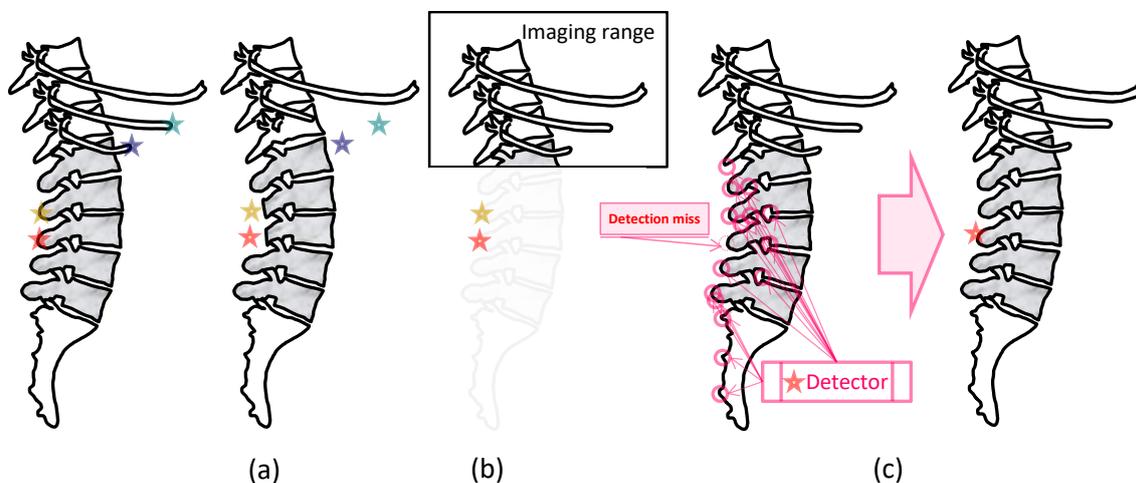
- The development of a novel Gibbs-sampling- and importance-sampling-based combinatory optimization framework. This framework can stochastically handle missing landmarks (Fig. 3) caused by limited detector sensitivity, a limited imaging range or a surgical/anatomical defect. The framework can also estimate the positions of missing landmarks (i.e., the most appropriate positions for defective, out-of-range or non-detected landmarks) using the other landmark positions. To the best of my knowledge, this is the first study on medical images in which the positions of such missed landmarks are systematically estimated.
- A stochastic model of a general landmark detector. In the model, the failure to detect a target landmark is regarded as a random event. Using Bayesian estimation, the algorithm can handle detectors with low reliability and utilize them optimally. This stochastic model is optimally integrated with the combinatorial optimization framework to detect or estimate difficult landmarks (e.g. the landmarks shown in Figs. 1 and 2). This integration is also critical to handling as many as 197

landmarks concurrently.

- An experimental validation in which the proposed framework reliably detected multiple landmarks, including vertebral bone landmarks, to identify all 24 vertebrae. The algorithm can identify and label the vertebrae even if the given CT volume includes only a short part of the spinal column. This is the first report of the simultaneous detection of multiple landmarks on each of all 24 vertebrae in CT datasets.
- A novel method that can estimate the positions of landmarks outside the imaging range. Through such estimation of landmark positions out of the imaging range, the alignment of an SSM can be improved greatly for datasets in which the segmentation target organ is partially included.

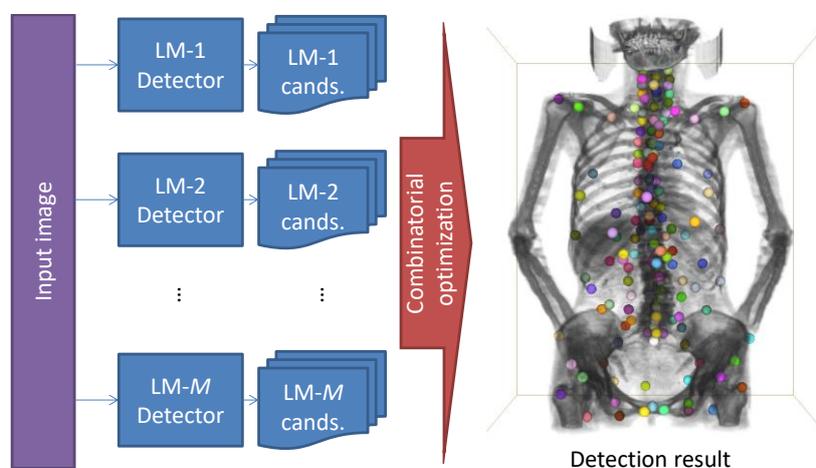
As illustrated in Fig. 4, the proposed framework begins by enumerating candidates for each landmark position using the corresponding single-landmark detector. Then, a Gibbs-sampling-based combinatorial optimization algorithm [11] searches for the combination of candidates that best fits the given landmark point distribution model (L-PDM). The L-PDM is used to provide prior knowledge of the variation in the spatial distribution of landmarks in the human body. Here, some landmarks may not be detected owing to detector failure, the absence of anatomical structures, or their being outside the imaging range (Fig. 3). In particular, a single detection failure of a spinal landmark may cause a shift in vertebral bone identification, leading to multiple landmark detection failures in the optimization phase. To solve this problem, a novel two-stage sampling algorithm is introduced (Fig. 5). In this algorithm, Gibbs sampling is combined with another artificial-candidate sampling algorithm that can estimate the spatial distribution of each missed landmark. Since this

estimation of the spatial distribution only uses information on other landmark positions and does not use local appearance information, it works even if a landmark exists outside the imaging range. From the estimated spatial distribution, several artificial candidates are sampled and merged with detector-derived candidates. After this artificial-candidate sampling, the Gibbs sampling algorithm samples one candidate from the merged candidate set and then proceeds to sample the next landmark. The artificial candidates are repeatedly updated before each cycle of Gibbs sampling. Through this two-stage sampling, the framework can handle CT volumes with an insufficient imaging range, in which many target landmarks are out of view (Fig. 3(b)). The framework can also work correctly even if many single-landmark detectors fail to detect their targets (Fig. 3(c)).

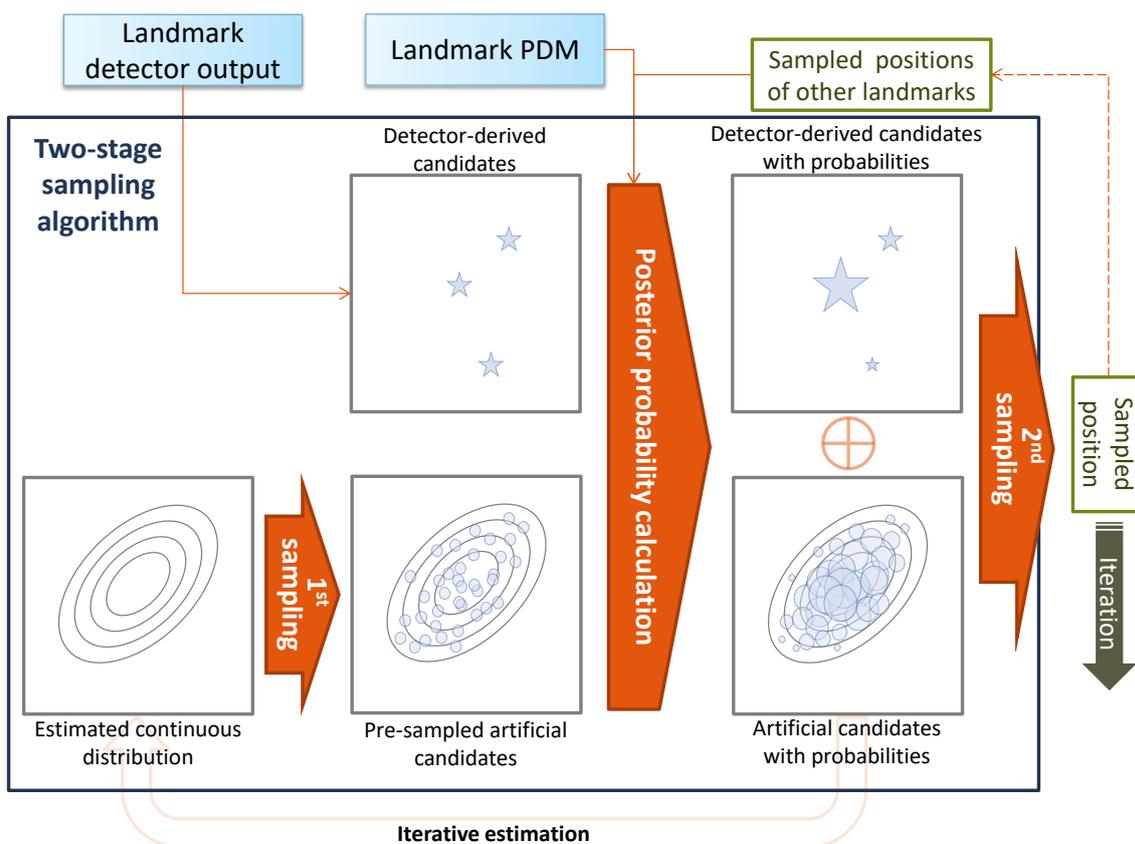


**Fig. 3.** Definitions of missing landmarks in this study. The proposed algorithm can estimate the most probable landmark position even in such situations. (a) Surgical or anatomical defect. The corresponding anatomical structure is defective owing to an anatomical anomaly or postsurgical state. (b) Being out of imaging range. The target landmark

is out of the imaging range. (c) Detection failure. Although the corresponding anatomical structure is within the imaging range, the true landmark position is not detected by the detector responsible for the target landmark. Such a detection failure is automatically diagnosed by the two-stage sampling algorithm and the algorithm estimates the most probable position for the landmark using other landmark positions.



**Fig. 4.** Outline of the landmark detection system.



**Fig. 5.** Outline of the proposed two-stage sampling.

In this paper, I describe the results of testing the proposed method using 104 human torso CT datasets with various imaging ranges with or without intravenous contrast agent injection. A total of 197 landmarks were detected, and their detection accuracies were evaluated.

**Table 1.** Comparison of landmark detection approaches.

	Landmarks	Strategy	Region	Methodology
(Potesil et al., 2011) [1]	22	simultaneous	torso	pictorial structure
(Liu and Zhou, 2012) [4]	63	sequential	torso	submodular optimization
(Major et al., 2013) [6]	23	simultaneous	spine	MRF
(Donner et al., 2013) [7]	57	simultaneous	whole body	MRF
(Glocker et al., 2012) [8]	24	simultaneous	vertebral body	HMM
(Glocker et al., 2013) [9]	24	simultaneous	vertebral body	random forest
proposed	197	simultaneous	whole body	Gibbs sampling, importance sampling

## 2.2. Methods

An outline of the proposed method is shown in Fig. 4. The method is composed of two phases: the independent landmark candidate detection phase and the combinatorial optimization phase. In the former, each landmark is independently searched for by a corresponding detector. Each

detector is trained for and optimized to the target landmark and outputs a list of candidate points. Then, the combinatorial optimization algorithm determines the final result. The final detection result of each landmark may be one of the candidate points, or it may be the point with the maximal posterior probability estimated by the probability distribution estimator.

I formulated the entire process of multiple landmark detection as a Bayesian estimation problem. Firstly, each landmark detector is modeled using a parametric stochastic model. This model can estimate not only the probability of each generated candidate being true positive (TP) but also the probability of all candidates being FPs. Secondly, the probability distribution of the spatial configuration of all landmarks in the human body (L-PDM) is also modeled and learned using training datasets. Thirdly, using detector models for all landmarks and the L-PDM together, the final combinatory optimization process is formulated as a maximum a posteriori (MAP) estimation problem and is solved by Markov chain Monte Carlo (MCMC) and simulated annealing algorithms. In the solving process, a two-stage sampling approach is used to efficiently estimate the positions of landmarks that are not detected by the corresponding detectors (owing to limited detector sensitivity, being out of the imaging range or an anatomical or surgical defect of the target landmark). In this approach, a number of artificial candidates are presampled from an estimated continuous probability distribution in each MCMC sampling cycle (Fig. 5). This pre-sampling process enables the proposed algorithm to estimate missing landmark positions even when the missing landmarks are outside the imaging range. Details of the detector model, L-PDM and the combinatory optimization solving process are described in the rest of this section.

### 2.2.1 Single-landmark detector

In the first stage, each landmark is independently searched for by a landmark-dedicated detector. Each detector searches for its target landmark from the entire CT volume and outputs a list of candidate positions. The detector used in this study is described in this section. The general stochastic model of an arbitrary detector will be described later.

Each detector consists of two components: A) an initial detector that functions by the sliding window method and B) a Madaboost-based [12] classifier for estimating the probabilities of detected candidates being TPs and for eliminating FPs. In B), a total of 482 image features are used (Table 2).

The detector used in this study is based on the work of [13] and [14]. For further details I referred readers to these two publications. Note that my MCMC-based method can handle any arbitrary detector that outputs a series of landmark candidate points and their likelihood values (as described later).

**Table 2.** Appearance-derived features used in Madaboost-based candidate likelihood score determination.

60 local-appearance-model-derived features: (Principal-component score, residual L2 norm, Mahalanobis distance from the mean) × Number of eigenvectors used to compose the model subspace (min. 1 to max. 20)
342 Haar features (Tu et al., 2006) 19 types of rectangular solid mask combination × 9 ROI cube sizes × 2 preprocessing (original volume or top-hat-filtered with 4 mm kernel radius)
40 Hu-moment features (Prokop and Reeves, 1992) 5 types of moments × 4 sizes of spherical ROI (2, 4, 6, 8 mm) × 2 preprocessing (original volume or top-hat-filtered with 4 mm kernel radius)
32 Hessian matrix-derived features 4 types (mean and Gaussian curvatures, shape index, curvedness) × 4 sizes of Gaussian smoothing $\sigma$ (2, 4, 6, 8 mm) × 2 preprocessing (original volume or top-hat-filtered with 4 mm kernel radius)
8 DoG features (Lowe, 2004) 4 pairs of Gaussian smoothers $\sigma$ : (2,4), (4,6), (6,8), (8, 10) mm × 2 preprocessing (original volume or top-hat-filtered with 4 mm kernel radius)
Total = 482

## 2.2.2 General stochastic model of landmark detector

I generally defined a landmark detector as a function whose input is a certain 3-D volume  $V: \mathbb{R}^3 \rightarrow \mathbb{R}$  and whose output is a set  $S = \{\theta_i\}$  where each element  $\theta_i = (\mathbf{c}_i, u_i), i = 1, 2, \dots, |S|$  is a pair of the  $i$ th candidate position  $\mathbf{c}_i \in \mathbb{R}^3$  and its detector-outputted likelihood score  $u_i \in \mathbb{R}$ . A detector can output an arbitrary number of candidates for the given volume.

### 2.2.2.1 Single-candidate model

First of all, a single candidate generated from a landmark detector is statistically modeled. Suppose that the tolerance of the detection distance error is  $d_{tol}$ . Then, each candidate is judged as TP if and only if  $\mathbf{c}_i$  is within the sphere whose center is the true landmark position and whose radius is  $d_{tol}$ . Let this sphere be  $R_{true}$  (Fig. 6). Additionally, let  $p_{true}(u) \equiv p(\mathbf{c}_i \in R_{true} | u_i = u)$  be the conditional probability of any

candidate being TP when its corresponding likelihood score is  $u$ . The conditional probability of it being FP (i.e., anywhere outside  $R_{true}$ ) must be  $p_{false}(u) = 1 - p_{true}(u)$ .

Then two categories of FPs are introduced in this study; pseudo-TPs and other FPs (Fig. 6). The reason for this categorization is to distinguish failed detections from inevitable detections of anatomical structures with virtually the same shape as the target landmark. In particular, the spinal column and ribs have a couple of such indistinguishable features. Thus, I divide the FPs into pseudo-TPs and other FPs and define their conditional probabilities as  $p_{pseudo}(u)$  and  $p_{other}(u)$ , respectively. These satisfy  $p_{false}(u) = p_{pseudo}(u) + p_{other}(u)$ .



**Fig. 6.** Example of the TP region and pseudo-TP regions. The pseudo-TP regions are regions around any confusing structures (i.e., the 1st, 3rd and 4th lumbar intervertebral disks), that can hardly be distinguished from the true landmark (i.e., the 2nd lumbar intervertebral disk).

To establish a statistical model for a detector, two assumptions are made: (a), the log odds ratio of TPs + pseudo-TPs to other FPs can be represented

as a linear function of  $u$ :

$$\frac{p_{true}(u) + p_{pseudo}(u)}{p_{other}(u)} = \exp(a_0 + a_1 u), \quad (1)$$

and (b) the odds ratio of TPs to pseudo-TPs is a constant:

$$\frac{p_{true}(u)}{p_{pseudo}(u)} = a_2, \quad (2)$$

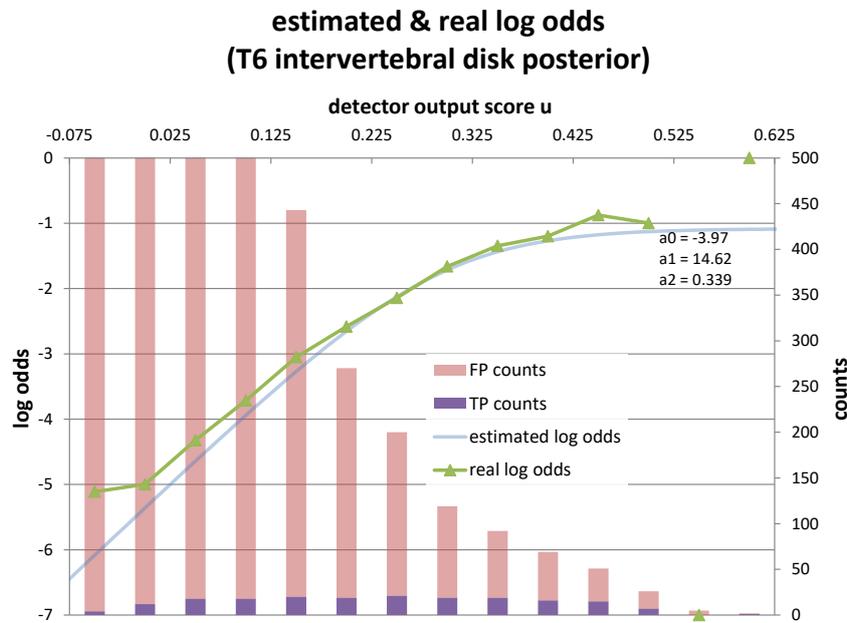
where  $\mathbf{a} = (a_0, a_1, a_2)$  is a model parameter. The first assumption means that the detector TP/FP odds can be parametrized using a logistic model (i.e., log-odds ratio is modeled as a linear function of the variable, that is, the detector-generated likelihood  $u$ ). The second assumption is that each detector cannot distinguish TPs from pseudo-TPs and the detection ratio of them are independent of the detector-generated likelihood  $u$ . Using this model,  $p_{true}(u)$  can be modeled as

$$p_{true}(u) = \frac{a_2}{1 + a_2} \cdot \frac{\exp(a_0 + a_1 u)}{1 + \exp(a_0 + a_1 u)}. \quad (3)$$

Practically, the model parameter  $\mathbf{a}$  has to be estimated in advance. In this study,  $\mathbf{a}$  is estimated from training datasets. After each detector is trained using the training datasets, the detector is in turn applied to all the training datasets themselves. Then the detection results from all datasets are added up to compose a pair of score sets  $\mathbf{u}_{TP}^{training}$  and  $\mathbf{u}_{FP}^{training}$ . The elements of  $\mathbf{u}_{TP}^{training}$  and  $\mathbf{u}_{FP}^{training}$  are the detector likelihood scores of TP and FP candidates (judged using ground-truth landmark positions), respectively. Then, the optimal parameter  $\hat{\mathbf{a}}$  can be estimated by a maximal likelihood (ML) method as follows:

$$\hat{\mathbf{a}} = \arg \max_{\mathbf{a}} \left\{ \sum_{u \in \mathbf{u}_{TP}^{training}} \log p_{true}(u) + \sum_{u \in \mathbf{u}_{FP}^{training}} \log(1 - p_{true}(u)) \right\}. \quad (4)$$

An example of the estimated log odds  $\log \frac{p_{true}(u)}{1-p_{true}(u)}$  for a spinal landmark is shown in Fig. 7.



**Fig. 7.** Example of model-estimated TP/FP log odds curve (sky blue). A real log odds curve (green), which is calculated from a real detector output histogram, is also shown. Note that the estimated curve closely fits well to the real curve.

### 2.2.2.2 Multiple-candidate model

On the basis of the single-candidate detector model described above,

another multiple-candidate detector model is built. Using the single candidate model,  $p_{true}(u) = p(\mathbf{c}_i \in R_{true} | u_i)$  was calculated in the previous section. This means that the probability of each candidate *without any knowledge of the other candidates* was calculated. Here I calculate the probability of each candidate considering all the candidates generated by the detector.

The details of the calculation are provided in Appendix 1. It was found that the probability of each candidate  $\mathbf{c}_i$  being TP when considering all the candidate likelihoods  $\mathbf{u} = (u_1, u_2, \dots, u_{|S|})$  can be calculated as follows:

$$p(\mathbf{c}_k \in R_{true} | \mathbf{u}) = \frac{1}{C} \cdot \frac{1 - p_0}{|S|} \cdot r_{prior}^{-1} \cdot \frac{p_{true}(u_k)}{1 - p_{true}(u_k)} \quad (5)$$

Here, the constants  $p_0$  and  $r_{prior}$  are parameters (determining certain prior probabilities) and have to be determined in advance. In this study  $p_0 = 0.02$  and  $r_{prior} = 0.01$  are used.  $C$  is a regularization term that makes the sum of the probabilities equal to 1.

On the other hand, the probability that all the candidate positions are false is also calculated using the model. That is,

$$p(\mathbf{c}_i \notin R_{true}, \forall i | \mathbf{u}) = \frac{1}{C} \cdot p_0 \quad (6)$$

where  $C = p_0 + \sum_{k=1}^{|S|} \frac{1-p_0}{|S|} \cdot r_{prior}^{-1} \cdot \frac{p_{true}(u_k)}{1-p_{true}(u_k)}$  is the same regularization term.

### 2.2.3 Landmark point distribution model

The reliable detection of multiple landmarks cannot be achieved only using the single detector models described above. Therefore, the spatial

configuration of all landmarks must be considered. In this section, I introduced two types of landmark point distribution model (L-PDM) for this purpose.

Consider the set of all landmark points as a random variable set  $\mathbf{X} = (\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_L)$ . Here,  $L$  is the number of landmarks,  $\mathbf{x}_l$  is a column vector of the coordinates of the  $l$ th landmark and column vector  $\mathbf{X}$  is the concatenated form of all  $\mathbf{x}_l$ . Thus,  $\mathbf{X}$  represents one particular spatial configuration of all landmarks in the human body. When  $\mathbf{X}$  is regarded as a random variable set, its probability distribution  $p(\mathbf{X})$  can be estimated from a finite number of training datasets. I define this distribution using an L-PDM. This L-PDM is defined as the function  $p: \mathbb{R}^{3L} \rightarrow \mathbb{R}$ , which represents the probability distribution of  $\mathbf{X}$ .

In this study, two different types of L-PDM were constructed from training datasets and their performances were compared. One is defined by a coordinate-based multivariate Gaussian distribution similar to that used in [15]. The other is my proposed L-PDM defined using a multivariate Gaussian distribution whose variables are inter-landmark logarithmic distances.

### 2.2.3.1 Coordinate-based L-PDM

Firstly, I constructed a PDM by simply estimating the probability distribution of  $\mathbf{X}$  as a Gaussian distribution with  $3L$  variables. Suppose that the ground-truth landmark points in the training datasets are manually inputted and rigidly registered in advance. Then, the sample average  $\bar{\mathbf{X}}$  and sample covariance matrix  $\text{Cov}(\mathbf{X})$  are calculated from these datasets. From  $\text{Cov}(\mathbf{X})$ , the precision matrix (the inverse of the covariance matrix)  $\Theta_{coord}$  is estimated using Tikhonov regularization [16] as

$$\Theta_{coord} = (\text{Cov}(\mathbf{X}) + \lambda_{coord}\mathbf{I})^{-1}, \quad (7)$$

where  $\lambda_{coord}$  is a constant used to control the strength of regularization. Using  $\bar{\mathbf{X}}$  and  $\Theta_{coord}$ , the PDF of the coordinate-based L-PDM  $p_{coord}(\mathbf{X})$  is defined as

$$p_{coord}(\mathbf{X}) = \frac{1}{Z_c} \exp\left(-\frac{1}{2}(\mathbf{X} - \bar{\mathbf{X}})^t \Theta_{coord} (\mathbf{X} - \bar{\mathbf{X}})\right), \quad (8)$$

where  $Z_c$  is a regularization coefficient.

### 2.2.3.2 Distance-based L-PDM

The other L-PDM is defined as a Gaussian distribution whose variables are logarithmic distances between pairs of landmarks. Let

$$d_{l,l'} = \ln(\max\{|\mathbf{x}_l - \mathbf{x}_{l'}|, d_{min}\}) \quad (9)$$

be the logarithmic distance between the  $l$ th and  $l'$ th candidate landmark points. Here,  $d_{min}$  is a constant used to suppress large negative logarithmic value. In this study  $d_{min} = \frac{1}{2}d_{tol}$  was used. Then, each distance is normalized by its average  $E(d_{l,l'})$  and variance  $V(d_{l,l'})$ , which are estimated from the training datasets. The normalized distance  $g_{l,l'}$  is given by

$$g_{l,l'} = \frac{d_{l,l'} - E(d_{l,l'})}{\sqrt{V(d_{l,l'})}}. \quad (10)$$

Let the vector  $\mathbf{G} = (g_{1,2} \ g_{1,3} \ \dots \ g_{1,L} \ g_{2,3} \ \dots \ g_{l,l'} \ \dots \ g_{(L-1),L})^t$  be a concatenated form of the normalized logarithmic distances between all the landmark pairs. Note that  $\mathbf{G}$  is a function of  $\mathbf{X}$  and that it has  $L(L-1)/2$  elements. Suppose that the sample covariance matrix of  $\mathbf{G}$ ,

namely,  $\text{Cov}(\mathbf{G})$ , is calculated from the training datasets. Then, the precision matrix of  $\mathbf{G}$  can be estimated by Tikhonov regularization as  $\Theta_{dist} = (\text{Cov}(\mathbf{G}) + \lambda_{dist} \mathbf{I})^{-1}$  ( $\lambda_{dist}$  is a constant). Then, the PDF of the distance-based L-PDM  $p_{dist}(\mathbf{X})$  is defined as

$$p_{dist}(\mathbf{X}) = \frac{1}{Z_d} \exp\left(-\frac{1}{2} \mathbf{G}^t \Theta_{dist} \mathbf{G}\right). \quad (11)$$

#### 2.2.4 Maximum *a posteriori* estimation

In the next stage, the best combination of all detector-generated candidates is searched for. Supposing that a total of 197 landmarks are to be detected and each detector outputs 100 candidates with their estimated likelihoods. Then the number of possible combinations becomes  $100^{197}$ . To solve this large combinatorial optimization problem, I use an approximate method in which Gibbs sampling and simulated annealing are combined [11].

Let  $I_l = (\mathbf{c}_l^1, \mathbf{c}_l^2, \dots, \mathbf{c}_l^{|S_l|}, p_l^1, p_l^2, \dots, p_l^{|S_l|}, p_l^\times)$  be the  $l$ th landmark-detector-derived information set (not only the detector output  $S_l$  but also the probabilities  $p_l^1, p_l^2, \dots$  which are calculated using the stochastic detector model). Here,  $\mathbf{c}_l^i$  is the  $i$ th candidate for the  $l$ th landmark and  $p_l^i$  is its corresponding posterior probability estimated as  $p(\mathbf{c}_i \in R_{true} | \mathbf{u})$  by Eq. (5). On the other hand,  $p_l^\times$  is the posterior probability of all candidates being FP, which can be estimated as  $p(\mathbf{c}_i \notin R_{true}, \forall i | \mathbf{u})$  by Eq. (6). That is,

$$p_l^i = \frac{1}{C_l} \cdot \frac{1 - p_0}{|S_l|} \cdot r_{prior}^{-1} \cdot \frac{p_{true;l}(\mathbf{u}_{i;l})}{1 - p_{true;l}(\mathbf{u}_{i;l})} \quad (12)$$

$$p_l^\times = \frac{1}{C_l} \cdot p_0, \quad (13)$$

where  $|S_l|$  is the number of candidates,  $u_{i;l}$  is the detector-generated likelihood for the  $i$ th candidate,  $p_{true;l}(\cdot)$  is the detector-model-derived probability function (Eq. (3)), and  $C_l$  is the regularization term for the  $l$ th landmark.

Given the entire detector-derived information set  $I = (I_1, I_2, \dots, I_L)$ , the proposed algorithm searches for the best candidate using an L-PDM. The L-PDM is a statistical model of the spatial distribution of landmarks in the human body and is trained using training datasets. My combinatory optimization framework searches for the most probable landmark position set  $\hat{\mathbf{X}}$  under the condition of the given detector information  $I$ . In other words, it solves the MAP problem  $\hat{\mathbf{X}} = \arg \max_{\mathbf{X}} p(\mathbf{X}|I)$ .

Applying Bayes' theorem to the posterior probability  $p(\mathbf{X}|I)$ ,

$$\begin{aligned} p(\mathbf{X}|I) &= \frac{p(I|\mathbf{X}) \cdot p(\mathbf{X})}{p(I)} \\ &\propto p(I|\mathbf{X}) \cdot p(\mathbf{X}) \end{aligned} \quad (14)$$

is satisfied. Note that the denominator  $p(I)$  is independent of  $\mathbf{X}$  and can be ignored when maximizing  $p(\mathbf{X}|I)$  in terms of  $\mathbf{X}$ .

In this study, one of the L-PDMs (defined in Section 2.3) is used as  $p(\mathbf{X})$ . To balance the weight of each L-PDM with the weight of the candidate probabilities, the L-PDMs are modified by a constant  $c_{detector}$  to give  $p(\mathbf{X}) = p_{coord}(\mathbf{X})^{\frac{1}{c_{detector}}}$  or  $p(\mathbf{X}) = p_{dist}(\mathbf{X})^{\frac{1}{c_{detector}}}$ . Here,  $c_{detector}$  is a constant to be determined in advance.

I assume that all detector outputs  $I_1, I_2, \dots, I_L$  are independent of each other. Furthermore, I also assume that, when the target landmark position is given as a condition, the corresponding detector output is conditionally

independent of the other landmark positions. This means that each detector output is only dependent on the position of the target landmark. Under this assumption, the term  $p(I|\mathbf{X})$  can be factorized as  $\prod_{l=1}^L p(I_l|\mathbf{x}_l)$ . Applying Bayes' theorem again,

$$\begin{aligned} p(I|\mathbf{X}) &= \prod_{l=1}^L p(I_l|\mathbf{x}_l) \\ &= \prod_{l=1}^L \frac{p(\mathbf{x}_l|I_l) \cdot p(I_l)}{p(\mathbf{x}_l)} \end{aligned} \quad (15)$$

is satisfied. Here,  $p(I_l)$  is independent of  $\mathbf{X}$  and can be ignored. The term  $p(\mathbf{x}_l)$  is the probability distribution of the  $l$ th landmark *without any prior knowledge* and is assumed to be uniform in the entire space in this study. Therefore, using Eqs. (14) and (15), the MAP estimation problem can be formulated as

$$\begin{aligned} \hat{\mathbf{X}} &= \arg \max_{\mathbf{X}} p(\mathbf{X}|I) \\ &= \arg \max_{\mathbf{X}} p(\mathbf{X}) \cdot p(I|\mathbf{X}) \\ &= \arg \max_{\mathbf{X}} p(\mathbf{X}) \cdot \prod_{l=1}^L p(I_l|\mathbf{x}_l) \\ &= \arg \max_{\mathbf{X}} p(\mathbf{X}) \cdot \prod_{l=1}^L p(\mathbf{x}_l|I_l). \end{aligned} \quad (16)$$

The purpose of this formulation is to factorize the MAP problem into the L-PDM  $p(\mathbf{X})$  and the single landmark terms  $p(\mathbf{x}_l|I_l)$ . Note that  $p(\mathbf{X}|I)$  cannot be directly factorized as  $p(\mathbf{X}|I) = \prod_{l=1}^L p(\mathbf{x}_l|I_l)$  because the landmark positions  $\mathbf{x}_l$  are not independent of each other. Instead, the correlations among the landmark positions are modeled by the L-PDM  $p(\mathbf{X})$

itself.

In Eq. (16), the term  $p(\mathbf{x}_l|I_l)$  represents the PDF of the landmark position  $\mathbf{x}_l$  when the detector output  $I_l$  is given. Here, suppose that  $p_l^\times = 0$  for the purpose of illustration. This means that one of the detector-generated candidate points  $\{\mathbf{c}_l^1, \mathbf{c}_l^2, \dots\}$  must be the true answer. Note that the assumption  $p_l^\times = 0$  cannot be satisfied in the real problem, because in my stochastic detector model each detector has a nonzero probability of missing landmark (owing to being out of the imaging range, limited detection sensitivity or an anatomical/surgical defect). Under the assumption of  $p_l^\times = 0$ , the term  $p(\mathbf{x}_l|I_l)$  can be computed from Eq. (12) as

$$p(\mathbf{x}_l|I_l) = \begin{cases} p_l^i, & \text{if } \mathbf{x}_l = \mathbf{c}_l^i \text{ for any } i \\ 0, & \text{if } \mathbf{x}_l \neq \mathbf{c}_l^i \text{ for all } i \end{cases}. \quad (17)$$

Therefore, the posterior probability in Eq. (16) becomes nonzero only at the candidate points  $\mathbf{x}_l \in \{\mathbf{c}_l^1, \mathbf{c}_l^2, \dots\}$ . This means that this MAP estimation is formulated as a simple combinatorial optimization problem if  $p_l^\times = 0$ .

Our goal is to estimate the target landmark position even when the detector output does not include the true candidate. When  $p_l^\times > 0$ , it is possible that no candidate point  $\mathbf{c}_l^i$  corresponds to the true landmark point  $\mathbf{x}_l$ . If this is the case,  $p(\mathbf{x}_l|I_l)$  will be nonzero even if  $\mathbf{x}_l \neq \mathbf{c}_l^i$  for all  $i$ . Therefore, I need to define  $p(\mathbf{x}_l|I_l)$  at any point in the domain  $\mathbf{x}_l \in \mathbb{R}^3$ . One possible way is to define  $p(\mathbf{x}_l|I_l)$  as a distribution that is uniform at any point other than the candidate points  $\{\mathbf{c}_l^1, \mathbf{c}_l^2, \dots\}$  (Fig. 8(b)). Consider a probability space  $\Omega$  that is sufficiently larger than both the subject human body and the CT volume. Additionally, suppose that the probability  $p(\mathbf{x}_l|I_l)$  is constant everywhere in  $\Omega$  except for the candidate points  $\{\mathbf{c}_l^1, \mathbf{c}_l^2, \dots\}$ . Then, the probability distribution  $p(\mathbf{x}_l|I_l)$  can be represented

as

$$p(\mathbf{x}_l|I_l) = \frac{1}{|\Omega|} p_l^\times + \sum_{i=1}^{|\mathcal{S}_l|} \delta(\mathbf{x}_l - \mathbf{c}_l^i) \cdot p_l^i \quad (18)$$

where  $\delta(\cdot)$  is the Dirac delta function and  $|\Omega|$  is the volume of  $\Omega$ .

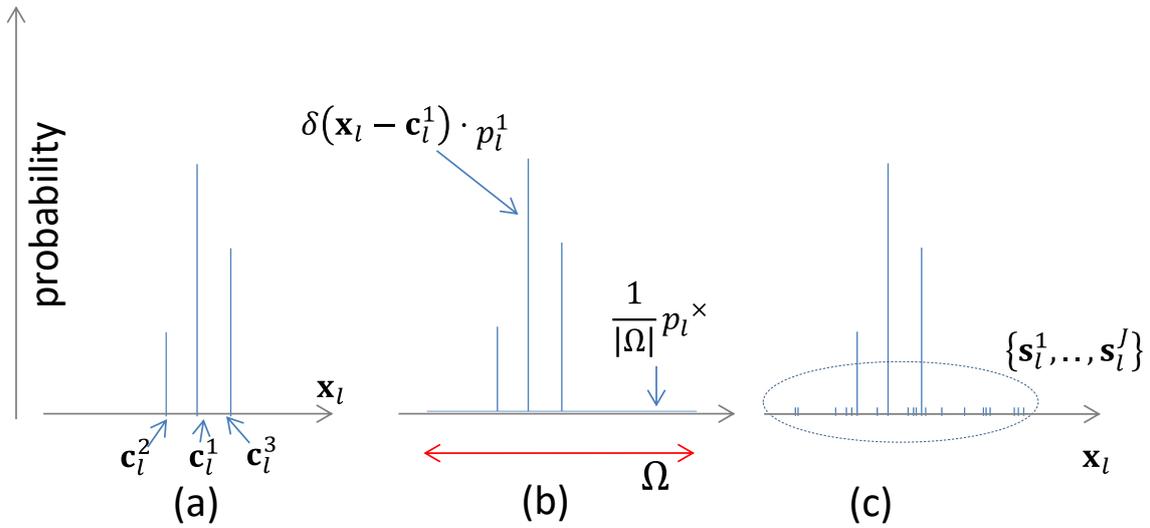
Since the PDF in Eq. (18) is not discrete, the MAP estimation cannot be formulated as a combinatorial optimization problem. To reduce this problem to a combinatory optimization one, I discretize and replace the domain  $\Omega$  with a large number of  $J$  artificial candidates  $\{\mathbf{s}_l^1, \dots, \mathbf{s}_l^J\}$  (Fig. 8(c)). That is,

$$p(\mathbf{x}_l|I_l) = \sum_{j=1}^J \frac{1}{J} \cdot \delta(\mathbf{x}_l - \mathbf{s}_l^j) \cdot p_l^\times + \sum_{i=1}^{|\mathcal{S}_l|} \delta(\mathbf{x}_l - \mathbf{c}_l^i) \cdot p_l^i, \quad (19)$$

or, in the discretized form,

$$p(\mathbf{x}_l|I_l) = \begin{cases} p_l^i, & \text{if } \mathbf{x}_l = \mathbf{c}_l^i \text{ for any } i \\ \frac{1}{J} \cdot p_l^\times, & \text{if } \mathbf{x}_l = \mathbf{s}_l^j \text{ for any } j \\ 0, & \text{otherwise} \end{cases} \quad (20)$$

In this study, the artificial candidates  $\{\mathbf{s}_l^1, \dots, \mathbf{s}_l^J\}$  are iteratively sampled from the estimated distribution of  $\mathbf{x}_l$  by the two-stage sampling framework (Fig. 5). This will be described in detail in Section 2.6.



**Fig. 8.** Schemas for explaining the probability distributions. (a) When  $p_l^\times = 0$ , the probability function  $p(\mathbf{x}_l|I_l)$  is discrete and becomes nonzero only at  $\mathbf{x}_l = \mathbf{c}_l^i$ . (b) When  $p_l^\times > 0$ , the distribution  $p(\mathbf{x}_l|I_l)$  is represented by the sum of a constant function and a series of delta functions. (c) Discretized probability distribution with randomly sampled artificial candidates  $\{\mathbf{s}_l^1, \dots, \mathbf{s}_l^J\}$  (Eq. (25)). In reality, the artificial candidates are sampled from an estimated multivariate Gaussian distribution, not from a uniform distribution (as described in Section 2.2.6).

### 2.2.5 Combinatory optimization with Gibbs sampling and simulated annealing

In this study, I used a Gibbs-sampling and simulated-annealing-based method [11] to solve the MAP estimation problem (Eq. (16)) and find the best combination of  $\mathbf{X} = (\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_L)$ , where  $\mathbf{x}_l \in \{\mathbf{c}_l^1, \mathbf{c}_l^2, \dots, \mathbf{c}_l^{|\mathcal{S}_l|}, \mathbf{s}_l^1, \mathbf{s}_l^2, \dots, \mathbf{s}_l^J\}$ ,  $1 \leq l \leq L$ . I chose this method because it can handle the huge extent of the problem domain in which my algorithm

has to search.

Gibbs sampling [17] is a simple algorithm for selecting samples from a given distribution. Let the initial condition of the algorithm be  $\mathbf{X}^{(0)} = (\mathbf{x}_1^{(0)}, \mathbf{x}_2^{(0)}, \dots, \mathbf{x}_l^{(0)}, \dots, \mathbf{x}_L^{(0)})$  and the condition after  $n$  iterations be  $\mathbf{X}^{(n)}$ . Then, the current estimate of the landmark point  $\mathbf{x}_l^{(n)}$  is updated by random sampling from the conditional distribution of the given PDF for  $l = 1, 2, 3, \dots, L, 1, 2, 3, \dots, L, \dots$ , in a sequential manner.

Suppose that the  $l$ th landmark position  $\mathbf{x}_l$  is being sampled in the  $n$ th iteration. Before the sampling, the current mode set of the landmark positions is  $\mathbf{x}_1 = \mathbf{x}_1^{(n)}, \dots, \mathbf{x}_{l-1} = \mathbf{x}_{l-1}^{(n)}, \mathbf{x}_l = \mathbf{x}_l^{(n-1)}, \mathbf{x}_{l+1} = \mathbf{x}_{l+1}^{(n-1)}, \dots, \mathbf{x}_L = \mathbf{x}_L^{(n-1)}$ . Then  $\mathbf{x}_l$  is sampled from the PDF  $p(\mathbf{X}) \cdot \prod_{l=1}^L p(\mathbf{x}_l | I_l)$  (Eq. (16)) under the condition of  $\mathbf{x}_1 = \mathbf{x}_1^{(n)}, \dots, \mathbf{x}_{l-1} = \mathbf{x}_{l-1}^{(n)}, \mathbf{x}_{l+1} = \mathbf{x}_{l+1}^{(n-1)}, \dots, \mathbf{x}_L = \mathbf{x}_L^{(n-1)}$ . That is,

$$\begin{aligned}
 & p_{\text{conditional}}(\mathbf{x}_l) \\
 &= p(\mathbf{X}) \\
 & \cdot \prod_{l=1}^L p(\mathbf{x}_l | I_l) \Big|_{\mathbf{x}_1 = \mathbf{x}_1^{(n)}, \dots, \mathbf{x}_{l-1} = \mathbf{x}_{l-1}^{(n)}, \mathbf{x}_{l+1} = \mathbf{x}_{l+1}^{(n-1)}, \dots, \mathbf{x}_L = \mathbf{x}_L^{(n-1)}} \\
 & \propto p(\mathbf{x}_1^{(n)}, \dots, \mathbf{x}_{l-1}^{(n)}, \mathbf{x}_l, \mathbf{x}_{l+1}^{(n-1)}, \dots, \mathbf{x}_L^{(n-1)}) \cdot p(\mathbf{x}_l | I_l).
 \end{aligned} \tag{21}$$

For the sake of convenience, let  $\mathbf{X}^{(n)} \setminus \mathbf{x}_l = (\mathbf{x}_1^{(n)}, \dots, \mathbf{x}_{l-1}^{(n)}, \mathbf{x}_{l+1}^{(n-1)}, \mathbf{x}_L^{(n-1)})$  be the current landmark mode set other than  $\mathbf{x}_l$ . Then, let

$$\begin{aligned}
 & p(\mathbf{x}_1^{(n)}, \dots, \mathbf{x}_{l-1}^{(n)}, \mathbf{x}_l, \mathbf{x}_{l+1}^{(n-1)}, \dots, \mathbf{x}_L^{(n-1)}) \\
 & \equiv p_{L-PDM}(\mathbf{x}_l; \mathbf{X}^{(n)} \setminus \mathbf{x}_l)
 \end{aligned} \tag{22}$$

be the value of L-PDM  $p(\mathbf{X})$  for the condition of  $\mathbf{X}^{(n)} \setminus \mathbf{x}_l$ . Using Eq. (22), Eq. (21) can be rewritten as

$$p_{\text{conditional}}(\mathbf{x}_l) = \frac{1}{Z} \cdot p_{L\text{-PDM}}(\mathbf{x}_l; \mathbf{X}^{(n)} \setminus \mathbf{x}_l) \cdot p(\mathbf{x}_l | I_l), \quad (23)$$

where  $Z$  is a regularization factor. Note that calculation of the real value of  $Z$  is not necessary in the Gibbs sampling algorithm. Using Eqs. (23) and (19), the posterior probability of each candidate in  $c_l^+ = \{\mathbf{c}_l^1, \dots, \mathbf{c}_l^{|\mathcal{S}_l|}, \mathbf{s}_l^1, \dots, \mathbf{s}_l^J\}$  is calculated as

$$\begin{aligned} p_{\text{conditional}}(\mathbf{x}_l = \mathbf{c}_l^i) &= \frac{1}{Z} \cdot p_{L\text{-PDM}}(\mathbf{c}_l^i; \mathbf{X}^{(n)} \setminus \mathbf{x}_l) \cdot p_l^i \\ p_{\text{conditional}}(\mathbf{x}_l = \mathbf{s}_l^j) &= \frac{1}{Z} \cdot p_{L\text{-PDM}}(\mathbf{s}_l^j; \mathbf{X}^{(n)} \setminus \mathbf{x}_l) \cdot \frac{1}{J} \cdot p_l^{\times} \end{aligned} \quad (24)$$

From  $p_{\text{conditional}}(\cdot)$ ,  $\mathbf{x}_l^{(n)}$  is randomly sampled from  $c_l^+$  in accordance with this discrete posterior probability function. After all  $l$  updates are finished,  $n$  is incremented and the entire update process is iterated.

The algorithm can be described in a pseudocode as follows:

for  $n=1$  to *total\_iteration\_number*

    for  $l=1$  to  $L$

        sample  $\mathbf{x}_l^{(n)}$  from the conditional probability distribution in

Eq. (24)

    end for

end for

This algorithm outputs a sequence of values  $\mathbf{X}^{(n)}$ , which are randomly

sampled from the target distribution  $p(\mathbf{X}|I)$ .

Additionally, a simulated annealing method is combined with Gibbs sampling to make the sampling sequence converge to the optimal point. In the annealing, a positive variable  $T$  is introduced as a temperature.  $T$  is first set at a very high value and then gradually lowered during Gibbs sampling. Before each sampling, the distribution  $p_{conditional}(\mathbf{x}_l)$  in Eq. (24) is modified using  $T$  to

$$p_{annealed}(\mathbf{x}_l = \mathbf{c}_l^i; T) = \frac{1}{Z'} \cdot \{p_{L-PDM}(\mathbf{c}_l^i; \mathbf{X}^{(n)} \setminus \mathbf{x}_l) \cdot p_l^i\}^{\frac{1}{T}} \quad (25)$$

$$p_{annealed}(\mathbf{x}_l = \mathbf{s}_l^j; T) = \frac{1}{Z'} \cdot \frac{1}{J} \cdot \{p_{L-PDM}(\mathbf{s}_l^j; \mathbf{X}^{(n)} \setminus \mathbf{x}_l) \cdot p_l^j\}^{\frac{1}{T}},$$

where  $Z'$  is another regularization term (dependent on  $T$ ). The factor  $1/J$  is not modified to avoid an excessive effect when  $T$  is very high or low.

|

In the simulated annealing, each  $\mathbf{x}_l^{(n)}$  is sequentially and repeatedly sampled from the modified distribution (Eq. (25)) by a Gibbs sampler. The sampling begins with a very high  $T$ , which decreases gradually and finally becomes so low that the system converges to the maximum point. After the annealing, the final  $\mathbf{X}^{(n)}$  is outputted as the combinatorial optimization result.

### 2.2.6 Proposed two-stage sampling

The key to the proposed method is how to sample missing landmarks

during the iterative Gibbs sampling. Each Gibbs sampling of  $\mathbf{x}_l^{(n)}$  from the distribution  $p_{conditional}(\mathbf{x}_l)$  (Eq. (23)) requires a set of pre-sampled artificial candidates  $\{\mathbf{s}_l^1, \dots, \mathbf{s}_l^J\}$ . Therefore, sampling of  $\{\mathbf{s}_l^1, \dots, \mathbf{s}_l^J\}$  is required prior to each Gibbs sampling. The sample  $\{\mathbf{s}_l^1, \dots, \mathbf{s}_l^J\}$  represents the missing landmarks (being out of imaging range, defect due to surgical/pathological changes, limited detector sensitivity, etc.). Therefore, the sampling points  $\{\mathbf{s}_l^1, \dots, \mathbf{s}_l^J\}$  can spread outside the imaging range border. It also means that the sampling of  $\{\mathbf{s}_l^1, \dots, \mathbf{s}_l^J\}$  cannot depend on any image appearance information. Although it is theoretically possible to sample  $\{\mathbf{s}_l^1, \dots, \mathbf{s}_l^J\}$  from a constant distribution whose domain is  $\Omega$ , this is extremely ineffective because a sampled point  $\mathbf{s}_l^i$  for which  $p_{L-PDM}(\mathbf{s}_l^i; \mathbf{X}^{(n)} \setminus \mathbf{x}_l)$  is small has little chance of being selected in the following Gibbs sampling. Thus, a more effective way to sample  $\{\mathbf{s}_l^1, \dots, \mathbf{s}_l^J\}$  is vital to the missing landmark handling property of the proposed method.

Here, I sample  $\{\mathbf{s}_l^1, \dots, \mathbf{s}_l^J\}$  from the target PDF  $p_{L-PDM}(\mathbf{x}_l; \mathbf{X}^{(n)} \setminus \mathbf{x}_l)$  itself. However, sampling directly from a complex PDF is too costly to perform iteratively. In this study, a novel two-stage sampling strategy is introduced, in which the sampling point  $\{\mathbf{s}_l^1, \dots, \mathbf{s}_l^J\}$  is repeatedly sampled from a probability distribution  $q_l(\mathbf{x}_l)$  that is similar to, but not equal to, the target PDF. Simultaneously, the distribution  $q_l(\mathbf{x}_l)$  is repeatedly re-estimated using the calculated values of the target PDF at the sampling points, that is,  $p_{L-PDM}(\mathbf{s}_l^j; \mathbf{X}^{(n)} \setminus \mathbf{x}_l)$ ,  $j = 1, 2, \dots, J$  [17].

The details of the importance sampling algorithm used in this study are

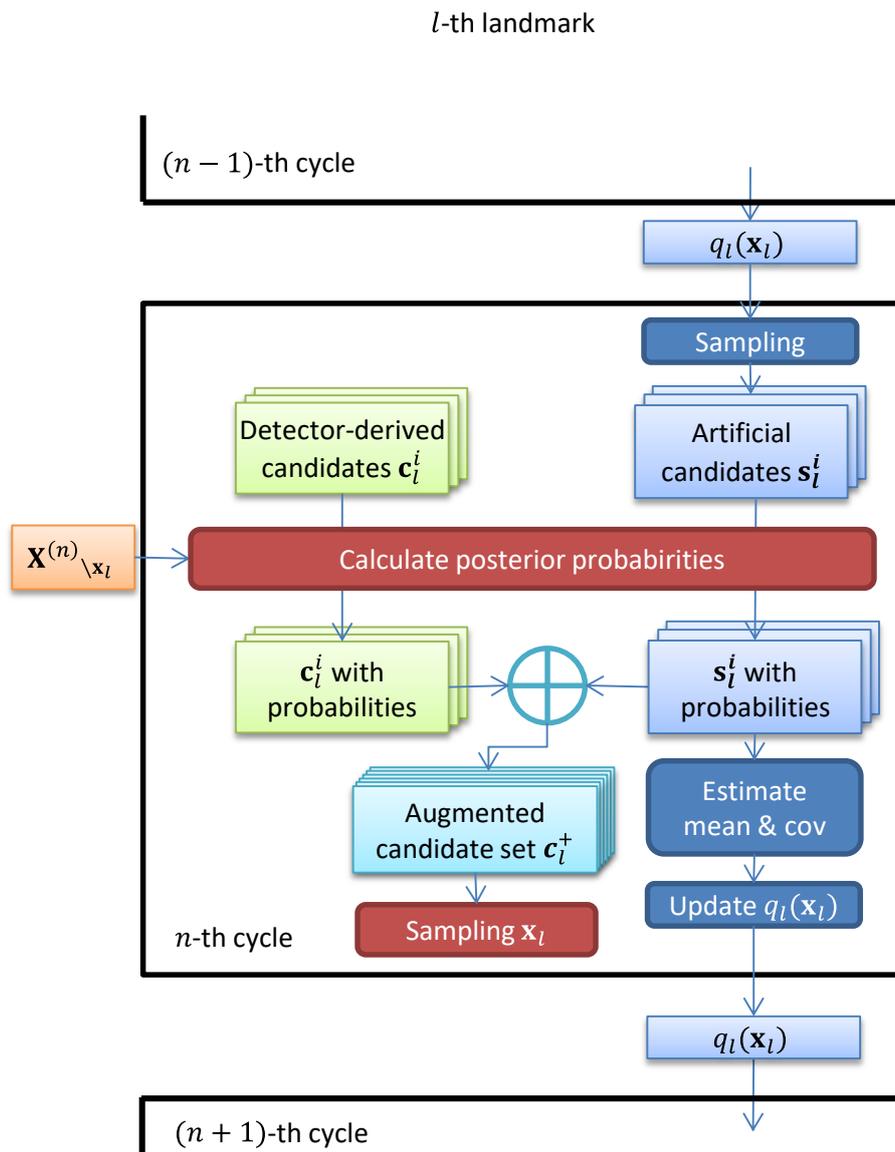
described in Appendix 2. In brief, the algorithm samples  $\{\mathbf{s}_l^1, \dots, \mathbf{s}_l^J\}$  from a three-dimensional Gaussian distribution  $q_l(\cdot)$  whose mean vector  $E$  and covariance matrix  $Cov$  are calculated as follows:

$$E = \frac{\sum_{j=1}^J \mathbf{s}_l^j \cdot \frac{p_{L-PDM}(\mathbf{s}_l^j; \mathbf{X}^{(n)} \setminus \mathbf{x}_l)^{1/T}}{q_l(\mathbf{s}_l^j)}}{\sum_{j=1}^J \frac{p_{L-PDM}(\mathbf{s}_l^j; \mathbf{X}^{(n)} \setminus \mathbf{x}_l)^{1/T}}{q_l(\mathbf{s}_l^j)}} \quad (26)$$

$$Cov = \alpha \cdot \frac{\sum_{j=1}^J (\mathbf{s}_l^j - E)(\mathbf{s}_l^j - E)^t \cdot \frac{p_{L-PDM}(\mathbf{s}_l^j; \mathbf{X}^{(n)} \setminus \mathbf{x}_l)^{\frac{1}{T}}}{q_l(\mathbf{s}_l^j)}}{\sum_{j=1}^J \frac{p_{L-PDM}(\mathbf{s}_l^j; \mathbf{X}^{(n)} \setminus \mathbf{x}_l)^{\frac{1}{T}}}{q_l(\mathbf{s}_l^j)}}.$$

Here  $\alpha$  is a parameter to be determined in advance. Note that the right side of Eq. (26) includes the sampled points  $\{\mathbf{s}_l^1, \dots, \mathbf{s}_l^J\}$  themselves and  $q_l(\cdot)$  itself. This means that the sampling of  $\{\mathbf{s}_l^1, \dots, \mathbf{s}_l^J\}$  from  $q_l(\cdot)$  and the recalculation of  $E$  and  $Cov$  are iteratively performed (Fig. 9). At the end of each Gibbs sampling cycle,  $E$  and  $Cov$  are recalculated using the current  $\{\mathbf{s}_l^1, \dots, \mathbf{s}_l^J\}$  and  $q_l(\cdot)$ , and then the distribution  $q_l(\cdot)$  is updated using  $E$  and  $Cov$ . This updated distribution  $q_l(\cdot)$  is used in the next cycle to sample the points  $\{\mathbf{s}_l^1, \dots, \mathbf{s}_l^J\}$ .

The pseudocode of the entire combinatorial optimization process is shown in Fig. 10.



**Fig. 9.** Flowchart of the proposed two-stage sampling.

Input:

$l_l = (N_l, \mathbf{c}_l^1, \dots, \mathbf{c}_l^{N_l}, p_l^1, \dots, p_l^{N_l}, p_l^{\bar{c}_l})$  ( $l = 1, 2, \dots, L$ ), Detector inputs

$p(\mathbf{X})$ , L-PDM

$T_0$ , initial temperature and  $r_T$ , cooling ratio

$\mathbf{g}$ , coordinates of volume center

$\beta$ , sufficiently large initialization parameter

Output:

The estimated landmark position set  $\hat{\mathbf{X}}$

1.  $\mathbf{X}^{(0)} = (\mathbf{x}_1^{(0)}, \mathbf{x}_2^{(0)}, \dots, \mathbf{x}_l^{(0)}, \dots, \mathbf{x}_L^{(0)}) \leftarrow (\mathbf{c}_1^1, \mathbf{c}_2^1, \dots, \mathbf{c}_l^1, \dots, \mathbf{c}_L^1)$ .

$q_l(\cdot) \leftarrow \mathcal{N}(\mathbf{g}, \beta \mathbf{I})$ ,  $l = 1, \dots, L$ .

$T \leftarrow T_0$ .

2. **FOR**  $n = 1, \dots, N_{iteration}$  :

**FOR**  $l = 1, \dots, L$ :

Sample  $\mathbf{s}_l^j \sim q_l(\cdot)$ ,  $j = 1, 2, \dots, J$ .

Calculate posterior likelihood of each candidate by

$$p_{annealed}(\mathbf{c}_l^i) = \frac{1}{Z_p} \cdot \{p_{L-PDM}(\mathbf{c}_l^i; \mathbf{X}^{(n)} \setminus \mathbf{x}_l) \cdot p_l^i\}^{\frac{1}{T}}$$

$$p_{annealed}(\mathbf{s}_l^j) = \frac{1}{Z_p} \cdot \frac{1}{J} \cdot \{p_{L-PDM}(\mathbf{s}_l^j; \mathbf{X}^{(n)} \setminus \mathbf{x}_l) \cdot p_l^{\times}\}^{\frac{1}{T}},$$

where  $Z_p = \sum_{i=1}^{N_l} \{p_{L-PDM}(\mathbf{c}_l^i; \mathbf{X}^{(n)} \setminus \mathbf{x}_l) \cdot p_l^i\}^{\frac{1}{T}} + \sum_{j=1}^J \frac{1}{J} \cdot \{p_{L-PDM}(\mathbf{s}_l^j; \mathbf{X}^{(n)} \setminus \mathbf{x}_l) \cdot p_l^{\times}\}^{\frac{1}{T}}$ .

Sample  $\mathbf{x}_l^{(n)} \sim p_{annealed}(\cdot)$ .

Update  $q_l(\cdot) \leftarrow$

$$\mathcal{N}\left(\frac{1}{Z_q} \cdot \sum_{j=1}^J \mathbf{s}_l^j \cdot \frac{p_{L-PDM}(\mathbf{s}_l^j; \mathbf{X}^{(n)} \setminus \mathbf{x}_l)^{\frac{1}{T}}}{q_l(\mathbf{s}_l^j)}, \frac{\alpha}{Z_q} \cdot \sum_{j=1}^J (\mathbf{s}_l^j - \mathbf{E}(\mathbf{x}_l)) (\mathbf{s}_l^j - \mathbf{E}(\mathbf{x}_l))^t \cdot \frac{p_{L-PDM}(\mathbf{s}_l^j; \mathbf{X}^{(n)} \setminus \mathbf{x}_l)^{\frac{1}{T}}}{q_l(\mathbf{s}_l^j)}\right),$$

where  $Z_q = \sum_{j=1}^J \frac{p_{L-PDM}(\mathbf{s}_l^j; \mathbf{X}^{(n)} \setminus \mathbf{x}_l)^{\frac{1}{T}}}{q_l(\mathbf{s}_l^j)}$ .

**End for**

$T \leftarrow T \cdot r_T$

**End for**

3. **Output**  $\hat{\mathbf{X}} \leftarrow \mathbf{X}^{(N_{iteration})}$ .

**Fig. 10.** Pseudocode of the proposed combinatorial optimization algorithm.

### 2.2.7 Evaluation

This study was approved by the Ethical Review Board of The University of Tokyo Hospital.

**Dataset A.** A total of 120 CT datasets were randomly collected from my CT examination database. Among them, 16 datasets showed an anomaly in the number of spinal bones and were thus excluded from this study. Thus, 104 CT datasets (Dataset A) were used in the following experiments. Among them, 50 were from healthy subjects without intravenous contrast agent injection with the neck-to-pelvis imaging range. The other 54 datasets were from diseased subjects with or without contrast agent injection, and the imaging range varied from the chest-to-pelvis region to the upper abdomen only. The voxel size was within the range from  $0.625 \times 0.625 \times 1.00$  mm to  $1.148 \times 1.148 \times 1.25$  mm. The positions of a total of 197 anatomically defined landmarks were manually inputted by medical experts. After that, a radiologist checked and modified (if necessary) all of them. These modified landmark positions were used as ground-truth positions. The landmarks were defined on soft tissue structures (25), the spinal column (120), pelvic bones (34) or other bones (18).

The experiments were performed using twofold cross-validation. Before training the detectors and L-PDMs, parameter optimization was performed. Both the training and parameter tuning of each detector and each L-PDM were performed using training datasets. Details of parameter optimization for the detectors are given in [13]. For the L-PDMs, two parameters were optimized:  $\lambda_{coord}$  or  $\lambda_{dist}$  (constants that determine the weight of the Tikhonov normalization term) and  $c_{detector}$  (the extent to which the optimization algorithm considers the likelihoods of candidates generated by

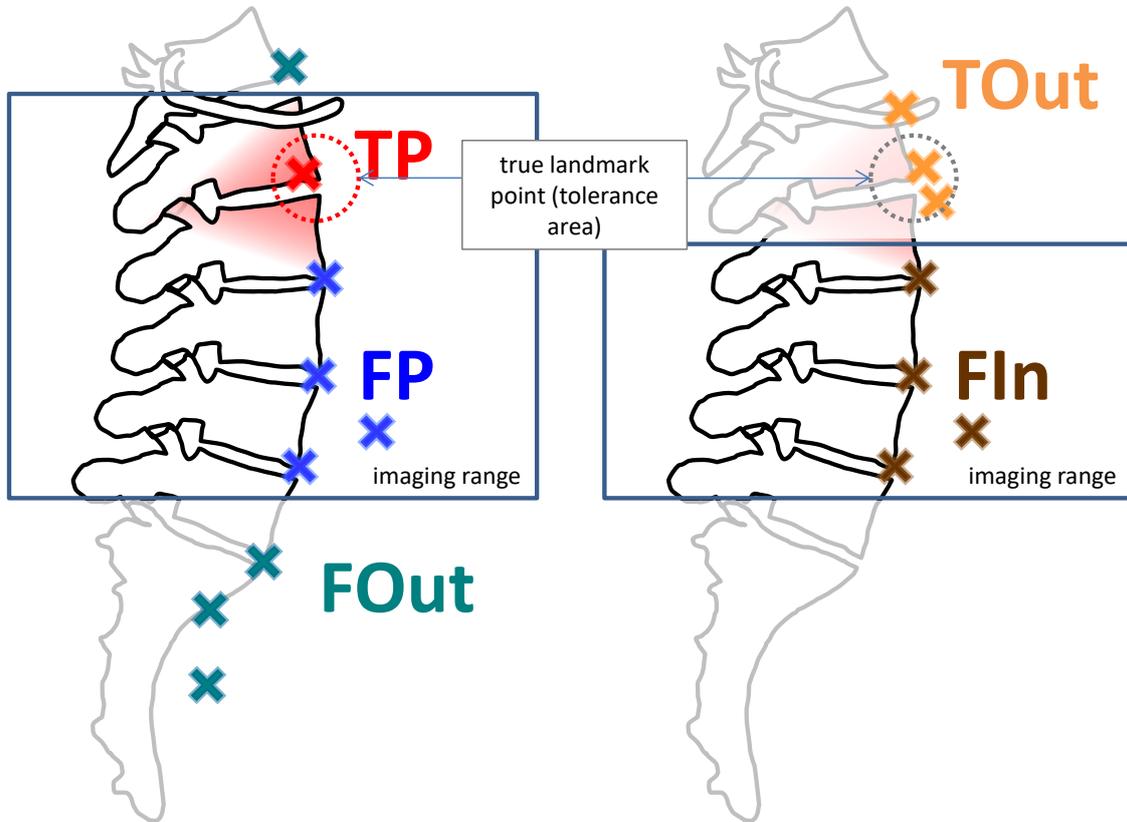
detectors). From the results,  $(\lambda_{coord} = 100, c_{detector} = 40)$  and  $(\lambda_{dist} = 0.25, c_{detector} = 100)$  were selected and used in the actual training of the coordinate-based and distance-based L-PDMs, respectively. The number of sample points in the importance sampling was set at  $J = 100$  in this study. In the simulated annealing, the initial temperature  $T_0 = 1000$ , the cooling ratio  $\gamma_T = 1000^{-1/1000}$  and the iteration number  $N_{iteration} = 1000$  were used. The tolerance of the distance error was set at  $d_{tol} = 10.0$  (mm). These parameters were determined empirically.

In the evaluation experiment with the distance-based L-PDM, my landmark detection framework attempted to detect all 197 landmarks without any prior knowledge of the imaging range, contrast agent injection status and so forth. On the other hand, in the experiment with the coordinate-based L-PDM, the ground truth was used to determine the gravity center and rotation angle (i.e., pose) of the target body in the unseen volume by rigidly registering the ground-truth landmark points to the L-PDM. Then, all the detector outputs were relocated in advance to fit those of the L-PDM. Note that this intentionally unfair initialization was used for only the coordinate-based L-PDM and was not used for the distance-based L-PDM.

The detection framework outputs one of the following statuses for each landmark position: (a) detected, (b) not detected but estimated (i.e., as an artificial candidate) in the imaging range or (c) not detected and estimated to be outside the imaging range. Then, each detection result was classified into one of the following four categories (Fig. 11):

- **True Inside (TIn).** The target landmark is within the imaging range and its position is detected or estimated. Each TIn is subcategorized into one of the following:

- **True Positive (TP).** TIn within  $d$  mm from the ground-truth position (i.e., detected or estimated correctly).
- **False Positive (FP).** TIn whose distance error is no less than  $d$  mm (i.e., detected or estimated wrongly).
- **True Outside (TOut).** The target landmark is outside the imaging range and is correctly estimated to be outside the imaging range.
- **False Inside (FIn).** The target landmark is detected but the position is not correct or the target landmark is detected but the true landmark does not exist in the imaging range.
- **False Outside (FOut).** The target landmark is estimated to be at a point outside the imaging range, despite the landmark existing in the imaging range.



**Fig. 11.** Definition of detection result criteria. (Left) TP, FP and FOut are defined when the ground-truth landmark point (L1/2 intervertebral disk posterior margin) is within the imaging range. (Right) TOut and FIn are defined when the ground-truth landmark point is outside of the imaging range.

The number of TPs (#TP) was evaluated using the criterion  $d = 10$  (mm). The TP ratio and TOut ratio were defined as

$$TP\% = \frac{\#TP}{\#TP + \#FP + \#FOut}$$

$$TOut\% = \frac{\#TOut}{\#TOut + \#FIn} ,$$

where #TP is the total number of TPs, and so forth.

**Dataset B.** To compare the vertebral identification performance of my method with other methods, I tested my method with a publicly available spinal CT dataset [9] (<http://research.microsoft.com/en-us/projects/spine>). The dataset (Dataset B) includes 242 scans of partial or whole spinal columns with manual annotation of each vertebral centroid. Here, my method is designed to detect the anterior or posterior mid-center point of intervertebral disks, not the centroid of the vertebra. To apply my method to their centroid detection problem, I assumed each centroid position to be the gravity center of four points; anterior and posterior mid-center points of the disks above and below the target vertebra. Vertebral centroid detection was judged as successful if its estimated centroid (i.e., the gravity center) was within 2 cm of the true center, and the closest centroid was the correct one (the same criteria as Glocker et al.). Since my method is optimized to detect multiple landmark points per vertebra rather than vertebral centroids, I did not retrain my system; instead, I simply trained my system with my datasets and then applied it to their dataset.

My implementation of the proposed method took approximately 12 min and 5 min for the initial landmark detection and the subsequent combinatorial optimization for one subject, respectively, using a workstation with an 8-core Intel i7-5960X CPU and an NVIDIA Tesla C2050 GPGPU.

## 2.2 Results

### 2.3.1 Detection accuracy

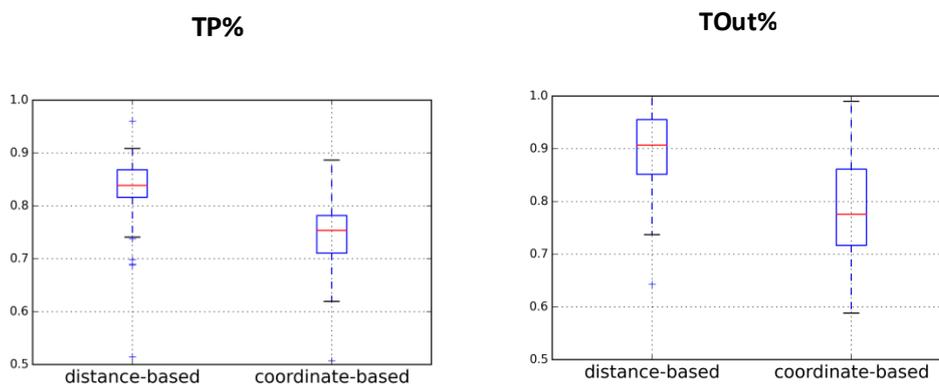
The detection performance results of Dataset A are shown in Fig. 12 and examples of landmark detection results are shown in Fig. 13. As shown, the TP% values were 83.8% and 74.6% for the distance-based and coordinate-based L-PDMs, respectively, with the criterion of  $d = 10$  mm. The averages and standard deviations of the distance errors (between the detected position and the ground truth) were  $6.6 \pm 6.4$  and  $8.1 \pm 7.9$  mm for the distance-based and coordinate-based L-PDMs, respectively. TOut% was greater for the distance-based L-PDM than for the coordinate-based L-PDM (92.1 and 82.4%, respectively). Therefore, I concluded that the distance-based L-PDM surpassed the coordinate-based one in all aspects, even though pose information on the target body was given to the latter L-PDM in advance. Therefore, the results of the distance-based L-PDM are described in detail in the rest of this section.

Figure 14 shows the detection accuracies (distance errors) of all 197 landmarks. The detection accuracies tended to be better for bony landmarks, especially the spinal and pelvic ones. The distance errors and standard deviations for the soft-tissue landmarks were relatively large.

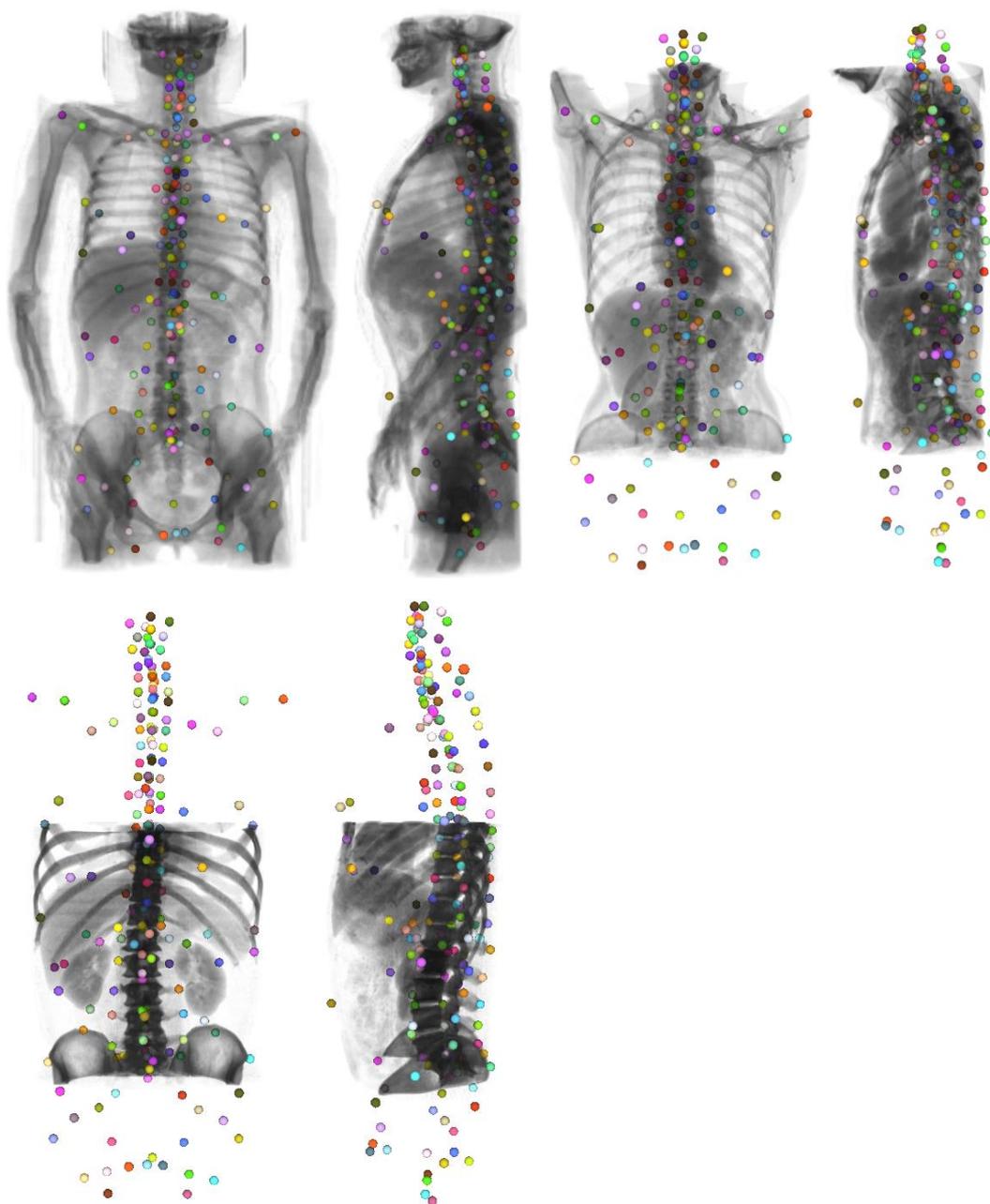
Figure 15 shows a histogram of detection distance errors for a total of 18,674 landmarks determined to be TIn (i.e., correctly determined to be within the imaging range). From Fig. 15, 83.8, 93.2 and 96.5% of landmarks were detected within 10, 15 and 20 mm from the ground truth, respectively. Detection accuracies for soft tissue landmarks and part of the bony landmarks (e.g., the tips of 11th and 12th ribs) tend to be lower than

those for spinal and pelvic landmarks, probably due to their large shape variations.

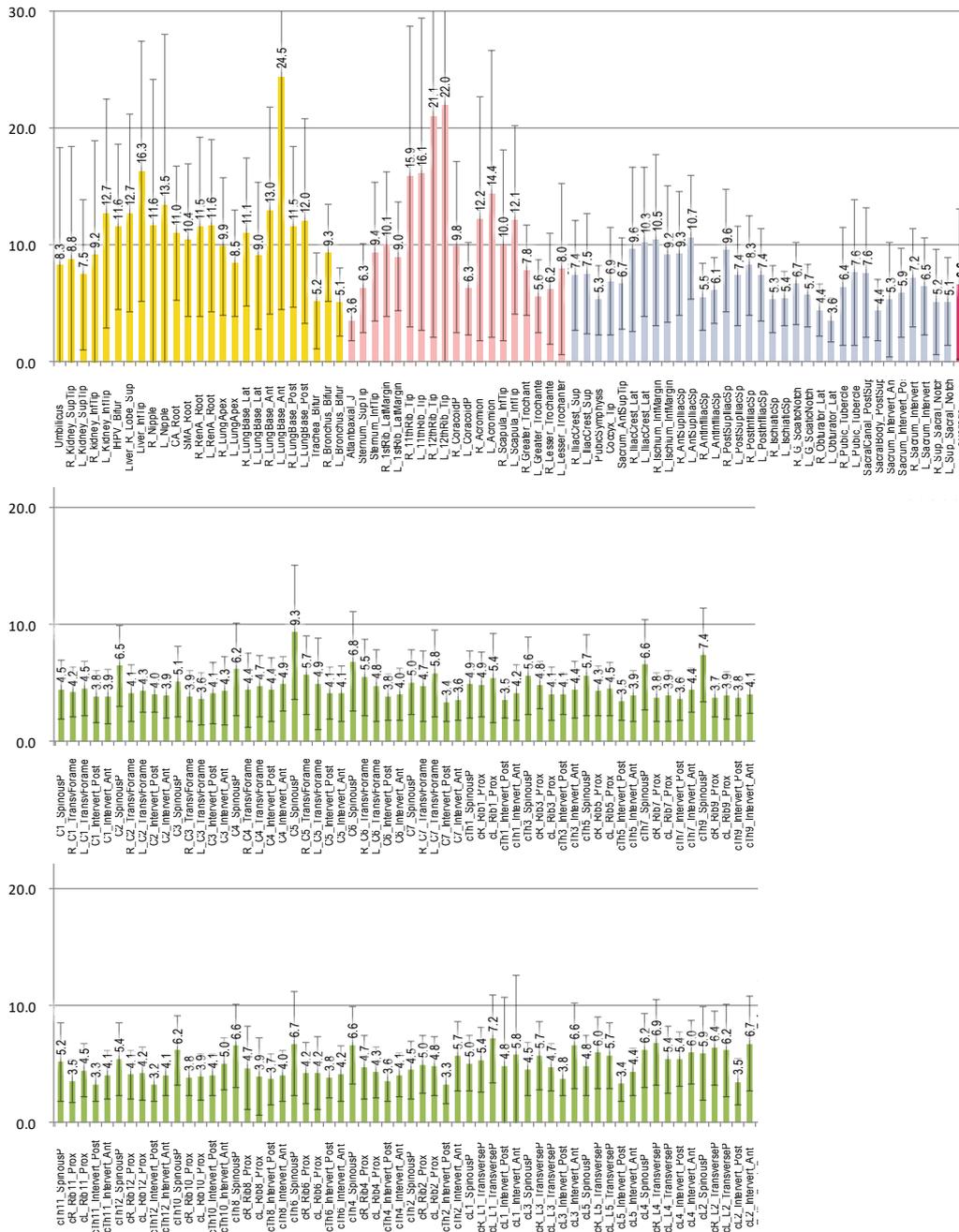
In Table 3, my landmark detection accuracies are compared with those reported by Liu and Zhou (Liu and Zhou, 2012) for landmarks suitable for comparison. Additionally, my result was also compared with the organ detection and localization method reported by Criminisi et al. [2]. As shown in Table 3, my results were less accurate than those in [4] for most of the comparable soft-tissue landmarks, especially for landmarks whose distance errors are relatively small. However, my results appear to be comparable for renal landmarks and the inferior tip of the liver when evaluated by 95 percentile values. Note that the proposed method can determine the tolerance of the distance error  $d_{tol}$ , which was set as  $d_{tol} = 10$  mm in this study. Therefore, it is natural that the distance errors of my method are approximately 10 mm for a large number of landmarks. Although direct comparison was impossible due to the the difference in criteria, the detection error distance of my method was lower than that of Criminisi et al. [2] for all comparable organs.



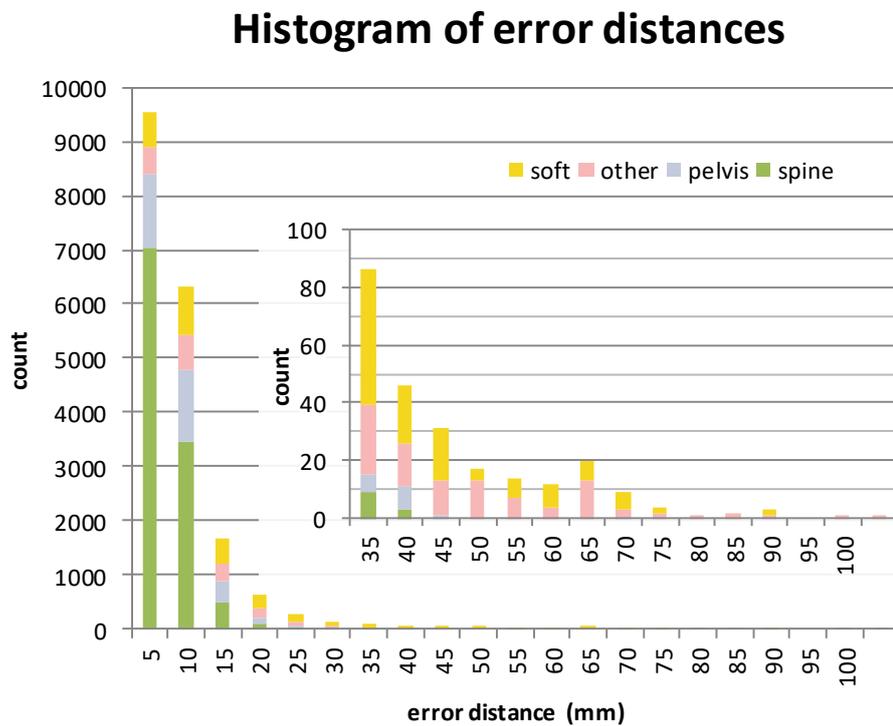
**Fig. 12.** Box plots of (left) TP% and (right) TOut% rate. The distance-based model surpassed the coordinate-based model in both two criteria. The differences between the two models were statistically significant for both criteria ( $p < 0.001$ , Student's paired t-test).



**Fig. 13.** Examples of landmark detection results. (Top left) Subject with neck-to-pelvis imaging range, anterior and lateral views. (Top right) Subject with chest-to-abdomen imaging range. (Bottom) Subject with upper abdomen imaging range. Note that the landmark positions out of the imaging range are estimated with reasonable accuracy.



**Fig. 14.** Average distance errors and standard deviations of detected landmarks using the distance-based L-PDM. Green: spine, yellow: soft tissue, blue: pelvis, pink: other bones.



**Fig. 15.** Histogram of distance errors of all detection results from all subjects.

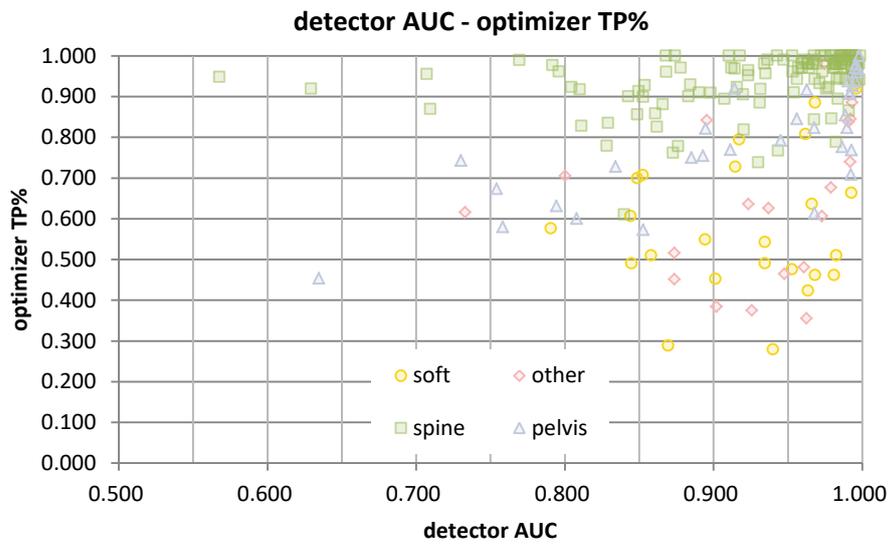
**Table 3.** Comparison of detection accuracies between my method and those in [4] and [2].

Landmark	average distance error (mm)			95 percentile (mm)	
	proposed	Criminisi et al.*	Liu et al.	proposed	Liu et al.
airway					
bifurcation of trachea	5.2		2.5	12.3	4.5
1st bifurcation of rt. bronchus	9.3			12.3	
1st bifurcation of lt. bronchus	5.1			17.6	
lung					
rt. lung		10.1*			
rt. lung apex	9.9		3.2	21.4	8.5
rt. lung base lateral margin	11.1			21.8	
lt. lung		12.9*			
lt. lung apex	8.5		2.6	15.8	6.0
lt. lung base lateral margin	9.0			22.7	
liver					
liver		15.7*			
superior margin of liver	12.7		2.5	30.3	4.0
inferior tip of liver	16.3		6.4	38.4	30.5
root of celiac artery	11.0			22.1	
kidney					
rt.kidney		16 <sup>*</sup>	6.4 <sup>†</sup>		39.2 <sup>†</sup>
superior tip	8.8			22.2	
inferior tip	9.2			24.9	
root of rt. renal artery	11.5			26.3	
lt.kidney		13.6 <sup>*</sup>	8.4 <sup>†</sup>		50.7 <sup>†</sup>
superior tip of lt. kidney	7.5			17.1	
inferior tip of lt. kidney	12.7			31.1	
root of lt. renal artery	11.6			27.0	
all 120 spinal landmarks	4.8			10.2	
all 25 soft tissue landmarks	11.1			29.3	
all 33 pelvic landmarks	7.0			16.4	
all 19 other bony landmarks	11.1			31.2	
all 197 landmarks	<b>6.6</b>			<b>17.0</b>	

\* (bounding box localization errors) † center of the organ

### 2.3.2 Correlation between performance of each single-landmark detector and overall detection accuracy

Since the proposed Gibbs-sampling-based combinatory optimization framework can select the best detector candidate or can even estimate landmark positions by itself, the correlation between the detector performance and the overall accuracy of the framework for each landmark is analyzed. Figure 16 is a scatter plot of the detector performance (area under the receiver operating characteristic curve) and the detection accuracy of the proposed framework (for all 197 landmarks). The vertical axis shows the optimizer TP% of a certain landmark and the horizontal axis shows detector AUC of the same landmark, where all landmarks are used by the optimizer. As shown, the framework can successfully detect spinal landmarks even when the AUC of the detector is no more than 0.6. It means that that my optimization system can find an optimal point for spinal landmarks even if the corresponding detector has very poor performance. Therefore, I strongly believe that my framework is suitable for detecting spinal landmarks with the highest possible reliability. On the other hand, the detection reliability for most of the soft-tissue landmarks seems to be lower than that for spinal and pelvic landmarks. The probable reason for this is the high deformability and large position variance of these soft tissue landmarks in the human body.

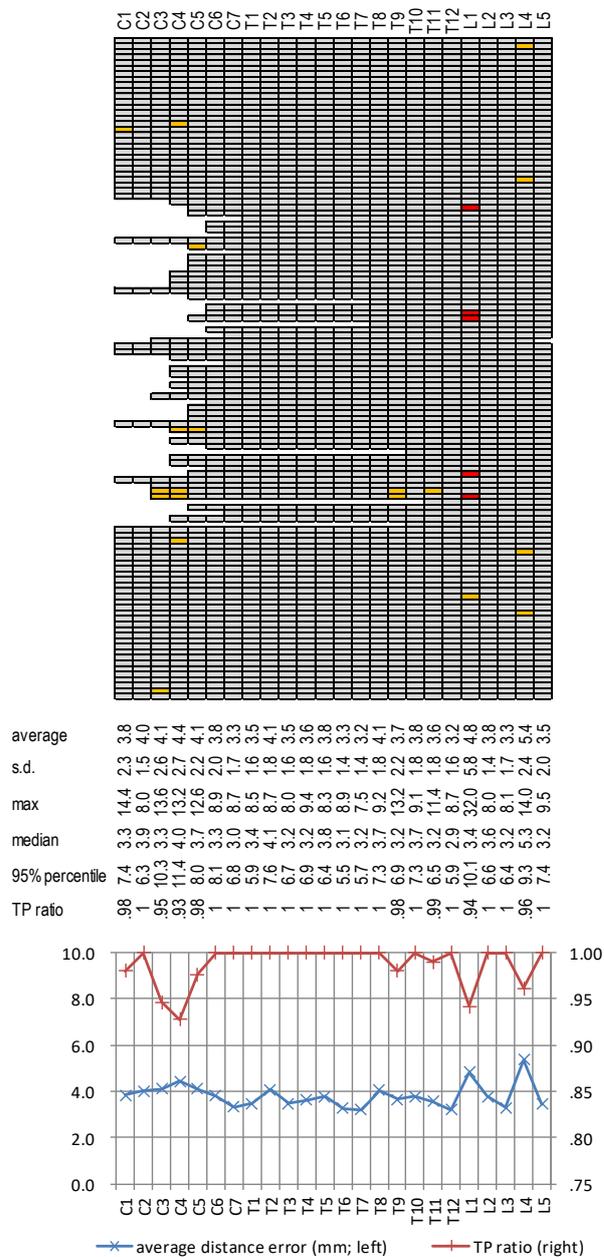


**Fig. 16.** Scatter plot showing relationship between the landmark detector performance and TP% in the final optimization result. The AUC was calculated assuming that each single-landmark detector outputs 100 FPs. Note that TP% of over 90% can be achieved even when the AUC is lower than 0.6.

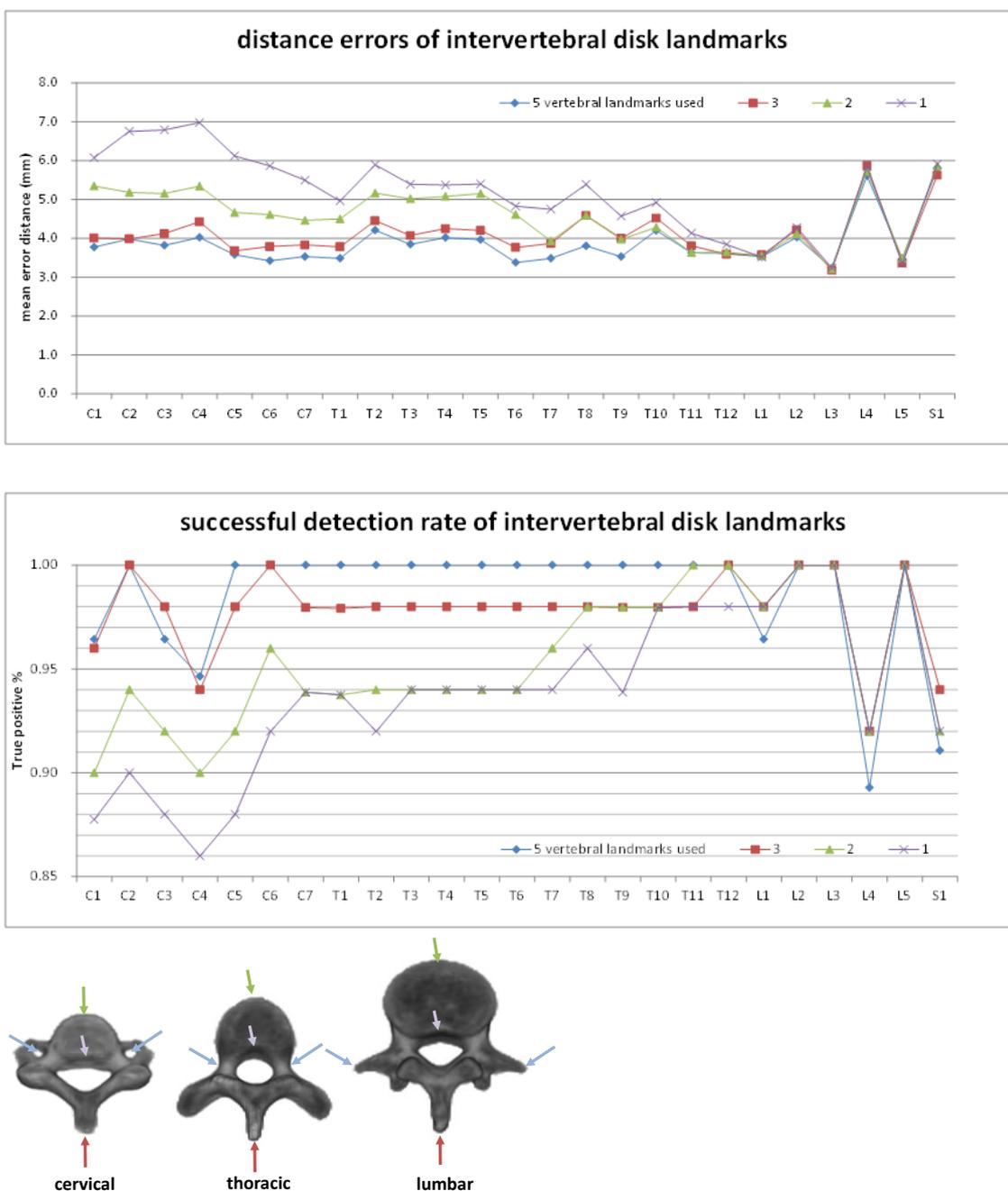
### 2.3.3 Accuracy of vertebra detection and identification using Dataset A

Since my application can detect five landmarks per vertebra, it can detect and identify each vertebra in the given CT datasets. Figure 17 summarizes the vertebra detection and identification accuracy using my dataset. As shown, a total of 2285/2317 (98.6%) of intervertebral disk (posterior mid-center) positions were correctly detected within 10 mm from the corresponding ground-truth positions. The average distance error and standard deviation were  $3.9 \pm 2.4$  mm, which is comparable to the result of Kelm et al. [10], in which the distance error was 3.22 mm. The disks can be correctly detected and identified even if the imaging range includes only 7 or 8 vertebrae. Additionally, the proposed method showed no cranial/caudal shifts, which were reported in the two related studies by Major et al. [6] and Kelm et al. [10].

I assume that the reliable suppression of cranial/caudal shifts shown in my result was achieved by multiple spinal and non-spinal landmark detection. To prove this, I intentionally removed some of the five landmarks from each vertebra and compared the disk detection / identification accuracy. Instead of all five landmarks per vertebra, only one, two or three landmarks per vertebra (together with all non-spinal landmarks) were used and detected by the proposed landmark detection framework. The result is illustrated in Fig. 18. Most of the disks showed improved detection accuracy as the number of landmarks per vertebra increased. Therefore, I concluded that my “multiple landmarks per vertebra” approach was effective for improving the spinal landmark detection accuracy.



**Fig. 17.** (Top) Result of vertebra position detection. Red: >20 mm distance error, yellow: >10 mm distance error. A white blank means the corresponding vertebra is out of the imaging range. (Bottom) Detection accuracy (average distance error and TP ratio) of vertebrae at each level.



**Fig. 18.** Detection accuracies of vertebrae when the number of vertebral landmarks was changed. Note that the detection accuracy improves as more landmarks per vertebra are detected simultaneously. (Top) Average distance errors. (Middle) TP ratio when the successful detection criterion was  $<10$  mm from the ground-truth mid-center posterior point of the intervertebral disk. (Bottom) Five landmarks used in this experiment.

The anterior/posterior margins of the intervertebral disks, spinal process, transverse foramina (only for cervical vertebrae), costovertebral angles (only for thoracic) and transverse processes (only for lumbar) were used. The colors of the arrows correspond to those in the above graphs.

#### 2.3.4 Accuracy of vertebra detection and identification using Dataset B

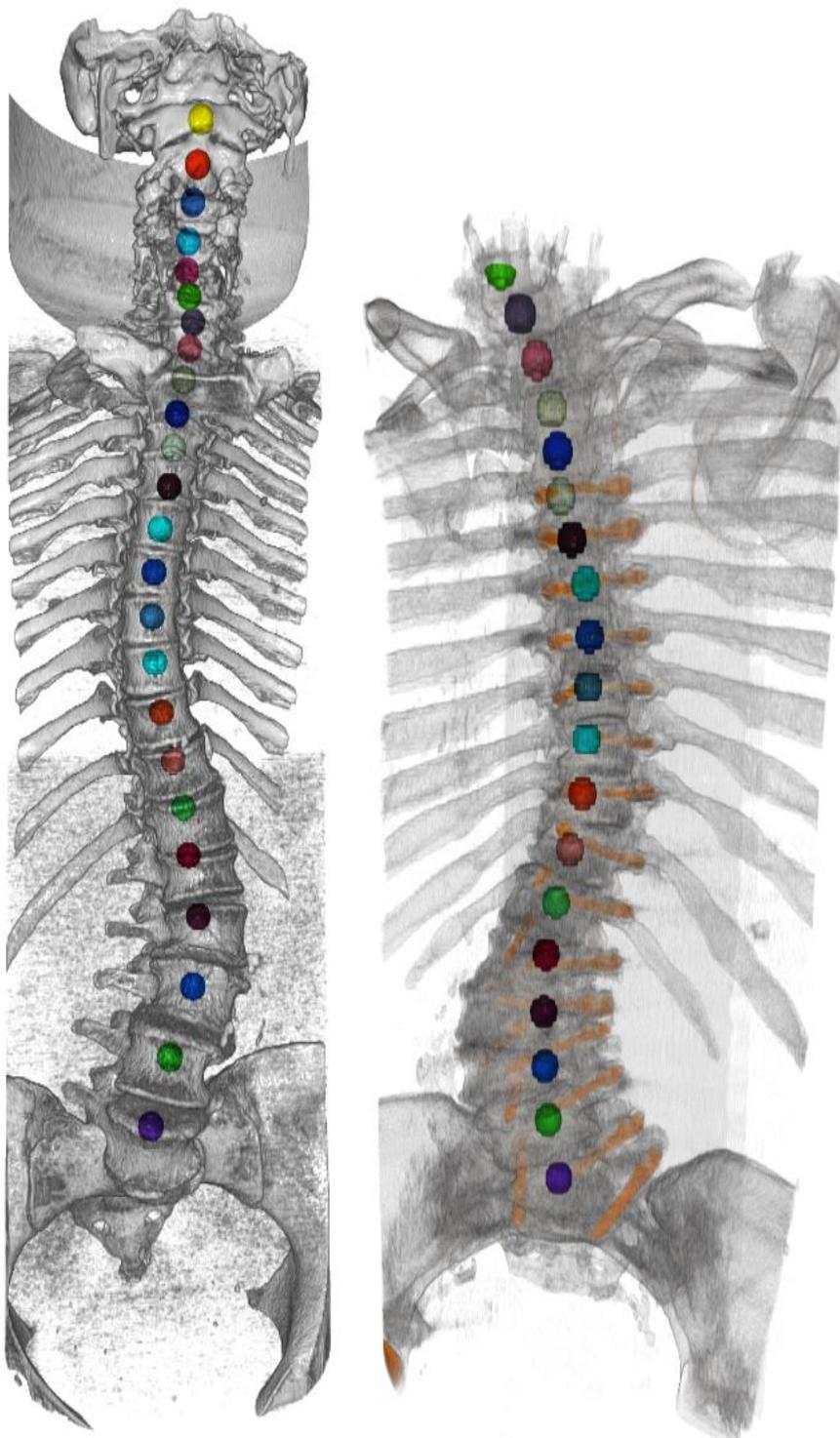
I also tested my method using the public dataset provided by [9]. It consists of 242 partial or whole spinal CT images. I applied my method to all 242 CT volumes and evaluated the results using the same criteria as [9].

The results are summarized in Table 4. As shown, the overall successful identification rate was 70%, which is the same as that of [9]. Therefore, I concluded that my method has comparable performance to their state-of-the-art method. Although the proposed method had a slightly smaller median of localization errors, it had larger standard deviations especially in the lumbar region. That is because my proposed algorithm sometimes failed to converge when the image only included a pathological lumbar spine with severe metal artifacts.

Examples of identification results are shown in Fig. 19.

**Table 4.** Localization errors (mm) and successful vertebral centroid identification rates (%) of the proposed method and the method in [9].

	Glocker et al.				Proposed			
	Median	Mean	Std.	Id. %	Median	Mean	Std.	Id. %
All	8.8	<b>12.4</b>	<b>11.2</b>	70	<b>7.5</b>	14.6	28.8	70
Cervical	5.9	7.0	4.7	80	<b>5.6</b>	<b>6.8</b>	4.7	<b>81</b>
Thoracic	9.8	13.8	<b>11.8</b>	62	<b>7.8</b>	13.8	16.7	<b>66</b>
Lumbar	<b>10.2</b>	<b>14.3</b>	<b>12.3</b>	<b>75</b>	10.4	24.4	52.0	69

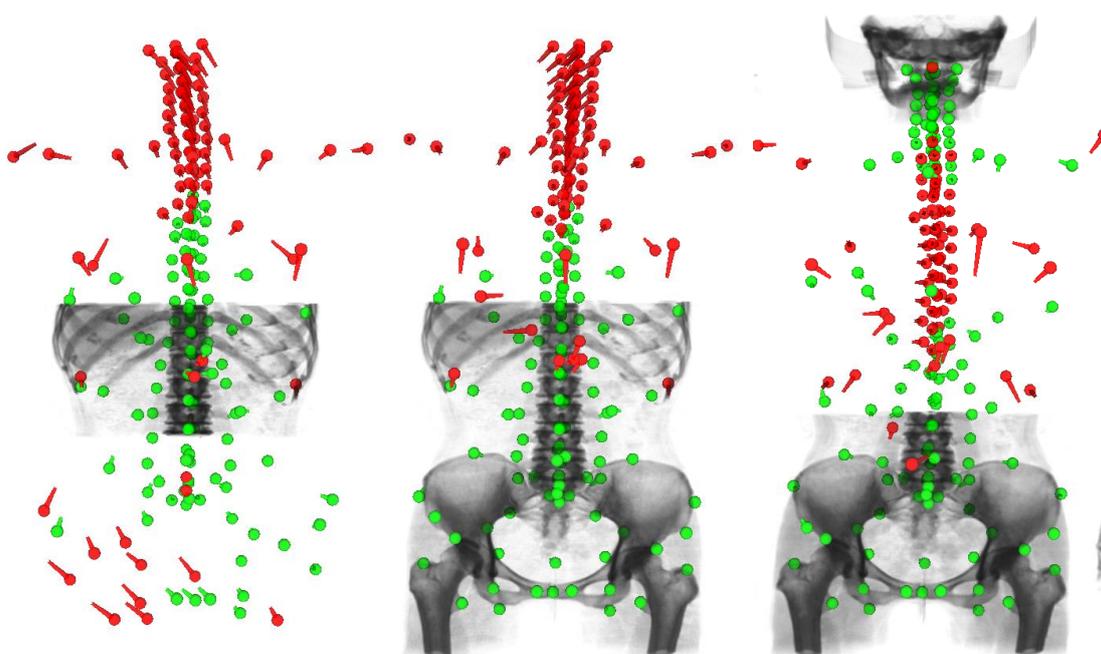
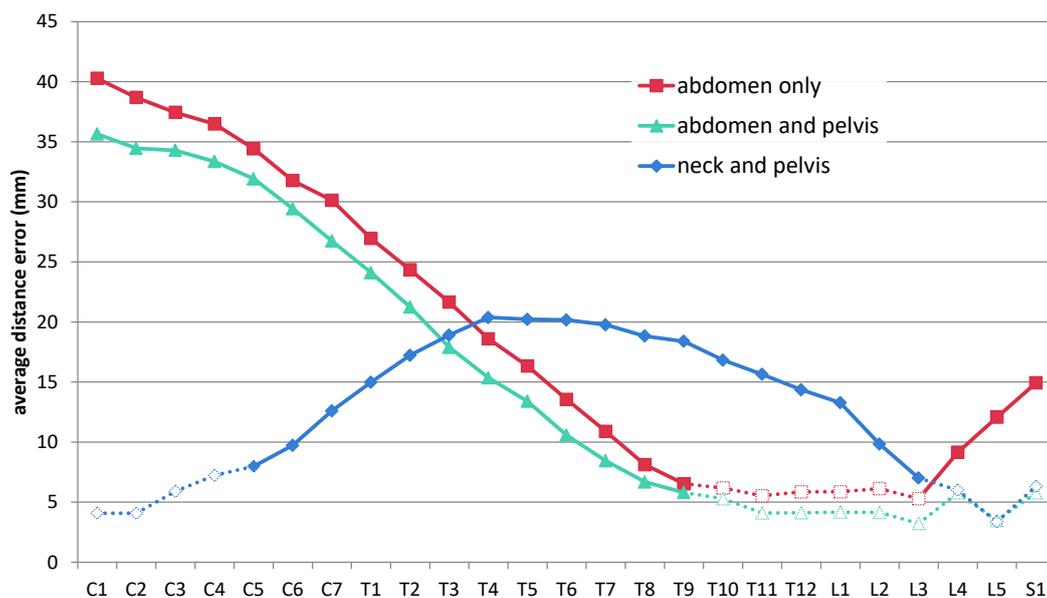


**Fig. 19.** (Left) An example of identification result for whole-spine vertebral centroids. (Right) Another example with metal screw implants.

### 2.3.5 Position estimation of landmarks outside the imaging range

As already shown in Fig. 13, my method can successfully estimate landmark positions that are not included in the given CT volume. In this section, the accuracy of such estimation is analyzed using artificial volumes in which the imaging range is intentionally truncated.

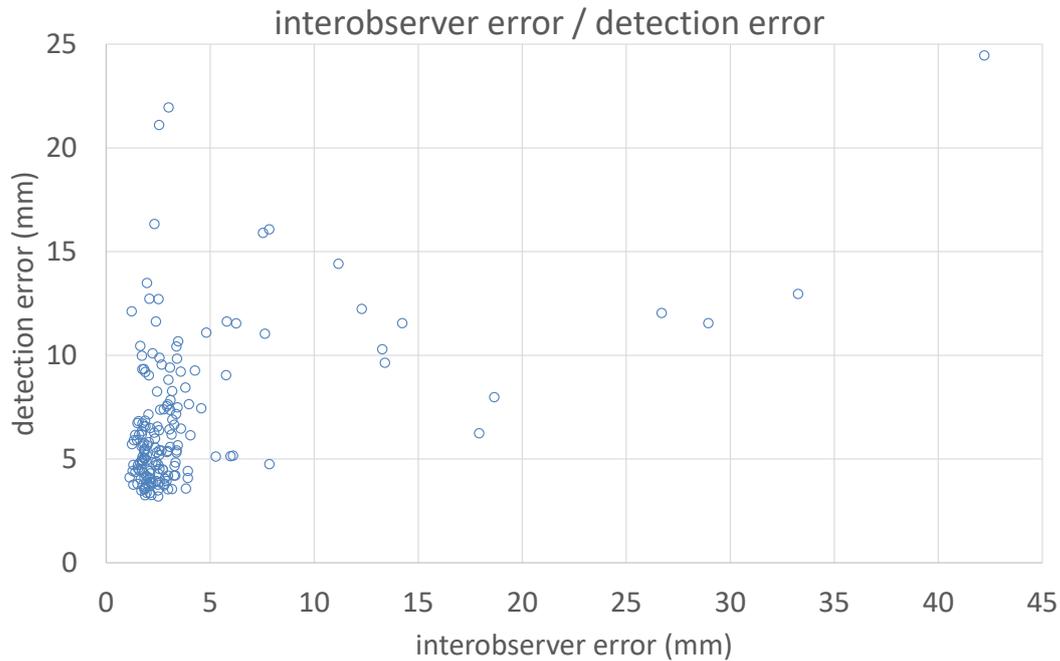
Figure 20 illustrates the intentionally truncated volumes and their landmark detection/estimation results. The truncated volumes were artificially created using 60 datasets with neck-to-pelvis imaging ranges and used as test datasets (30 cases per fold). The training datasets were the same as in the other experiments (60 cases per fold). Although the estimation accuracy of each landmark decreased as the landmark became further from the imaging range, the body structure were recovered with a reasonable shape. The estimation accuracy was lower in the cervical area, owing to the difficulty in estimating the pose information of the neck. The estimation accuracy improved when both the head-neck and pelvic regions were inside the imaging range and the other regions were estimated.



**Fig. 20.** (Top) Average distance errors of estimated landmark positions inside (dotted line) and outside (solid line) the intentionally truncated imaging range. (Bottom) Examples of estimation results with various truncated imaging ranges. Red pins indicate a distance error of over 10 mm (the head and tail of each pin indicate the estimated/detected and ground-truth landmark points, respectively).

### 2.3.6 Landmark position ambiguity analysis using interobserver errors among human experts

Most of anatomical landmarks cannot be strictly determined at one point. Otherwise, each landmark has particular amount of positional ambiguity due to the size and shape of the structure on which the landmark is defined. To estimate these ambiguities, interobserver errors among human experts were evaluated. Four human experts inputted landmark positions for three CT datasets. Then, for each landmark, the interobserver error distance (i.e., root mean square error (RMSE) of coordinates) among inputted points was calculated. The average and standard deviation of the interobserver errors of all 197 landmarks were  $3.83 \pm 5.29$  mm. A scatter plot between the interobserver errors and detection errors is shown in Fig. 21. Although most of landmarks had less than 5 mm of interobserver errors, 11 landmarks had interobserver errors larger than 10 mm. In 10 of these 11 landmarks (including the anterior and posterior bottom tips of the lungs and the lateral tips of the iliac crests), the interobserver errors were larger than the detection errors. It may suggest that these points were not suitable as manually-defined landmarks. On the other hand, the detection errors were larger than interobserver errors for most of landmarks. Here, I defined an ambiguity-subtracted error distance as (detection error distance) – (interobserver error distance). Then, for 40 out of 197 landmarks (20.0%), the ambiguity-subtracted error distances were larger than 5mm. In other words, 80% of landmarks were successfully determined within 5mm of error distances when the positional ambiguities were subtracted.



**Fig. 21.** Scatter plot between the detection error distances of the proposed system and interobserver error distances of human experts.

## 2.4 Discussion

In this study, a framework for multiple-landmark detection in CT volumes with various imaging ranges was presented. It gave true positive detection ratios of 84.3 and 96.5% when the tolerances of the detection distance error were 10 and 20 mm, respectively. Furthermore, 92.1% of the landmarks outside the imaging range were classified correctly into TOut (i.e., estimated as being out of the imaging range). Therefore, I believe that the feasibility of my approach in detecting over 100 landmarks simultaneously in CT

datasets with various imaging ranges has been validated. To the best of my knowledge, this is the first study in which over 100 anatomical landmarks have been automatically detected and their accuracies evaluated using CT datasets with various imaging ranges.

In the evaluation, the detection performance of the distance-based L-PDM is superior to that of the coordinate-based one, despite the unfair evaluation settings in which pose information was given to the coordinate-based L-PDM in advance. Several possible reasons for this can be considered. Firstly, my L-PDMs are defined as multiple Gaussian distributions of variables. In the coordinate-based L-PDM, the probability is defined as the exponential of a sum of terms (as in Eq. (14)), and each term is dependent on only one or two landmark positions. Thus, this L-PDM can be regarded as a Markov random field (MRF) with a complete graph, where each node corresponds to a landmark. On the other hand, in the distance-based L-PDM, the probability is also defined as the exponential of a sum of terms. However, each term is not determined from two landmarks but two inter-landmark distances, which are dependent on at most four landmarks. Thus, the distance-based L-PDM can be regarded as a fourth-order MRF, i.e., a higher-order MRF (HOMRF). It is possible that this complexity helps to model the landmark distribution better. Secondly, the logarithmic distance has some advantages over coordinates. For instance, it has translation and rotation invariance. Note that an inter-landmark distance not only has global invariance but is also robust to local rotations or translations when the rotating or translating structure includes both landmarks. Additionally, a scaling transformation adds a common constant to all the logarithmic distances; for example, a twofold scale transformation adds  $\ln 2$  to all

logarithmic distances. I surmise that these properties are beneficial in constructing the model.

In the comparison of landmark detection accuracies, the distance errors of my algorithm tended to be greater than those reported by Liu and Zhou [4] (Table 3). Because they used relatively large training datasets of 500 cases, it is possible that my training datasets (52 for each fold) were too small to represent the variety of landmark appearances. It will be my future work to survey the impact of the training dataset size on the detection accuracy using a larger dataset. Note that a precise comparison is not possible because the number of landmarks in [4] was 60, and detection accuracies were reported for only 13 of them. I am currently attempting to improve the detection accuracy by postprocessing, which can fine-tune the detected landmark positions [18].

One of the applications expected for my landmark detection framework is the identification of each vertebra (such as the 1st cervical, 6th thoracic or 5th lumbar vertebra) in CT datasets. This is the first study in which multiple landmarks on all 24 vertebrae have been simultaneously detected in clinical CT volumes. In my results, the average distance error and successful identification rate were 3.9 mm and 98.6%, respectively. Recently, Glocker et al. [8] reported a study in which all 24 vertebral centroids were detected automatically from CT datasets with mostly healthy spinal columns. Their target landmarks were the 26 vertebrae from C1 to S2, and their overall localization error and successful identification rate were 6.10 mm and 81%, respectively. Although different datasets were used in their study and my study, my results surpass both their localization and identification results. Furthermore, I performed another comparison study using the same dataset as that in [9] which includes spines with severe

pathological changes. My successful identification rate was 70%, which is the same as that of [9]. Therefore, I believe that my method has a comparable performance to other state-of-the-art methods.

Among the 197 landmarks detected, 120 were spinal ones. One of the purposes of this choice of landmarks in my study was to validate the ability of my combinatorial optimization framework to identify and distinguish a series of very similar landmarks. It is not an easy task to determine the index of each vertebra in CT images. For example, Klinder et al. [5] reported an automatic segmentation method for all 24 vertebrae, and their segmentation results were excellent with an average distance error of 1.12 mm. However, prior to segmentation, their vertebral identification process failed in three out of 59 cases owing to a unit shift in the vertebral index, and these three cases were excluded from their evaluation of segmentation accuracy. Note that the incorrect determination of vertebral levels in clinical interventions such as surgery can have a hazardous effect [19]. This means that a vertebral identification failure in computer-assisted surgery may be harmful to the patient, which is why I need a reliable identification method for vertebral indices. In this study, no cranial or caudal shifts in the vertebral number were observed in 104 subjects. Thus, I conclude that the ability of my framework to identify vertebrae was validated. I am now planning to develop segmentation and registration methods for spinal bones utilizing the method presented in this study.

The identification of each vertebra becomes more difficult when only part of the spinal column is included in the given volume. Nevertheless, in my results the detection accuracy did not markedly decrease as the imaging range became narrower (Fig. 19). In particular, no deterioration in detection accuracy occurred even when the neck and pelvic regions were

outside the imaging range. This suggests that my framework can estimate the indices of vertebrae from nonspinal landmarks. Recently Connor et al. [20] reported that the determination of vertebral indices from a single nonspinal landmark is not reliable. In their study, the only reliable method was to count the vertebrae from the 2nd cervical vertebra. However, in this study, the indices of vertebrae were successfully estimated even in datasets that did not include the neck region, and thus did not include the 2nd cervical vertebra. Therefore, I consider that multiple landmark information is necessary to reliably estimate the vertebral indices when the whole spine is not included in the volume. My simultaneous landmark detection approach can respond flexibly to such a situation.

This work has some limitations. Firstly, subjects with anomalous numbers of vertebrae were excluded from the evaluation. Such anatomical anomalies are quite common. For example, an occurrence of 10.8% was reported in [21]. I am attempting to solve this problem by virtually converting the landmark configurations of anomalous spines to those of normal ones [22]. However, it is almost impossible to determine the existence or nonexistence of such anomalies when the imaging range is limited and does not include the whole spine. It is probable that only an application-dependent answer can be defined in such a case.

Another problem is the long calculation time. My current implementation takes approximately 17 min per subject. Although this may not be problematic when it is used for preprocessing before other very time-consuming tasks such as precise nonrigid registration, it may not be affordable for many medical image analysis applications. One possible way to reduce the execution time is to apply one of the more efficient HOMRF optimization algorithms, such as that proposed in [23], to my

optimization framework. However, it will be challenging to apply an HOMRF optimization algorithm to my framework, which has to handle missing landmarks caused by detection failure and a limited imaging range.

## 2.5 Conclusion

A framework for simultaneously detecting 197 anatomical landmarks was presented. In the framework, a novel two-stage sampling algorithm was introduced to appropriately handle missing landmarks due to limited detector sensitivity, a partial imaging range and surgical/anatomical defects of landmarks. Each landmark detector was parametrically modeled, and the landmark configuration in the human body was also statistically modeled as an L-PDM. The two-stage sampling algorithm optimally utilizes these stochastic models and all detector-derived information to detect or estimate all the landmark positions, even if a significant number of landmarks are missed by the detector or outside of the imaging range. Its feasibility was validated through experiments with various CT datasets and an overall mean detection error of 6.6 mm was achieved. The feasibility of the estimation algorithm for out-of-range landmark positions was validated via experiments with intentionally cropped volumes. The detection accuracy of all 24 vertebrae via my “multiple landmarks per vertebra” approach was also confirmed since the identification and detection accuracies were comparable to those of another state-of-the-art method.

My future works may include improving the detection accuracy, the handling of vertebral number anomalies and the implementation of a more efficient solver for HOMRF optimization.

## Appendix 1. Calculation of the multiple candidate model

Here, I make the assumption that the candidate set  $\mathbf{c}_i, i \in \{1, 2, \dots, |S|\}$  does not include more than one TP candidate. Therefore, the total number of possible probability events is  $(|S| + 1)$  : one of the  $\mathbf{c}_i, i = 1, 2, 3, \dots, |S|$  is correct or none of the  $\mathbf{c}_i$  are correct. Let  $p_0 \equiv p(\mathbf{c}_i \notin R_{true}, \forall i)$  be the prior probability that none of the  $\mathbf{c}_i$  are TP. Here,  $p_0$  is a parameter to be determined in advance. In this study  $p_0 = 0.02$  is used. Then, the prior probability that the  $i$ th candidate is TP can be estimated as

$$p(\mathbf{c}_i \in R_{true}) = \frac{1 - p_0}{|S|} \quad (1 \leq i \leq |S|), \quad (27)$$

so that the sum of the probabilities  $p_0 + \sum_{i=1}^{|S|} p(\mathbf{c}_i \in R_{true})$  is 1. Note that the prior probabilities of all candidates are regarded as being uniform because these are prior probabilities *without any prior knowledge* of the likelihood scores (a so-called uninformative prior).

Let the simultaneous distribution of all likelihood scores  $\mathbf{u} = (u_1 \ u_2 \ \dots \ u_{|S|})$  be  $p(\mathbf{u})$ . Because of the assumption that no more than one candidate can be correct, the elements of  $\mathbf{u}$  are not independent of each other. For example, it is expected that a certain candidate will have a large score if all the other candidates are known to be FPs. Therefore, only one element of  $\mathbf{u}$  tends to have a much larger value than the others. However, this tendency vanishes when I know which candidate is correct. Because all other candidates are known to be incorrect, their likelihood scores become much less dependent on each other. On the basis of this reasoning, I assume that  $u_i, i \in \{1, 2, \dots, N_l\}$  are independent of each other under the condition

that the correct candidate  $\mathbf{c}_k$  is fixed. Thus, the conditional simultaneous distribution of  $p(\mathbf{u})$  with the condition of  $\mathbf{c}_k \in R_{true}$  can be represented as follows:

$$\begin{aligned}
p(\mathbf{u}|\mathbf{c}_k \in R_{true}) &= p(u_k|\mathbf{c}_k \in R_{true}) \\
&\cdot \prod_{i=1, i \neq k}^{|\mathcal{S}|} p(u_i|\mathbf{c}_i \notin R_{true}) \\
&= \frac{p(u_k|\mathbf{c}_k \in R_{true})}{p(u_k|\mathbf{c}_k \notin R_{true})} \\
&\cdot \prod_{i=1}^{|\mathcal{S}|} p(u_i|\mathbf{c}_i \notin R_{true}).
\end{aligned} \tag{28}$$

On the other hand, the detector candidates can also be regarded as independent of each other when all the candidates are known to be FPs (i.e.,  $\mathbf{c}_i \notin R_{true}, \forall i$ ). That is,

$$p(\mathbf{u}|\mathbf{c}_i \notin R_{true}, \forall i) = \prod_{i=1}^{|\mathcal{S}|} p(u_i|\mathbf{c}_i \notin R_{true}). \tag{29}$$

Now we are able to calculate the posterior probability of each candidate, i.e.,  $p(\mathbf{c}_k \in R_{true}|\mathbf{u})$ , as follows:

$$\begin{aligned}
p(\mathbf{c}_k \in R_{true} | \mathbf{u}) &= \frac{p(\mathbf{u} | \mathbf{c}_k \in R_{true}) \cdot p(\mathbf{c}_k \in R_{true})}{p(\mathbf{u})} \\
&= \frac{1}{p(\mathbf{u})} \cdot \frac{1 - p_0}{|S|} \cdot \frac{p(u_k | \mathbf{c}_k \in R_{true})}{p(u_k | \mathbf{c}_k \notin R_{true})} \\
&\quad \cdot \prod_{i=1}^{|S|} p(u_i | \mathbf{c}_i \notin R_{true}) \\
&= \frac{1 - p_0}{|S|} \cdot \left( \frac{p(\mathbf{c}_k \in R_{true})}{p(\mathbf{c}_k \notin R_{true})} \right)^{-1} \\
&\quad \cdot \frac{p_{true}(u_k)}{1 - p_{true}(u_k)} \cdot \frac{1}{p(\mathbf{u})} \prod_{i=1}^{|S|} p(u_i | \mathbf{c}_i \notin R_{true}).
\end{aligned} \tag{30}$$

In the same way, the probability of all candidates being FP becomes

$$\begin{aligned}
p(\mathbf{c}_i \notin R_{true}, \forall i | \mathbf{u}) &= \frac{p(\mathbf{u} | \mathbf{c}_i \notin R_{true}, \forall i) \cdot p(\mathbf{c}_i \notin R_{true}, \forall i)}{p(\mathbf{u})} \\
&= p_0 \cdot \frac{1}{p(\mathbf{u})} \cdot \prod_{i=1}^{|S|} p(u_i | \mathbf{c}_i \notin R_{true}).
\end{aligned} \tag{31}$$

Because of the assumption that no more than one candidate can be correct,  $\sum_{k=1}^{|S|} p(\mathbf{c}_k \in R_{true} | \mathbf{u}) + p(\mathbf{c}_i \notin R_{true}, \forall i | \mathbf{u}) = 1$  must be satisfied to make the sum of the probabilities of all events equal to 1. Thus,

$$\begin{aligned}
p(\mathbf{c}_k \in R_{true} | \mathbf{u}) &= \frac{1}{C} \cdot \frac{1 - p_0}{|S|} \cdot \left( \frac{p(\mathbf{c}_k \in R_{true})}{p(\mathbf{c}_k \notin R_{true})} \right)^{-1} \\
&\quad \cdot \frac{p_{true}(u_k)}{1 - p_{true}(u_k)}
\end{aligned} \tag{32}$$

$$p(\mathbf{c}_i \notin R_{true}, \forall i | \mathbf{u}) = \frac{1}{C} \cdot p_0 \tag{33}$$

are satisfied. Here,  $C = p_0 + \frac{1-p_0}{|S|} \cdot \left( \frac{p(\mathbf{c}_k \in R_{true})}{p(\mathbf{c}_k \notin R_{true})} \right)^{-1} \cdot \sum_{k=1}^{|S|} \frac{p_{true}(u_k)}{1-p_{true}(u_k)}$  is a regularization term. Note that, although Eqs. (30) and (31) are not calculable because we have not modeled the probability density function (PDF) of  $p(\mathbf{u})$ , Eqs. (32) and (33) are calculable using only the odds ratios  $\sum_{k=1}^{|S|} \frac{p_{true}(u_k)}{1-p_{true}(u_k)}$ . The constant  $r_{prior} = \frac{p(\mathbf{c}_k \in R_{true})}{p(\mathbf{c}_k \notin R_{true})}$  is the ratio of prior probabilities, which must be determined in advance. In this study  $r_{prior} = 0.01$  is used. Note that this problem formulation is a sort of semiparametric approach in which  $p(\mathbf{u})$  is ‘erased’ in the calculation and only the odds ratios are left. Thanks to this approach, we do not have to parametrize the detector output histogram  $p(\mathbf{u})$  explicitly.

## Appendix 2. Importance sampling of the artificial candidates $\{\mathbf{s}_l^1, \dots, \mathbf{s}_l^J\}$ .

In this study, the artificial candidates  $\{\mathbf{s}_l^1, \dots, \mathbf{s}_l^J\}$  are sampled from a 3-D Gaussian distribution  $q_l(\cdot)$  whose mean vector and covariance matrix are determined via importance sampling.

The principle of importance sampling is as follows. Consider a three-dimensional random variable vector  $\mathbf{z}$  and its probability distribution  $p(\mathbf{z})$ . Suppose that  $\mathbf{d}_j$  ( $j = 1, 2, \dots, J$ ) are sampled from a different distribution  $q(\mathbf{z})$ . Then, the expectation value of an arbitrary function  $f(\mathbf{z})$  can be approximated by importance sampling as

$$\begin{aligned}
\mathbb{E}(f) &= \int_{\mathbf{z} \in \mathbb{R}^3} f(\mathbf{z}) \cdot p(\mathbf{z}) d\mathbf{z} \\
&\cong \frac{\sum_{j=1}^J f(\mathbf{d}_j) \cdot \frac{p(\mathbf{d}_j)}{q(\mathbf{d}_j)}}{\sum_{j=1}^J \frac{p(\mathbf{d}_j)}{q(\mathbf{d}_j)}}.
\end{aligned} \tag{34}$$

Recall that  $\mathbf{X}^{(n)} \setminus \mathbf{x}_l = (\mathbf{x}_1^{(n)}, \dots, \mathbf{x}_{l-1}^{(n)}, \mathbf{x}_{l+1}^{(n-1)}, \mathbf{x}_L^{(n-1)})$  comprises the recently sampled positions of landmarks other than the  $l$ th landmark (whose position is  $\mathbf{x}_l$ ). Thus, to sample the artificial candidates  $\{\mathbf{s}_l^1, \dots, \mathbf{s}_l^J\}$  from  $p_{L-PDM}(\mathbf{x}_l; \mathbf{X}^{(n)} \setminus \mathbf{x}_l)$ , it is necessary to (1) first estimate the spatial probability distribution of  $\mathbf{x}_l$  using the other recently sampled landmark positions  $\mathbf{X}^{(n)} \setminus \mathbf{x}_l$ , and then (2) sample  $\{\mathbf{s}_l^1, \dots, \mathbf{s}_l^J\}$  from the estimated distribution. In other words, it is necessary to sample  $\{\mathbf{s}_l^1, \dots, \mathbf{s}_l^J\}$  from the L-PDM  $p(\mathbf{X})$  while fixing the landmark positions other than  $\mathbf{x}_l$  to the recent sample set  $\mathbf{X}^{(n)} \setminus \mathbf{x}_l$ . If the modification by the temperature is taken into account,  $p_{L-PDM}(\mathbf{x}_l; \mathbf{X}^{(n)} \setminus \mathbf{x}_l)^{1/T}$  is the target distribution to be sampled. However, this distribution is complicated and sampling directly from it is a difficult process. Instead,  $\{\mathbf{s}_l^1, \dots, \mathbf{s}_l^J\}$  are sampled from another distribution,  $q_l(\mathbf{x}_l)$ . Here,  $q_l(\mathbf{x}_l)$  is not the same as, but must be similar to, the target distribution  $p_{L-PDM}(\mathbf{x}_l; \mathbf{X}^{(n)} \setminus \mathbf{x}_l)^{1/T}$ .

The question is how to determine the sampling distribution  $q_l(\mathbf{x}_l)$ . Because the target distribution  $p_{L-PDM}(\mathbf{x}_l; \mathbf{X}^{(n)} \setminus \mathbf{x}_l)^{1/T}$  changes at every  $n$ th Gibbs sampling cycle,  $q_l(\mathbf{x}_l)$  should also be updated at each cycle. In this study, a three-dimensional Gaussian distribution is used as  $q_l(\mathbf{x}_l)$  to

approximate the target distribution  $p_{L-PDM}(\mathbf{x}_l; \mathbf{X}^{(n)} \setminus \mathbf{x}_l)^{1/T}$ . Both the mean vector and the covariance matrix of  $q_l(\mathbf{x}_l)$  are updated at each cycle. This update is performed by estimating the mean vector and covariance matrix of  $p_{L-PDM}(\mathbf{x}_l; \mathbf{X}^{(n)} \setminus \mathbf{x}_l)^{1/T}$  by importance sampling.

Using Eq. (34) and substituting  $\mathbf{z} \leftarrow \mathbf{x}_l$ ,  $\mathbf{d}_j \leftarrow \mathbf{s}_l^j$ ,  $p(\cdot) \leftarrow p(\cdot, \mathbf{X}^{(n)} \setminus \mathbf{x}_l)^{1/T}$  and  $q(\cdot) \leftarrow q_l(\cdot)$ , the mean vector  $E(\mathbf{x}_l)$  and covariance matrix  $\text{Cov}(\mathbf{x}_l)$  of the distribution  $p_{L-PDM}(\mathbf{x}_l; \mathbf{X}^{(n)} \setminus \mathbf{x}_l)^{1/T}$  can be estimated as follows:

$$\begin{aligned}
E(\mathbf{x}_l) &= \int_{\mathbf{x}_l \in \mathbb{R}^3} \mathbf{x}_l \cdot p_{L-PDM}(\mathbf{x}_l; \mathbf{X}^{(n)} \setminus \mathbf{x}_l)^{1/T} d\mathbf{x}_l \\
&\cong \frac{\sum_{j=1}^J \mathbf{s}_l^j \cdot \frac{p_{L-PDM}(\mathbf{s}_l^j; \mathbf{X}^{(n)} \setminus \mathbf{x}_l)^{1/T}}{q_l(\mathbf{s}_l^j)}}{\sum_{j=1}^J \frac{p_{L-PDM}(\mathbf{s}_l^j; \mathbf{X}^{(n)} \setminus \mathbf{x}_l)^{1/T}}{q_l(\mathbf{s}_l^j)}} \\
\text{Cov}(\mathbf{x}_l) &= E\left( (\mathbf{x}_l - E(\mathbf{x}_l))(\mathbf{x}_l - E(\mathbf{x}_l))^t \right) \\
&= \int_{\mathbf{x}_l \in \mathbb{R}^3} (\mathbf{x}_l - E(\mathbf{x}_l))(\mathbf{x}_l - E(\mathbf{x}_l))^t \cdot p_{L-PDM}(\mathbf{x}_l; \mathbf{X}^{(n)} \setminus \mathbf{x}_l)^{1/T} d\mathbf{x}_l \\
&\cong \frac{\sum_{j=1}^J (\mathbf{s}_l^j - E(\mathbf{x}_l))(\mathbf{s}_l^j - E(\mathbf{x}_l))^t \cdot \frac{p_{L-PDM}(\mathbf{s}_l^j; \mathbf{X}^{(n)} \setminus \mathbf{x}_l)^{1/T}}{q_l(\mathbf{s}_l^j)}}{\sum_{j=1}^J \frac{p_{L-PDM}(\mathbf{s}_l^j; \mathbf{X}^{(n)} \setminus \mathbf{x}_l)^{1/T}}{q_l(\mathbf{s}_l^j)}}.
\end{aligned} \tag{35}$$

These estimation results are stored and used in the subsequent  $(n + 1)$ th iteration to determine  $q_l(\mathbf{x}_l)$  as  $q_l(\mathbf{x}_l) \sim \mathcal{N}(E(\mathbf{x}_l), \alpha \cdot \text{Cov}(\mathbf{x}_l))$  (Fig. 11).

Here,  $\alpha$  is a parameter used to enlarge the distribution to avoid too early shrinkage. In this study  $\alpha = 4$  is used. The initial state of  $q_l(\mathbf{x}_i)$  is given as  $\mathcal{N}(\mathbf{g}, \beta\mathbf{I})$ , where  $\mathbf{g}$  is the geometrical center of a given CT volume and  $\beta$  is a very large constant set to cover the entire human body.  $\beta = 1000$  mm is used in this study. Note that  $\mathbf{g}$  and  $\beta$  are only used in the first iteration of the Gibbs sampling and have little effect on the optimization result.

## References

1. Potesil V, Kadir T, Platsch G, Brady SM (2011) Personalization of pictorial structures for anatomical landmark localization. *Information processing in medical imaging : proceedings of the conference* 22:333-345
2. Criminisi A, Robertson D, Konukoglu E, Shotton J, Pathak S, White S, Siddiqui K (2013) Regression forests for efficient anatomy detection and localization in computed tomography scans. *Medical Image Analysis* 17 (8):1293-1303
3. Liu D, Zhou KS, Bernhardt D, Comaniciu D Search strategies for multiple landmark detection by submodular maximization. In: *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*, 2010. IEEE, pp 2831-2838
4. Liu D, Zhou SK (2012) Anatomical landmark detection using nearest neighbor matching and submodular optimization. *Medical image computing and computer-assisted intervention : MICCAI International Conference on*

Medical Image Computing and Computer-Assisted Intervention 15 (Pt 3):393-401

5. Klinder T, Ostermann J, Ehm M, Franz A, Kneser R, Lorenz C (2009) Automated model-based vertebra detection, identification, and segmentation in CT images. *Medical image analysis* 13 (3):471-482

6. Major D, Hladůvka J, Schulze F, Bühler K (2013) Automated landmarking and labeling of fully and partially scanned spinal columns in CT images. *Medical Image Analysis* 17 (8):1151-1163

7. Donner R, Menze BH, Bischof H, Langs G (2013) Global localization of 3D anatomical structures by pre-filtered Hough Forests and discrete optimization. *Medical Image Analysis* 17 (8):1304-1314

8. Glocker B, Feulner J, Criminisi A, Haynor DR, Konukoglu E (2012) Automatic localization and identification of vertebrae in arbitrary field-of-view CT scans. In: *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2012*. Springer, pp 590-598

9. Glocker B, Zikic D, Konukoglu E, Haynor DR, Criminisi A (2013) Vertebrae localization in pathological spine CT via dense classification from sparse annotations. In: *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2013*. Springer, pp 262-270

10. Michael Kelm B, Wels M, Kevin Zhou S, Seifert S, Suehling M, Zheng Y, Comaniciu D (2013) Spine detection in CT and MR using iterated marginal space learning. *Medical Image Analysis* 17 (8):1283-1292

11. Geman S, Geman D (1984) Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images. *IEEE transactions on pattern analysis and machine intelligence PAMI-6* (6):721-741.

doi:10.1109/TPAMI.1984.4767596

12. Murata N, Takenouchi T, Kanamori T, Eguchi S (2004) Information Geometry of U-Boost and Bregman Divergence. *Neural Computation* 16 (7):1437-1481. doi:10.1162/089976604323057452

13. Nemoto M, Masutani Y, Hanaoka S, Nomura Y, Yoshikawa T, Hayashi N, Yoshioka N, Ohtomo K (2011) A unified framework for concurrent detection of anatomical landmarks for medical image understanding. In: *SPIE Medical Imaging 2011*. 7962, 03/14/2011 2011. pp 79623E-79623E-79613. doi:10.1117/12.878327

14. Hanaoka S, Masutani Y, Nemoto M, Nomura Y, Miki S, Yoshikawa T, Hayashi N, Ohtomo K A Multiple Anatomical Landmark Detection System for Body CT Images. In: *Computing and Networking (CANDAR), 2013 First International Symposium on, 2013*. IEEE, pp 308-311

15. Dikmen M, Zhan Y, Zhou XS (2008) Joint detection and localization of multiple anatomical landmarks through learning. In: *Proc. SPIE, Medical Imaging 2008, 17/03/2008 2008*. pp 691538-691538-691539

16. Crimi A, Sporring J, de Bruijne M, Lillholm M, Nielsen M (2009) Prior knowledge regularization in statistical medical image tasks. Program Committee• John Ashburner (University College London)• Sylvain Bouix (Harvard Medical School)• Tim Cootes (University of Manchester):1

17. Bishop CM (2006) *Pattern Recognition and Machine Learning (Information Science and Statistics)*. Springer-Verlag New York, Inc.,

18. Nemoto M, Masutani Y, Hanaoka S, Nomura Y, Ohtomo K, Miki S, Yoshikawa T, Hayashi N Post-processing of Anatomical Landmark Detection: Distance Error Reduction by Pictorial Structure Matching-Based

Method. In: Computing and Networking (CANDAR), 2013 First International Symposium on, 4-6 Dec. 2013 2013. pp 316-319. doi:10.1109/CANDAR.2013.56

19. Konin GP, Walz DM (2010) Lumbosacral transitional vertebrae: classification, imaging findings, and clinical relevance. *AJNR American journal of neuroradiology* 31 (10):1778-1786. doi:10.3174/ajnr.A2036

20. Connor SE, Shah A, Latifoltojar H, Lung P (2013) MRI-based anatomical landmarks for the identification of thoracic vertebral levels. *Clinical radiology* 68 (12):1260-1267. doi:10.1016/j.crad.2013.07.009

21. Paik NC, Lim CS, Jang HS (2013) Numeric and morphological verification of lumbosacral segments in 8280 consecutive patients. *Spine* 38 (10):E573-578. doi:10.1097/BRS.0b013e31828b7195

22. Hanaoka S, Masutani Y, Nemoto M, Nomura Y, Yoshikawa T, Hayashi N, Yoshioka N, Ohtomo K Probabilistic modeling of landmark distances and structure for anomaly-proof landmark detection. In: *Proceedings of the Third International Workshop on Mathematical Foundations of Computational Anatomy 2011*, 2011. pp 159-169

23. Ishikawa H Higher-order clique reduction in binary graph cut. In: *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on, 20-25 June 2009 2009*. pp 2993-3000. doi:10.1109/CVPR.2009.5206689



# Chapter 3 Automatic definition of landmarks

A fully automatic method to define anatomically meaningful landmarks is presented. Firstly, I assume that possible anatomical landmark points must be registered correctly and consistently in most of volume triplets by a given image registration algorithm. Under this assumption, a novel landmark-ness criterion named triangular consistency criterion (TCC) is introduced. Landmarks are determined as points with sufficiently small TCCs in most of volume triplets. The proposed method was evaluated with 50 whole torso CT volumes and 50 landmarks were automatically defined.

## 3.1 Introduction

Anatomical landmarks in medical imaging have a wide variety of

applications. For example, automatic detection of landmarks is frequently used in initialization of statistical shape model (SSM)-based segmentation process for organs [1]. However, it is very time-consuming to define anatomical landmarks manually, as well as inputting anatomical landmark positions in medical images (e.g. for machine learning).

On the other hand, we can also use non-anatomically defined landmarks such as SIFT-based ones [2]. However, using SIFT-like landmarks is sometimes difficult due to its limited inter-modality and inter-individual correspondence and also its limited reproducibility among datasets. Especially, SIFT-based landmark definition is hard to be used in SSM-based methods, because SSM requires a predetermined set of landmarks which are embedded into the model itself. Therefore, a new methodology will be needed in which new landmarks with anatomical background can be defined automatically.

Similarly, more and more landmarks are demanded in registration [3]. Especially, if a large number of landmarks are accurately detected in the given volume pair, the following registration task may become far easier. The more landmarks, the better. In this sense, automatic landmark definition can play a key role in registration.

In this study, a registration-based method is presented in which new landmarks are defined based on a novel triangular consistency criterion (TCC). TCC can estimate how the target anatomical structure is determined as one point in all of the given training datasets. The proposed method is validated with 50 whole torso CT datasets and the automatically defined landmarks are illustrated. I also analyze each defined landmark and evaluate whether each landmark is defined on any anatomically meaningful structure.

## 3.2 Methods

The proposed method can basically use any arbitrary registration method. In this study, I utilized my domestic landmark-guided registration method based on diffeomorphic demons algorithm [4]. (The details of the registration method are described in Chapter 4.) This registration method can utilize both grayscale image information and the manually-inputted landmark positions.

### 3.2.1 Datasets and registration

Total  $N = 50$  whole-torso CT datasets without intravenous contrast agent were included in this study. Their voxel size was  $0.977 \times 0.977 \times 1.250$  mm. Total 197 landmark positions were inputted for each volume. These landmarks were worked as guides for registering a couple of CT datasets precisely. Note that automatically-defined landmarks will be determined at points at least 20 mm away from these predefined landmarks.

Registration was processed by one-to-one manner; all pairs of datasets were registered using manually-inputted landmarks and grayscale images. In the result, total  $N \times (N - 1) = 2450$  registration results were given.

### 3.2.2 Triangular consistency criterion calculation

Each registration result has its mapping vector field. Let the mapping vector field which deforms  $i$ th dataset to fit to  $j$ th dataset be  $\mathbf{M}_{ij}(\mathbf{x})$ . That means, the point  $\mathbf{x}$  in the  $i$ th image corresponds to the point  $\mathbf{M}_{ij}(\mathbf{x})$  in the  $j$ th image.

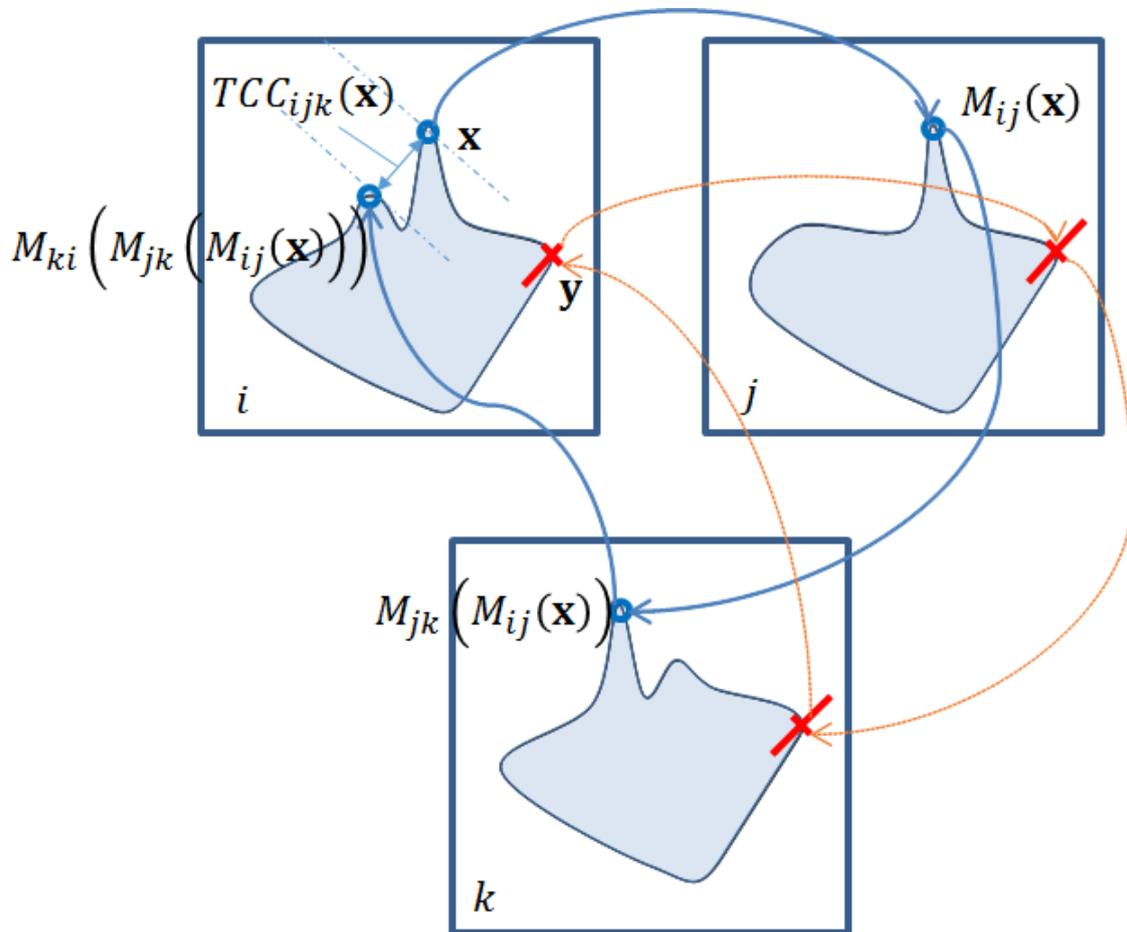
Then, the triangular consistency criterion is introduced (Fig. 1). For each

triplet of datasets  $i$ ,  $j$  and  $k$ , the TCC value is defined as  $TCC_{ijk}(\mathbf{x}) = \left| \mathbf{x} - \mathbf{M}_{ki} \left( \mathbf{M}_{jk} \left( \mathbf{M}_{ij}(\mathbf{x}) \right) \right) \right|$ . That means, the TCC value evaluates the inconsistency of the given three mapping vector fields. Small TCC means that the corresponding points are conserved during registration and thus I regard it as a possible landmark point detected on some anatomically prominent structure. In this study, the registration result is regarded as consistent if and only if the TCC is lesser than a threshold,  $d = 5$  mm (Figs. 2 and 3).

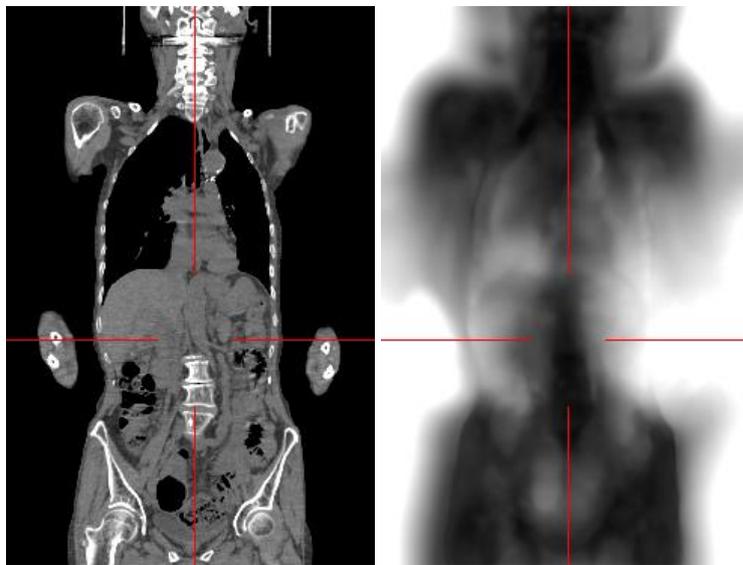
### 3.2.3 Automatic landmark definition

For one fixed dataset  $i$ , the sum of counts where TCC is lesser than  $d$  was calculated as  $S_i(\mathbf{x}) = \frac{1}{(N-1)(N-2)} \sum_{k=1, i \neq k}^N \sum_{j=1, i \neq j, j \neq k}^N I(TCC_{ijk}(\mathbf{x}) < d)$ . Here, function  $I$  is 1 if  $TCC_{ijk}(\mathbf{x}) < d$  or 0 otherwise. An example of  $S_i(\mathbf{x})$  is shown in Fig. 2(b).

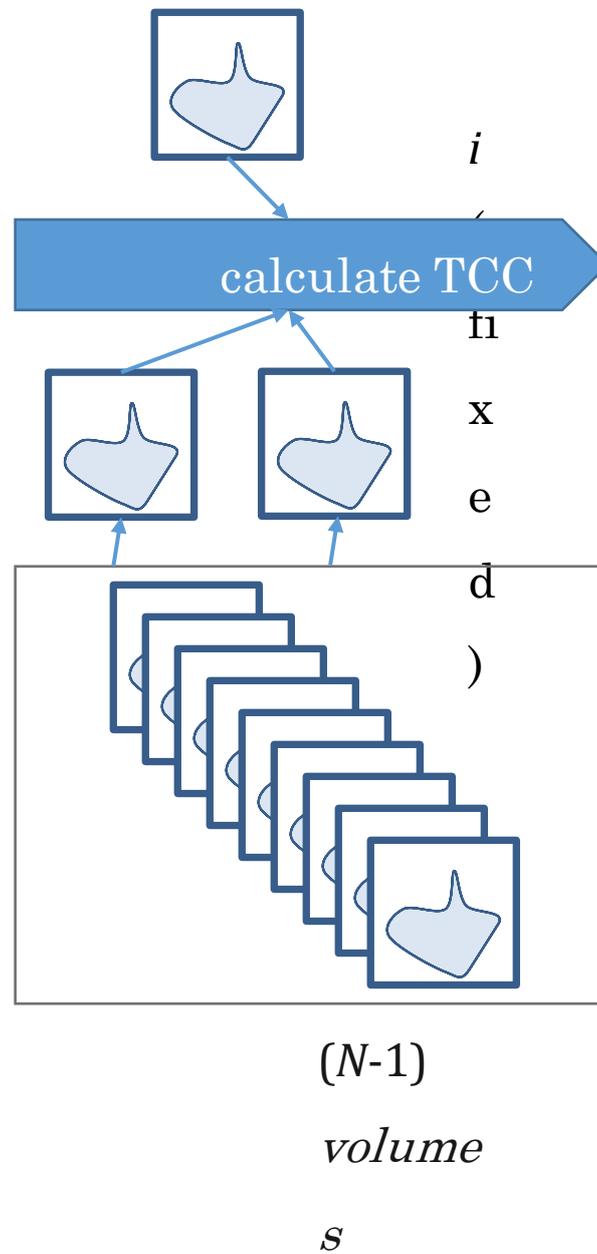
In  $S_i(\mathbf{x})$ , the algorithm searched for new landmarks by a sequential manner. Firstly, all the local maxima in  $S_i(\mathbf{x})$  are extracted. The local maxima within  $d_{elim} = 20$  mm from any of already-defined landmarks are eliminated. Also, local minima with  $S_i(\mathbf{x}) < 0.3$  are also eliminated. Then, the local maximum with the largest  $S_i(\mathbf{x})$  value is newly defined as a landmark. This process is repeated until no residual local maxima meet the criterion (Fig. 3).



**Fig.1.** The definition of the TCC. It is defined using three images and deformation fields between them. In this example, the distance between the original point  $\mathbf{x}$  and the threefold-moved point  $M_{ki}(M_{jk}(M_{ij}(\mathbf{x})))$  is defined as the  $TCC_{ijk}(\mathbf{x})$ . Note that  $TCC_{ijk}(\mathbf{y})$  is less than  $TCC_{ijk}(\mathbf{x})$  in this example, that means  $\mathbf{y}$  is a better landmark candidate than  $\mathbf{x}$ .



**Fig. 2.** (Left) A coronal CT slice and (Right) the corresponding mean TCC value image.



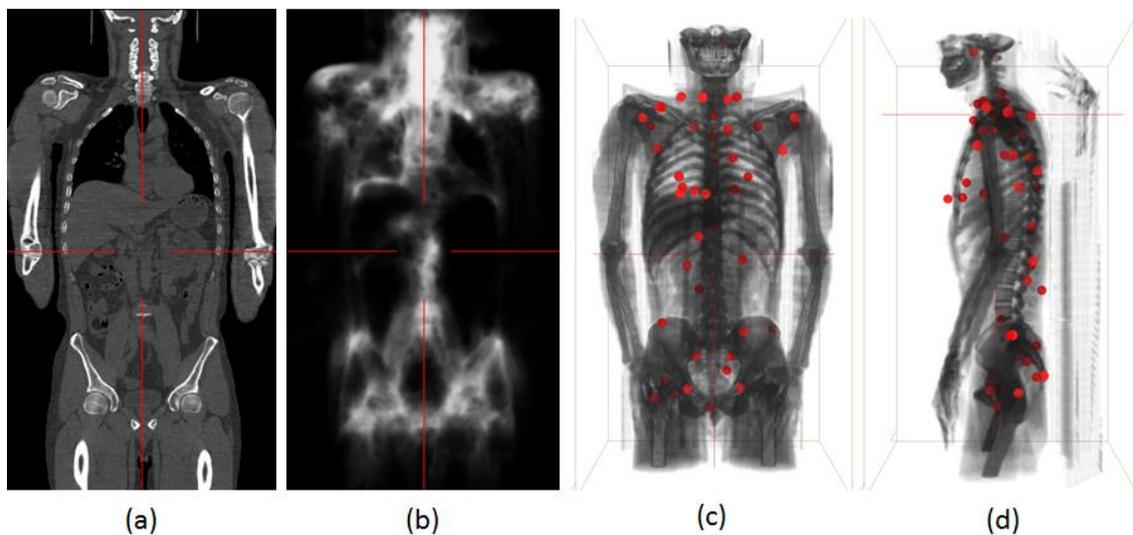
**Fig. 3.** Calculation of the TCC volumes.

### 3.3 Results

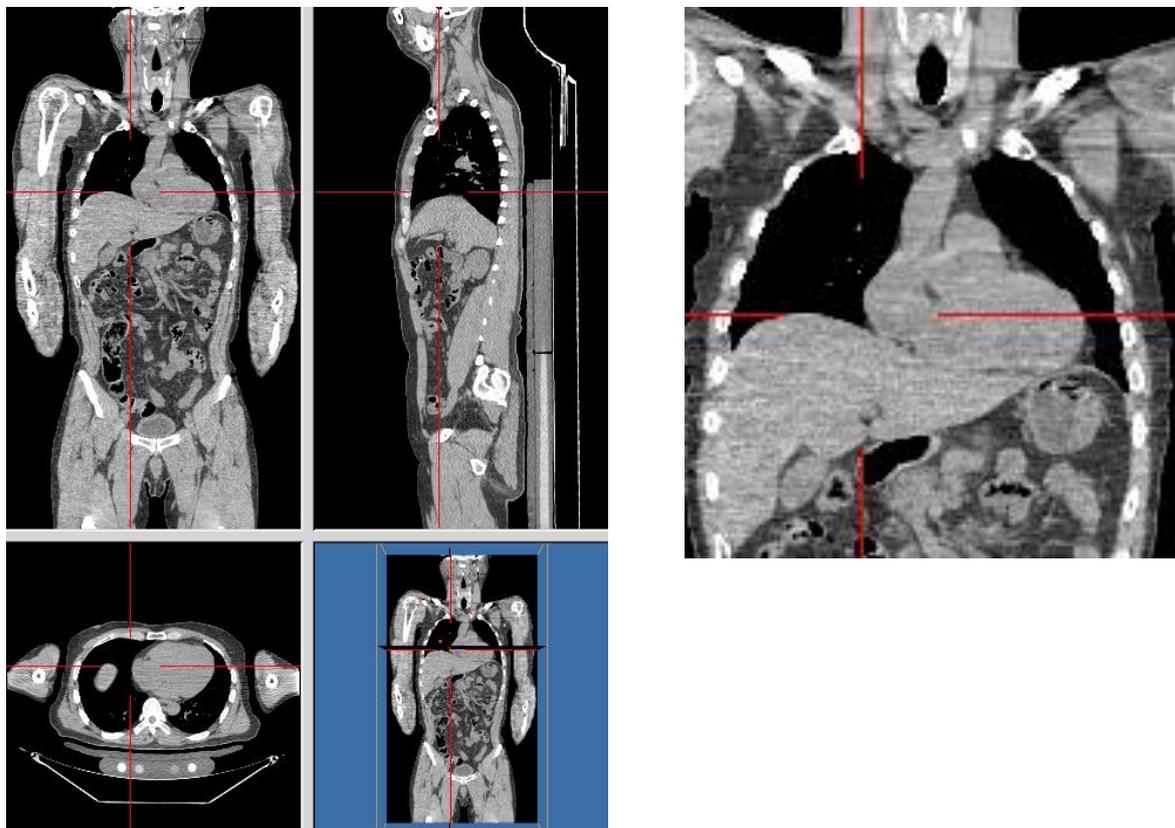
An exemplar result of the automatic landmark definition is shown in Figs. 4

to 9. Total 48 landmarks were defined. The automatically-defined landmarks include the posterior margin of bilateral kidneys, the inferior wall of center of the aortic arch, nasopharyngeal, interior sides of the bilateral femoral heads, bilateral sides of vocal cords, and so on. On the other hand, some landmarks were defined not bilaterally symmetric, which may reflect the fact that my method did not stably detect symmetric anatomical structures. Nevertheless, most of defined landmarks can be interpreted as anatomically meaningful points.

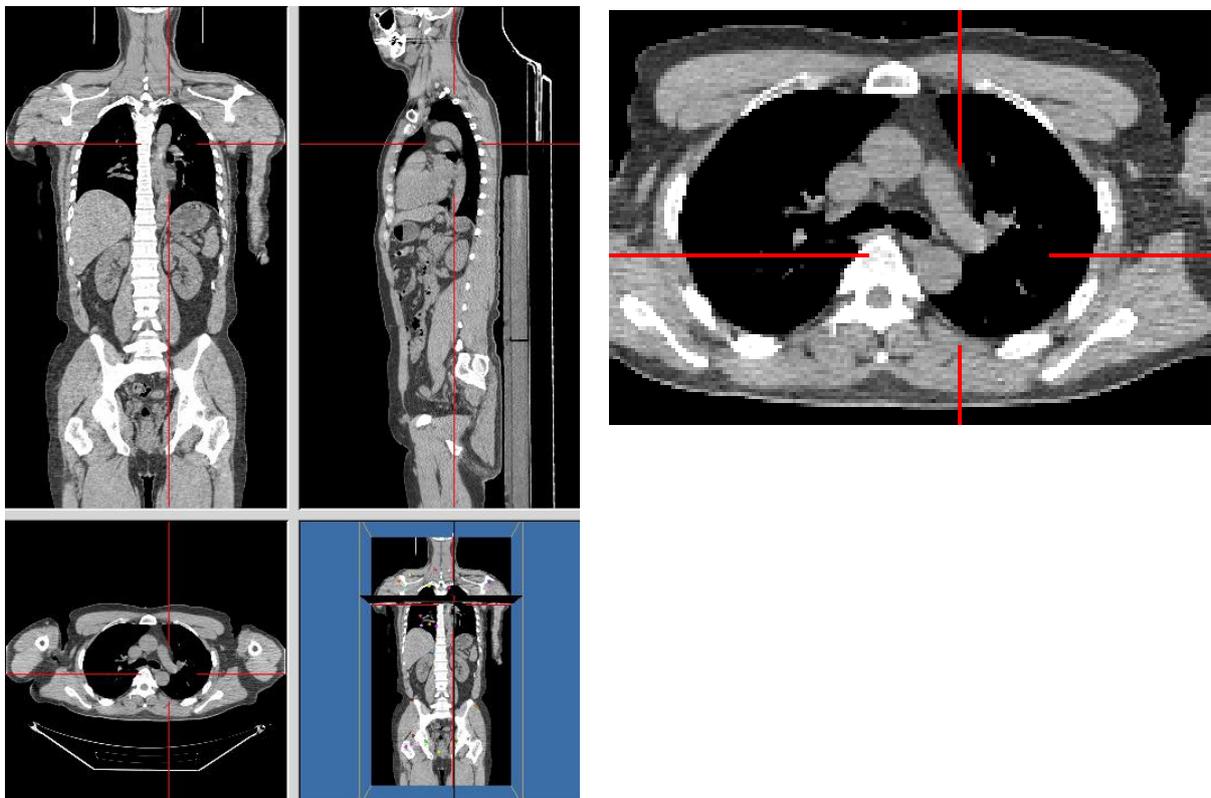
At least 27/48 landmarks can be regarded as a salient position and can be given an anatomical name (e.g., as the angle between the inferior vena cava and the bottom surface of the liver; Fig 7). Many landmarks were defined on bony surfaces, but some landmarks were defined on soft tissues such as the liver.



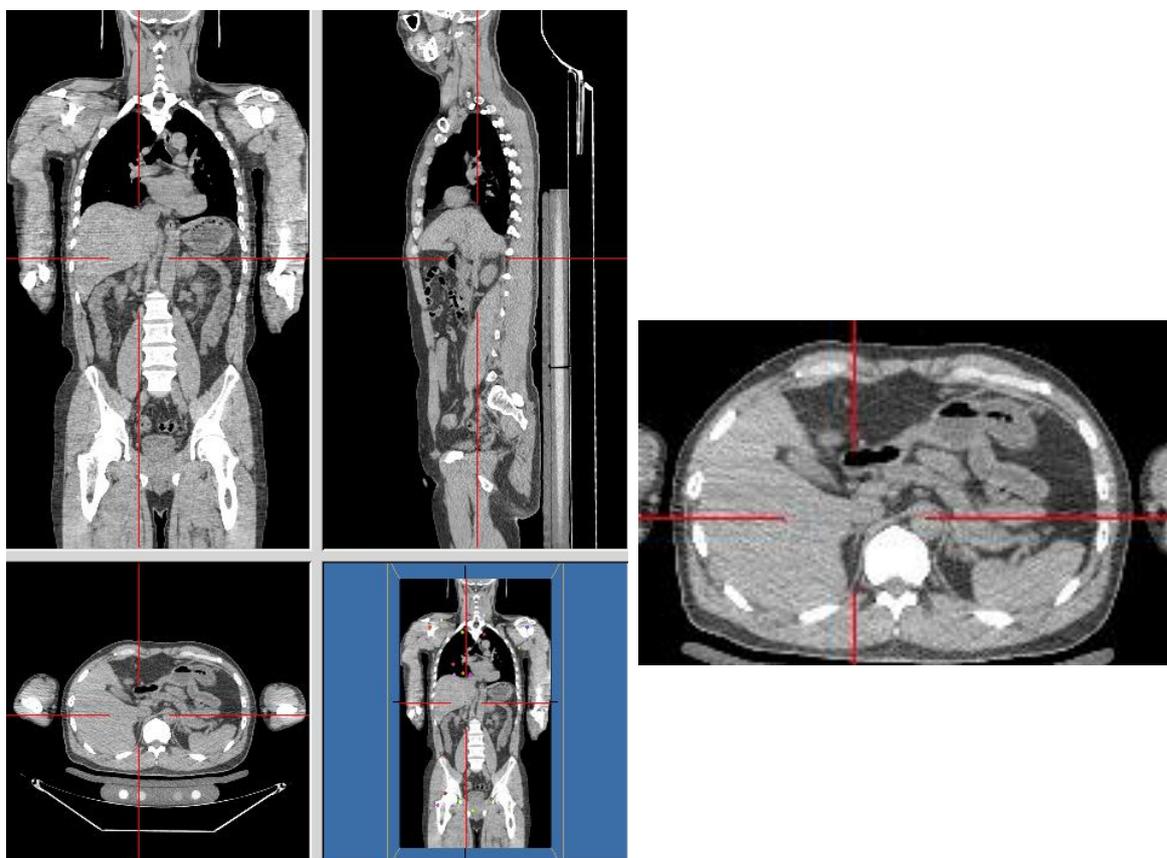
**Fig. 4.** (a) An example of original CT volume, a coronal cross-section. (b) The corresponding  $S_i$  values. (c) Automatically defined landmarks, frontal view. (d) Lateral view.



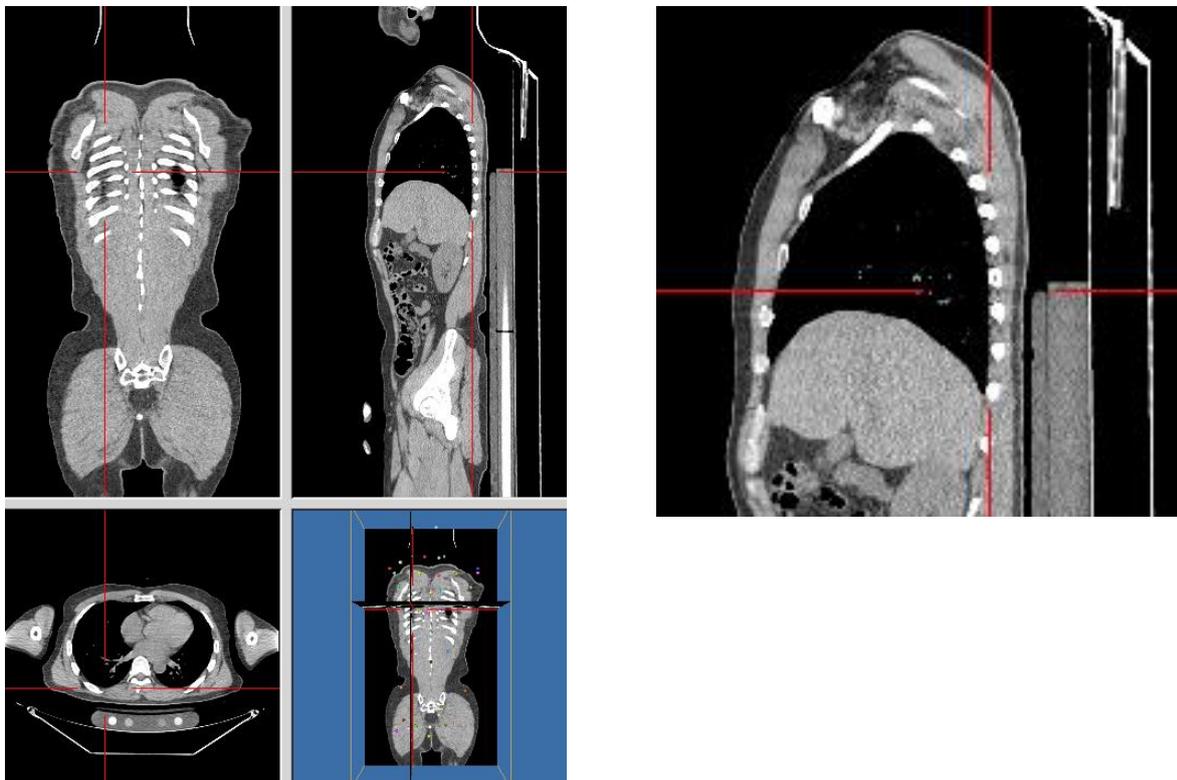
**Fig. 5.** An example of automatically defined landmark (rt. Cardiophrenic angle).



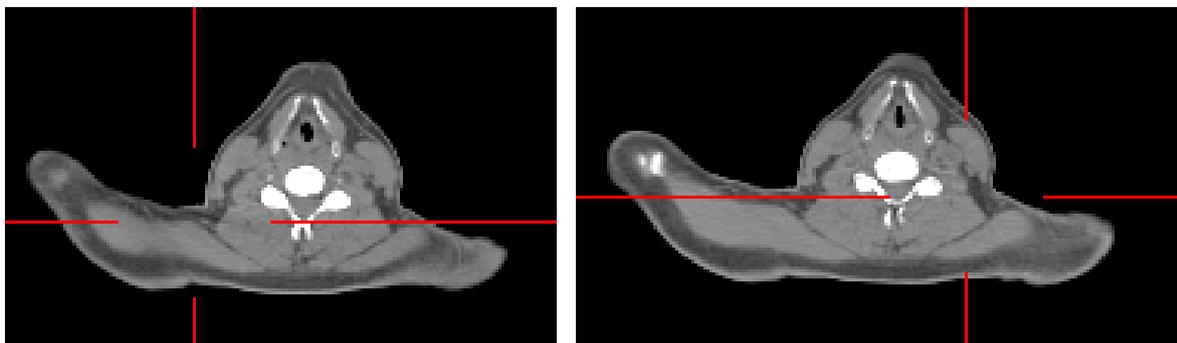
**Fig. 6.** An example of automatically defined landmark (the proximal point of the aorta and the left pulmonary artery).



**Fig. 7.** An example of automatically defined landmark (the angle between the inferior vena cava and the bottom surface of the liver).



**Fig. 8.** An example of automatically defined landmark. This landmarks was difficult to interpret as a salient anatomical entity, but can possibly be named as the posterior margin of the left thoracic cavity.



**Fig. 9.** Examples of automatically defined landmarks (the angles between the shoulders and the neck).

## 3.4 Discussion

A novel method to define landmarks from a large CT dataset was presented. In the method, stably registered positions in the given images are extracted as landmarks, using TCC criterion. Therefore, the landmark definition results are largely affected by the registration method used. Although in this study landmark-guided demon's algorithm worked well, it will be my future work to test other registration methods. On the other hand, if the registration method used can handle different modalities (e.g. CT and MRI), the algorithm can handle mixture of datasets with multiple modalities. Therefore, it will also be my future work to test the method with multiple modalities.

This study has several limitations. First, any quantitative analysis of the results has been performed. It is difficult to validate the results because the automatically defined landmarks do not have manually inputted "ground truth" positions. One possible way is to input ground truth landmark points manually, so that I can compare the ground truth positions and automatically defined positions. Furthermore, I can compose detectors for the newly defined landmarks, train the detectors as shown in Chapter 2, and evaluate their accuracy. It is expected that good detection accuracies are shown for well-defined landmarks, and vice versa. Another possible way to evaluate meaningfulness of each defined landmark is to use it in other applications. For example, I can add newly defined landmarks to my landmark-guided image registration system. It is probable that the registration accuracy will be improved by adding automatically defined landmarks.

The second limitation is that the defined landmarks are mainly located on the bony or skin structures and not on the guts or the pancreas. Generally

speaking, landmarks on the deformable soft tissue structures such as the guts and the pancreas have more impact in applications than bony landmarks. In my result, the landmark definition framework avoided such soft tissue organs, because the registration method used cannot accurately register such organs and thus the TCC values became large. It is a dilemma that such difficult organs were avoided whereas many easy bony landmarks were automatically defined. As described above, the proposed method relies upon the registration method used, and it is very difficult even for modern registration techniques to register such soft-tissue organs. On the other hand, it is even difficult for medical experts to manually define intestinal or pancreatic landmarks because of their deformability and wide variety of shapes. In this sense, it is not surprising that the proposed method avoided such difficult organs.

### 3.5 Conclusion

A novel method to automatically define landmarks was presented. In evaluation with 50 whole torso CT datasets, total 48 landmarks were automatically and successfully defined.

## References

1. Heimann T, Van Ginneken B, Styner MA, Arzhaeva Y, Aurich V, Bauer C, Beck A, Becker C, Beichel R, Bekes G (2009) Comparison and evaluation of methods for liver segmentation from CT datasets. *Medical Imaging, IEEE Transactions on* 28 (8):1251-1265
2. Lowe DG (1999) Object recognition from local scale-invariant features. In: *Computer Vision, 1999. The Proceedings of the Seventh IEEE International Conference on*, 1999. pp 1150-1157 vol.1152
3. Rohr K, Stiehl HS, Sprengel R, Buzug TM, Weese J, Kuhn MH (2001) Landmark-based elastic registration using approximating thin-plate splines. *IEEE Transactions on Medical Imaging* 20 (6):526-534
4. Vercauteren T, Pennec X, Perchant A, Ayache N (2008) Symmetric log-domain diffeomorphic registration: A demons-based approach. In: *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2008*. Springer, pp 754-761



# Chapter 4 Multiatlas-based segmentation of the vertebral and pelvic bones

A fully automatic multiatlas-based method for segmentation of the spine and pelvis in a torso CT volume is proposed. A novel landmark-guided diffeomorphic demons algorithm is used to register a given CT image to multiple atlas volumes. This algorithm can utilize both grayscale image information and given landmark coordinate information optimally.

The segmentation has four steps. Firstly, 170 bony landmarks are detected in the given volume. Using these landmark positions, an atlas selection procedure is performed to reduce the computational cost of the following registration. Then the chosen atlas volumes are registered to the given CT image. Finally, voxelwise label voting is performed to determine the final

segmentation result.

The proposed method was evaluated using 50 torso CT datasets as well as the public SpineWeb dataset. In the result, a mean distance error of  $0.59 \pm 0.14$  mm and a mean Dice coefficient of  $0.90 \pm 0.02$  were achieved for the whole spine and the pelvic bones, which are competitive with other state-of-the-art methods.

From the experimental results, the usefulness of the proposed segmentation method was validated.

## 4.1 Introduction

Automatic segmentation of the spinal bones in computed tomography (CT) images has various applications, such as automatic detection of bone metastasis [1], radiation planning [2] and surgical planning [3]. For these applications, a highly reliable and precise segmentation method for the whole spine is desired. However, automatic segmentation of all 24 vertebrae is still an open problem for which many new methods have been reported [4-9].

One of the reasons for the difficulty of segmenting vertebral bones is the high complexity of vertebral shapes and topologies. For example, lumbar vertebrae have long and thin bony processes of various lengths and angles. All the vertebrae have at least one hole (through which the spinal cord passes), but most of the cervical vertebrae have two more holes (for vertebral arteries). To accurately segment these fine and complex structures using a statistical shape model (SSM), a large number of degrees of freedom (DOFs) of the model may be required. However, more DOFs result in more local minima

(supposing that the model-fitting problem is nonconvex) and may increase the probability of failure in model fitting. To tackle this problem, some researchers have improved SSMs by making them adaptable to large morphological variation [4,5].

Another difficulty is the repetitiveness of vertebral shapes. Since neighboring vertebrae have similar shapes, it is not easy to identify each vertebra (e.g., the 4th cervical, 6th thoracic or 3rd lumbar). Identification of the vertebra is required prior to most state-of-the-art segmentation procedures such as SSM fitting, and a failure to identify them will cause a cranial or caudal shift of the identification result. A cranial/caudal shift leads to large errors in the final segmentation result. Such a shift can also occur when applying automatic image registration algorithms such as free-form deformation (FFD) or demons algorithms to spatially fit a pair of spinal bone structures. The difficulty of registering the spine is a major reason why multiatlas-based segmentation approaches [10], which are widely used for other human organs [11] [12], have seldom been applied to multiple spinal bone segmentation problems.

A multiatlas method is a registration-based segmentation method introduced in [13] and [14]. In a multiatlas method, all training datasets (manually labeled by an expert) are registered to a given unseen image and then the labels of the training datasets are propagated to the target image. The final segmentation result is built through integration (e.g., voting) of all the registered label images. The multiatlas approach has the flexibility to better capture anatomical variation and thus has superior segmentation accuracy to other methods [10]. One of the disadvantages of a multiatlas method is its

high computational cost; the most naïve implementation has a computational cost proportional to the number of training datasets. To avoid this problem, atlas selection is preferred. This is a technique to select and reduce the number of atlases used according to the reliability of each atlas. Simultaneously, the weight of each atlas in the following decision fusion phase may be decided. Atlas selection can be performed before or after registration, and the weights of atlases can be determined globally or locally. Note that if the weights of some atlases become zero globally, these atlases can be omitted in the following registration process. Thus, a global approach is preferred to reduce the computational cost. On the other hand, applying different weights to local regions leads to better segmentation accuracy instead of a nonsignificant reduction in the computational cost. In particular, to reduce the computational cost of the registration, it is sometimes effective to roughly preregister all atlases and then perform atlas selection using these preregistered images [10]. After atlas selection, the high-cost precise registration algorithm is applied to the small number of selected atlases. The most frequently used atlas selection criterion is the image similarity between two volumes. Recently, however, several sophisticated atlas selection techniques based on machine learning (e.g., random forests) have been reported [15-17]. For example, in [15] the final labeling performance itself is estimated by an algorithm and used as the atlas selection criterion.

A registration method used in the multiatlas framework must have sufficient accuracy as well as computational speed. For example, it should be able to correctly register a wide variety of fine and thin tubelike structures (such as spinal processes) and thin platelike structures (such as endplates and other cortical bones). To register such thin structures, the deformation field

calculated in the registration process must have a high number of DOFs.

Demons algorithms are voxelwise registration algorithms and were introduced by Thirion [18]. A diffeomorphic version of the demons algorithm was also presented by Vercauteren et al. [19]. In demons algorithms, all the voxels in a given volume have an independent deformation field vector, which is iteratively updated using the local intensity difference and Jacobian matrices. Since the value of the deformation vector field is determined at every voxel, the demons algorithm has an extremely high number of DOFs that is proportional to the total number of voxels. On the other hand, one of the disadvantages of demons algorithms is that the registration problem is solved in a steepest-descent manner; thus, the algorithms may be affected by local minima of the cost function. Because of its weakness against local minima, the quality of the given initial condition greatly affects the registration result. In other words, the algorithms are not well suited for a problem in which a large deformation is needed to register the two images. The algorithms also do not work well for repetitive spinal shapes, which cause a sequence of local minima. This problem may be partially solved by using a multiresolutional approach in which a pair of images are first roughly registered in a coarser scale and then precisely registered in a finer scale. However, in my experience, it is very difficult for a conventional demons algorithm to precisely register the spinal columns of different subjects.

Among the related spinal segmentation studies, Klinder et al. [6] first reported an SSM-based method in which all 24 vertebrae are identified and segmented automatically from 3-D CT volumes. Other SSM-based methods for thoracolumbar vertebrae have been reported [4,5,7,8]. Among them,

Kadoury et al. reported a method in which the vertebral identification and segmentation problems are formulated as a single higher-order Markov random field (HOMRF) problem [5]. The method was successfully applied to scoliotic spines and achieved an interlandmark distance error of  $1.6 \pm 0.6$  mm for CT images. Forsberg et al. reported a multiatlas-based thoracolumbar spine segmentation method [20]. After the vertebral positions and poses were estimated, the spine was divided into four subregions (L5-L1, L1-T9, T9-T5 and T5-T1) and each atlas was registered to them. Because they used a relatively small set of 10 atlases, atlas selection was not performed. They achieved a mean Dice index of  $0.94 \pm 0.03$ . Recently, Wang et al. [21] reported a novel method for thoracolumbar vertebral bone segmentation in which a multidimensional support vector regressor is used for direct regression from image features to the target object boundary. Other state-of-the-art methods and their segmentation performances using a publicly available thoracolumbar CT dataset (SpineWeb; <http://spineweb.digitalimaginggroup.ca/>) have been reported in [22]. However, to the best of my knowledge, no method that can simultaneously segment all the structures of the spinal and pelvic bones has been reported.

So far, a few landmark-guided registration methods have been reported [23-26]. Among them, a method for registering a pair of landmark sets distributed on a lumbar vertebral bone surface was reported in [26]. Using the fact that a lumbar spine has a hole, the task was formulated as a registration between two genus-one surfaces (i.e., tori). On the other hand, a demons algorithm using automatically generated landmarks was presented in [24]. After landmarks were automatically generated on the bone or skin borders, a landmark-stabilized demons image registration was performed.

In this study, a new method for segmentation of all the vertebral and pelvic bones in a CT volume is presented. In the method, multiple atlases are registered to the target unseen volume by a novel landmark-guided diffeomorphic demons algorithm. Given a set of vertebral and pelvic landmark positions (such as those obtained by my previously developed landmark detection method [27,28], which can detect 170 anatomically annotated landmarks, as described in Chapter 2), the algorithm can register two volumes using both grayscale image information and the landmark position information simultaneously. In each iteration, not only the diffeomorphic deformation field but also the trajectory of each landmark is calculated using a “speed image”. Each trajectory moves a landmark position in one of the given images to the corresponding landmark position in the other image. From these landmark trajectories, a landmark-derived update vector field is calculated to fit the two landmark positions. Then, the landmark-derived update field and the grayscale-image-derived update field are summed and used to update the speed image, from which the diffeomorphic deformation field is iteratively updated. Using this landmark-guided method, optimally selected multiple atlases are registered to the target image, which is followed by voxelwise voting to calculate the final segmentation result.

The contributions of this study are as follows:

- A novel landmark-guided demons registration approach is presented. Using the coordinates of corresponding landmark pairs, it can register two shapes with large deformations.
- The proposed method is evaluated using 50 whole-torso CT datasets. Using the proposed method with my automatic landmark detection system,

the spinal and pelvic bones are automatically segmented. To the best of my knowledge, this is the first study in which a segmentation method for the whole spine and the pelvic bones has been evaluated.

In the rest of this paper, the proposed landmark-guided diffeomorphic demons algorithm is described, followed by the atlas selection method used in this study. Then the hyperparameter optimization method and evaluation method are described and the results of the evaluation are given. Finally, the characteristics, advantages and disadvantages of the proposed segmentation framework are discussed.

## 4.2 Methods

### 4.2.1 Landmark-guided diffeomorphic demons algorithm

My registration algorithm is based on the log-domain demons algorithm [19]. In the log-domain demons algorithm, the diffeomorphism is ensured by deriving the deformation field from a speed image that represents an infinitesimal deformation. The speed image is iteratively updated using the given grayscale volumes and their Jacobians. The log-domain demons algorithm is described in detail in the following section, which is followed by details of the proposed landmark-guided demons algorithm.

#### 4.2.1.1 Log-domain demons algorithm

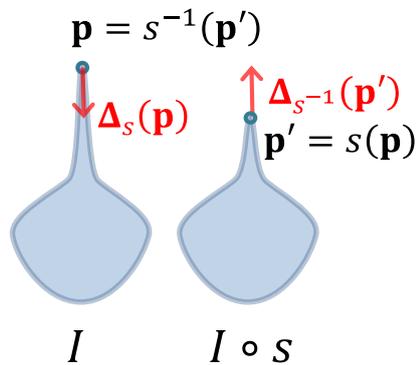
In the log-domain demons algorithm, the deformation vector field is represented as the exponential of another vector field, namely, the speed vector field, so that the diffeomorphism of the deformation field is guaranteed. Here, the exponential calculation is derived from Lie group theory and can be

calculated efficiently as described in [29]. Thus, we need to calculate the speed vector field to determine and update the deformation field.

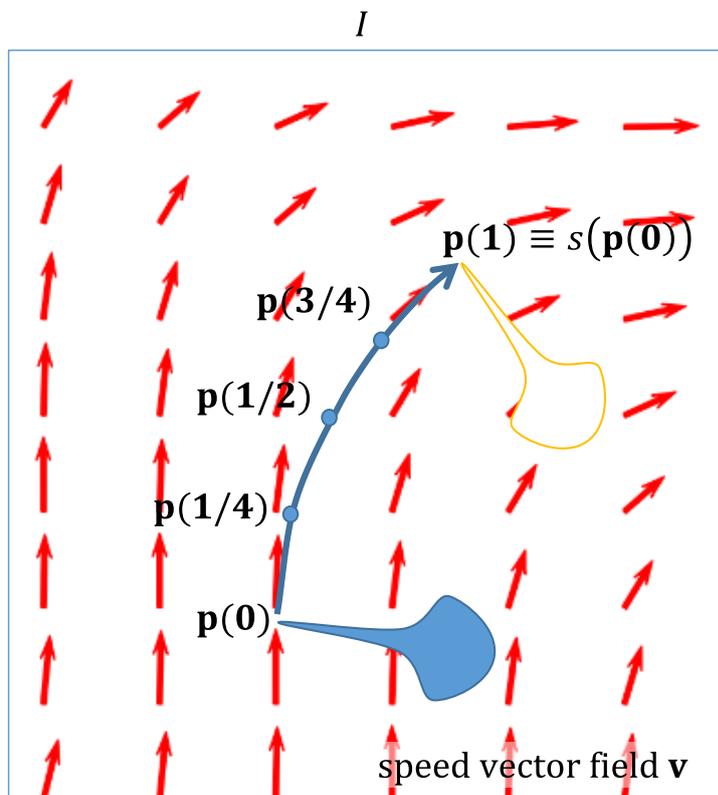
Consider three-dimensional image  $I$ , and let  $I(\mathbf{p}) \in \mathbb{R}$  be the grayscale voxel value at a point  $\mathbf{p} \in \mathbb{R}^3$  in  $I$ . A deformation  $s: \mathbb{R}^3 \rightarrow \mathbb{R}^3$  is defined by the corresponding deformation vector field  $\Delta_s(\cdot) \in \mathbb{R}^3$  such that point  $\mathbf{p}$  is warped to point  $\mathbf{p}' = s(\mathbf{p}) = \mathbf{p} + \Delta_s(\mathbf{p})$  (see Fig. 1). Then image  $I \circ s$ , which is image  $I$  after being deformed by  $s$ , can be calculated using the formula  $(I \circ s)(\mathbf{p}') = I(\mathbf{p})$ . Suppose that a deformation  $s^{-1}$  is defined as the inverse deformation of  $s$  (i.e.,  $s \circ s^{-1} = Id$ , the identity mapping). Then, using the corresponding deformation vector field  $\Delta_{s^{-1}}$ , the coordinates of  $\mathbf{p}$  in image  $I$  can be calculated from  $\mathbf{p}'$  by  $\mathbf{p} = s^{-1}(\mathbf{p}') \stackrel{\text{def}}{=} \mathbf{p}' + \Delta_{s^{-1}}(\mathbf{p}')$ . Thus, for an arbitrary point  $\mathbf{p}'$ ,

$$(I \circ s)(\mathbf{p}') = I(\mathbf{p}' + \Delta_{s^{-1}}(\mathbf{p}')) \quad (1)$$

is satisfied. This means that the inverted dislocation vector field  $\Delta_{s^{-1}}$  is required to calculate the deformed grayscale image  $I \circ s$  from the original image  $I$ . On the other hand, the original dislocation field  $\Delta_s$  is required to calculate the destination point  $\mathbf{p}' = \mathbf{p} + \Delta_s(\mathbf{p})$  from the original point  $\mathbf{p}$ . Thus, to warp both the landmark positions and the grayscale image, we require both  $\Delta_s$  and  $\Delta_{s^{-1}}$ . However, calculating the inverted dislocation vector field  $\Delta_{s^{-1}}$  from  $\Delta_s$  is not a trivial problem.



**Fig. 1.** Definition of deformations and deformation vector fields.



**Fig. 2.** Speed vector field and trajectory of the moving point  $\mathbf{p}(t)$ .

To solve this duality problem and to ensure the invertibility of the deformation field, a stationary speed vector field  $\mathbf{v}(\cdot) \in \mathbb{R}^3$  is introduced.

The vector  $\mathbf{v}(\mathbf{p})$  represents the infinitesimal movement, or moving speed, of point  $\mathbf{p}$ . The vector field  $\mathbf{v}$  can be regarded as a flow; at any moment in time, point  $\mathbf{p}$  moves in the direction of vector  $\mathbf{v}(\mathbf{p})$  and with speed  $|\mathbf{v}(\mathbf{p})|$ . Because the speed vector field is stationary, it does not change with time. On the other hand, point  $\mathbf{p}$  moves along the speed vector field. Suppose that point  $\mathbf{p} \equiv \mathbf{p}(0)$  at time 0 is moved to position  $\mathbf{p}(t)$  at time  $t$ . Then,

$$\mathbf{p}(t) = \mathbf{p}(0) + \int_0^t \mathbf{v}(\mathbf{p}(t')) dt', \quad (2)$$

or, in the differentiated form,

$$\frac{\partial \mathbf{p}(t)}{\partial t} = \mathbf{v}(\mathbf{p}(t)) \quad (3)$$

is satisfied. Then  $\mathbf{p}(1)$ , the position of point  $\mathbf{p} \equiv \mathbf{p}(0)$  after one unit of time, is redefined as  $s(\mathbf{p})$  (Fig. 2). That is,

$$s(\mathbf{p}) \stackrel{\text{def}}{=} \mathbf{p}(1) = \mathbf{p}(0) + \int_0^1 \mathbf{v}(\mathbf{p}(t)) dt. \quad (4)$$

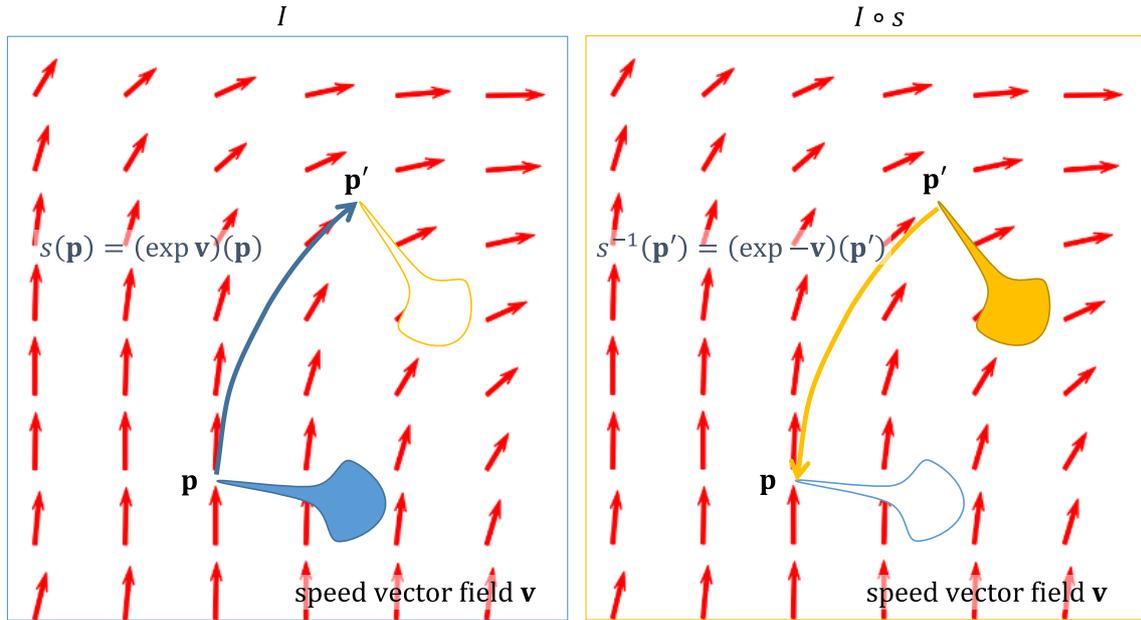
Note that Eq. (4) redefines the deformation field  $s$  for an arbitrary point  $\mathbf{p}$  using the speed vector field  $\mathbf{v}$ . Because  $s(\mathbf{p}) \stackrel{\text{def}}{=} \mathbf{p} + \Delta_s(\mathbf{p})$ , the deformation vector field  $\Delta_s$  can also be calculated from the speed field  $\mathbf{v}$  as follows:

$$\Delta_s(\mathbf{p}) = \int_0^1 \mathbf{v}(\mathbf{p}(t)) dt. \quad (5)$$

Note that the inverted deformation field  $s^{-1}$  can be calculated from the inverted speed field  $-\mathbf{v}$ . Because inversion of the deformation means inverting the temporal progress, it is equivalent to inverting the speed vector

field. Suppose that another point  $\mathbf{p}' \equiv \mathbf{p}'(0)$  at time 0 is moved to position  $\mathbf{p}'(t)$  at time  $t$  along the inverted vector field  $-\mathbf{v}$ . In other words, suppose that  $\mathbf{p}' \equiv \mathbf{p}'(0)$  and  $\mathbf{p}'(t) = \mathbf{p}'(0) + \int_0^t -\mathbf{v}(\mathbf{p}'(t'))dt'$ . Then,  $s^{-1}(\mathbf{p}') = \mathbf{p}'(1) = \mathbf{p}'(0) + \int_0^1 -\mathbf{v}(\mathbf{p}'(t))dt$  is satisfied. This means that both  $s$  and  $s^{-1}$  can be directly calculated from  $\mathbf{v}$ .

Using a term in Lie group theory, the mapping  $\mathbf{v} \mapsto s$  is called the exponential of the speed vector field. Therefore, in the description below, the definition  $s \stackrel{\text{def}}{=} \exp \mathbf{v}$  is used. Note that  $s^{-1} = \exp(-\mathbf{v})$  is satisfied as described above (Fig. 3).



**Fig. 3.** Speed vector field and its exponential.

It is known that the exponential calculation  $\mathbf{v} \mapsto \exp(\mathbf{v})$  can be performed efficiently as a sequence of compositions of finite transformations [29]. Note that halving the speed and doubling the moving time will lead to the same result.

Thus,

$$\left(\exp \frac{1}{2} \mathbf{v}\right) \circ \left(\exp \frac{1}{2} \mathbf{v}\right) = \exp \mathbf{v} \quad (6)$$

is satisfied. Here,  $\circ$  denotes the composition of the deformations. Thus,  $\exp \mathbf{v}$  can be calculated as

$$\exp \mathbf{v} = \underbrace{\left(\exp \frac{1}{2^N} \mathbf{v}\right) \circ \left(\exp \frac{1}{2^N} \mathbf{v}\right) \circ \dots \circ \left(\exp \frac{1}{2^N} \mathbf{v}\right)}_{2^N}. \quad (7)$$

Under the assumption that  $\frac{1}{2^N} \mathbf{v}$  is sufficiently small, the approximation  $\exp \frac{1}{2^N} \mathbf{v} \cong Id + \frac{1}{2^N} \mathbf{v}$  can be used. That means that  $\left(\exp \frac{1}{2^N} \mathbf{v}\right)(\mathbf{p}) \cong \mathbf{p} + \frac{1}{2^N} \mathbf{v}(\mathbf{p})$  at any point  $\mathbf{p}$ . Then,  $\exp \mathbf{v}$  can be calculated using a sequence of compositions of deformation fields:

$$\begin{aligned} \exp \frac{1}{2^{N-1}} \mathbf{v} &= \left(\exp \frac{1}{2^N} \mathbf{v}\right) \circ \left(\exp \frac{1}{2^N} \mathbf{v}\right) \\ \exp \frac{1}{2^{N-2}} \mathbf{v} &= \left(\exp \frac{1}{2^{N-1}} \mathbf{v}\right) \circ \left(\exp \frac{1}{2^{N-1}} \mathbf{v}\right) \\ &\vdots \\ \exp \mathbf{v} &= \left(\exp \frac{1}{2} \mathbf{v}\right) \circ \left(\exp \frac{1}{2} \mathbf{v}\right). \end{aligned} \quad (8)$$

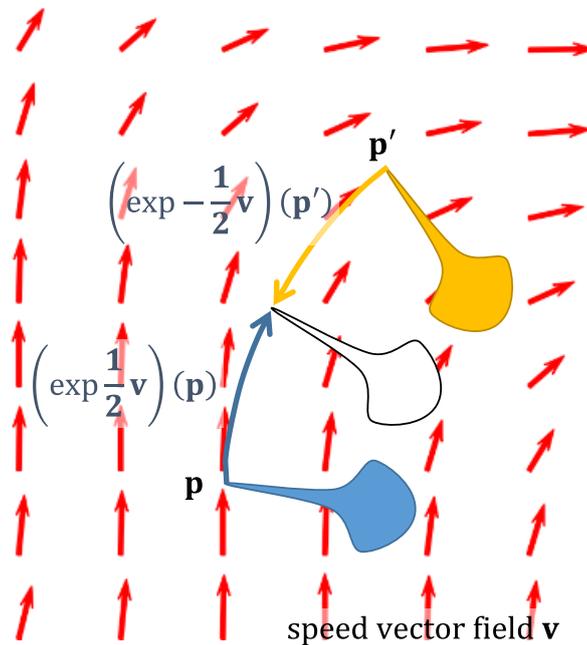
Using this exponential operation, the log domain demons algorithm used in this study is described as follows:

**Algorithm 1. Symmetric log-domain demons algorithm**

- Input: two images  $I_a$  and  $I_b$

- Set the initial speed image  $\mathbf{v}(\mathbf{p}) \leftarrow \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \forall \mathbf{p}$
- Iterate:
  - Calculate forward and backward deformation fields  $\exp \frac{1}{2} \mathbf{v}$  and  $\exp -\frac{1}{2} \mathbf{v}$
  - Calculate deformed images  $I_a^d = I_a \circ \left( \exp \frac{1}{2} \mathbf{v} \right)$  and  $I_b^d = I_b \circ \left( \exp -\frac{1}{2} \mathbf{v} \right)$
  - Compute the demons forces  $\mathbf{u}(\mathbf{p}) = -\frac{\Delta I}{\|\mathbf{J}_p\|^2 + (\Delta I)^2} \cdot \mathbf{J}^p$  where  $\Delta I = I_a^d(\mathbf{p}) - I_b^d(\mathbf{p})$  and  $\mathbf{J}_p = -\frac{1}{2} \left( \nabla I_a^d(\mathbf{p}) + \nabla I_b^d(\mathbf{p}) \right)$
  - For fluid-like regularization let  $\mathbf{u} \leftarrow K_{fluid} \star \mathbf{u}$
  - For diffusion-like regularization let  $\mathbf{v} \leftarrow K_{diff} \star (\mathbf{v} + \mathbf{u})$

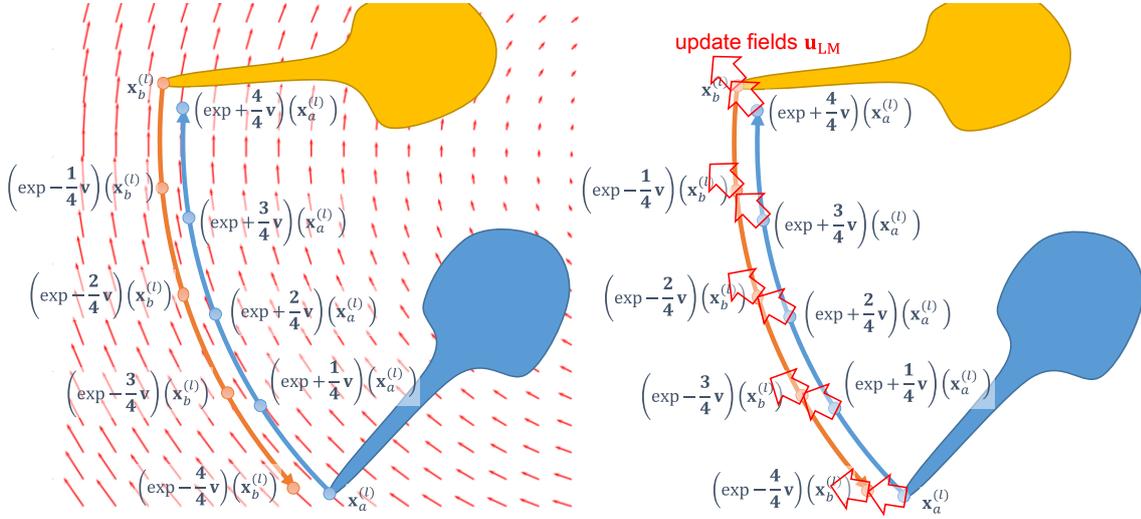
Here, the operators  $K_{fluid} \star$  and  $K_{diff} \star$  denotes the application of Gaussian smoothing filters with appropriate kernel sizes. Upon convergence, the pair of deformed images  $I_a^d$  and  $I_b^d$  become spatially fitted, that is,  $I_a^d(\mathbf{p}) \approx I_b^d(\mathbf{p})$ . Furthermore,  $I_a \circ (\exp \mathbf{v}) \approx I_b$  and  $I_b \circ (\exp -\mathbf{v}) \approx I_a$  are also satisfied (Fig. 4).



**Fig. 4.** Fitting of two objects in the symmetrical log-domain demons algorithm.

#### 4.2.1.2 Proposed landmark-guided demons algorithm

In my proposed landmark-guided demons algorithm, not only grayscale image information but also landmark position information is used to update the speed vector field. Suppose that a total of  $L$  landmarks to be fitted are determined in each image. Let the positions of the  $l$ th landmark in images  $I_a$  and  $I_b$  be  $\mathbf{x}_a^{(l)}$  and  $\mathbf{x}_b^{(l)}$ , respectively. Then, as described in (2.1.1), we can calculate and draw the trajectory of each landmark along the speed vector field  $\mathbf{v}$  (Fig. 5 left).



**Fig. 5.** Fitting of two landmarks in the symmetrical log-domain demons algorithm with  $K = 4$  as an example.

Here, our aim is to fit the warped  $l$ th landmark point in image  $I_a$ , or  $\exp(\mathbf{v})(\mathbf{x}_a^{(l)})$ , to the corresponding landmark point  $\mathbf{x}_b^{(l)}$  in image  $I_b$ . That is,  $\exp(\mathbf{v})(\mathbf{x}_a^{(l)}) \approx \mathbf{x}_b^{(l)}$ . Note that, from the invertibility of the deformation field,  $\exp(-\mathbf{v})(\mathbf{x}_b^{(l)}) \approx \mathbf{x}_a^{(l)}$  is satisfied at the same time. Additionally, in this study, the *trajectories* of the pair of landmarks are also fitted (Fig. 5 right). Consider a set of points at regular intervals  $\exp\left(\frac{k}{K}\mathbf{v}\right)(\mathbf{x}_a^{(l)})$ ,  $k = 0, 1, 2, 3, \dots, K$  along the trajectories of  $\mathbf{x}_a^{(l)}$ . Also consider another set of points at regular intervals  $\exp\left(-\frac{k}{K}\mathbf{v}\right)(\mathbf{x}_b^{(l)})$ ,  $k = 0, 1, 2, 3, \dots, K$  along the trajectories of  $\mathbf{x}_b^{(l)}$ . Then, as illustrated in the right of Fig. 5, the corresponding pairs of points  $\exp\left(\frac{k}{K}\mathbf{v}\right)(\mathbf{x}_a^{(l)})$  and  $\exp\left(-\frac{K-k}{K}\mathbf{v}\right)(\mathbf{x}_b^{(l)})$ ,  $k = 0, 1, 2, 3, \dots, K$ , are also fitted to the same position.

Then  $\mathbf{u}_{LM}$ , the update vector field (i.e. demons forces) from landmark position information, is calculated as follows:

$$\mathbf{u}'_{LM}(\mathbf{x}) = \begin{cases} \exp\left(-\frac{K-k}{K}\mathbf{v}\right)(\mathbf{x}_b^{(l)}) - \exp\left(\frac{k}{K}\mathbf{v}\right)(\mathbf{x}_a^{(l)}) & \text{if } \mathbf{x} = \exp\left(-\frac{K-k}{K}\mathbf{v}\right)(\mathbf{x}_b^{(l)}) \\ \text{or } \mathbf{x} = \exp\left(\frac{k}{K}\mathbf{v}\right)(\mathbf{x}_a^{(l)}), \exists k \exists l, & \\ \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} & \text{otherwise} \end{cases} \quad (9)$$

$$\mathbf{u}_{LM}(\mathbf{x}) = \begin{cases} \mathbf{u}'_{LM}(\mathbf{x}) & \text{if } |\mathbf{u}'_{LM}(\mathbf{x})| \leq u_{max} \\ \frac{u_{max}}{|\mathbf{u}'_{LM}(\mathbf{x})|} \cdot \mathbf{u}'_{LM}(\mathbf{x}) & \text{if } u_{max} < |\mathbf{u}'_{LM}(\mathbf{x})| \end{cases}$$

where  $u_{max}$  is a parameter used to prevent update vectors from having a too large norm. In this study  $u_{max} = 5$  was used. Using this landmark-based update field, the landmark-guided demons algorithms is composed as follows:

**Algorithm 2. Landmark-guided symmetric log-domain demons algorithm**

- Input: two images  $I_a$  and  $I_b$  and landmark positions  $\mathbf{x}_a^{(l)}$  and  $\mathbf{x}_b^{(l)}$ ,  $l = 1, 2, 3, \dots, L$
- Set the initial speed image  $\mathbf{v}(\mathbf{p}) \leftarrow \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \forall \mathbf{p}$
- Iterate:
  - Calculate forward and backward deformation fields  $\exp\frac{1}{2}\mathbf{v}$  and  $\exp-\frac{1}{2}\mathbf{v}$

- Calculate deformed images  $I_a^d = I_a \circ \left( \exp \frac{1}{2} \mathbf{v} \right)$  and  $I_b^d = I_b \circ \left( \exp -\frac{1}{2} \mathbf{v} \right)$
- Compute the grayscale demons forces  $\mathbf{u}_{GS}(\mathbf{p}) = -\frac{\Delta I}{\|\mathbf{J}_p\|^2 + (\Delta I)^2} \cdot \mathbf{J}_p$ , where  $\Delta I = I_a^d(\mathbf{p}) - I_b^d(\mathbf{p})$  and  $\mathbf{J}_p = -\frac{1}{2} \left( \nabla I_a^d(\mathbf{p}) + \nabla I_b^d(\mathbf{p}) \right)$
- For fluid-like regularization let  $\mathbf{u}_{GS} \leftarrow K_{fluid;GS} \star \mathbf{u}_{GS}$
- Compute the landmark demons forces  $\mathbf{u}_{LM}(\mathbf{x})$  using Eq. (9)
- For fluid-like regularization let  $\mathbf{u}_{LM} \leftarrow K_{fluid;LM} \star \mathbf{u}_{LM}$
- For diffusion-like regularization let  $\mathbf{v} \leftarrow K_{diff} \star (\mathbf{v} + \mathbf{u}_{GS} + \alpha_{LM} \cdot \mathbf{u}_{LM})$

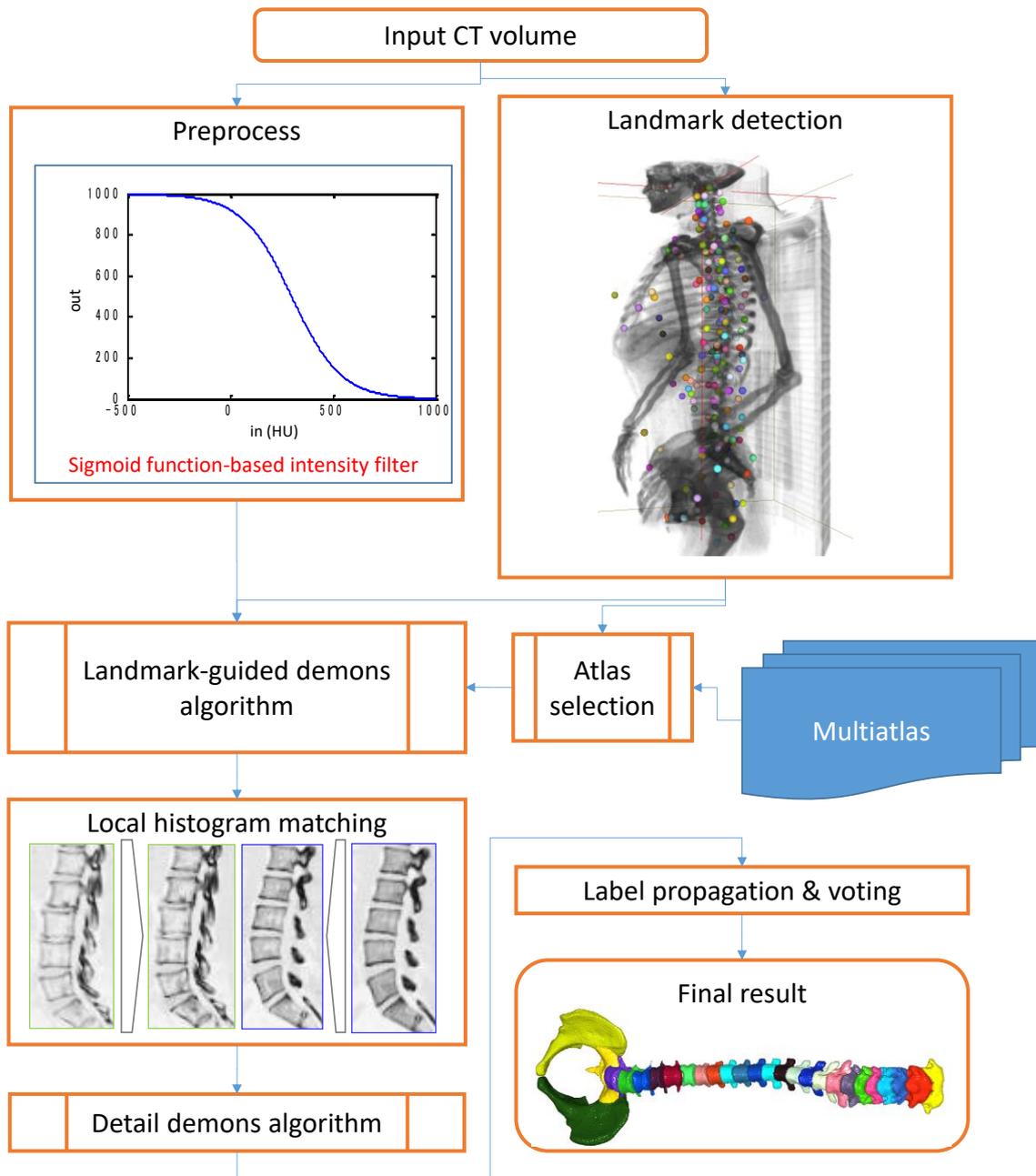
Note that although the landmark update field  $\mathbf{u}_{LM}(\mathbf{x})$  is only non-zero at discrete positions, its effect is diffused to its neighborhoods by convolving  $K_{fluid;LM}$  and  $K_{diff}$ . The parameters of the Gaussian filtering  $\sigma_{fluid;LM}, \sigma_{fluid;GS}, \sigma_{diff}$  and the weight coefficient  $\alpha_{LM}$  are hyperparameters to be determined in advance.

#### 4.2.2 Multiatlas segmentation framework

On the basis of the landmark-guided demons algorithm described above, I developed a novel framework to segment the whole spine and the pelvic bones. The framework has several advantageous features. Firstly, cranial or caudal shifts in the spinal registration process are effectively suppressed by using the preceding landmark detection result. Secondly, an extremely fast atlas selection method is utilized to reduce the computational cost of the following series of registrations. Thirdly, a local histogram matching method is also

applied to compensate for a wide variety of local bone mineral densities among datasets.

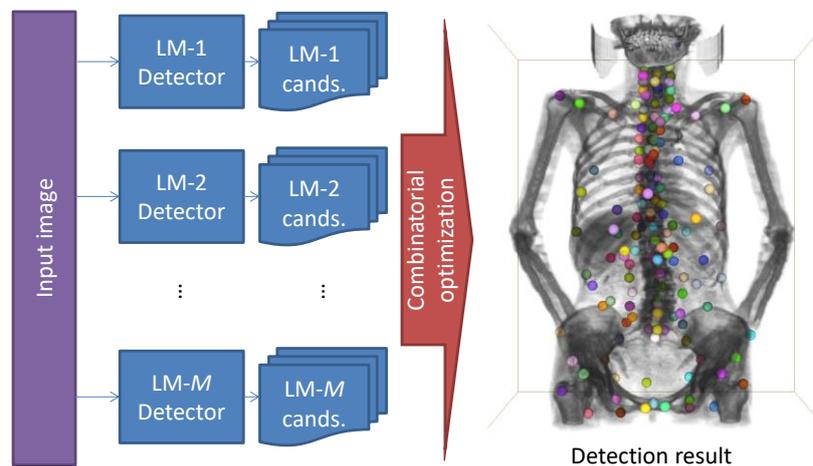
Figure 6 shows a flowchart of the proposed segmentation framework. Firstly, total a 170 of bony landmarks are detected by my landmark detection system. Before registration, as a preprocess, the voxel values of images are modified by a sigmoid function to emphasize bony structure. Additionally, a feature vector is extracted from voxels near each landmark to perform atlas selection. After atlas selection, the selected atlases are registered to the given unseen target image using the proposed landmark-guided demons algorithm. Then, local histogram matching and additional fine registration are performed. Finally, label propagation is performed by voxelwise voting and the final segmentation result is given.



**Fig. 6.** Outline of the proposed method.

#### 4.2.2.1 Landmark detection system

Figure 7 shows an outline of my landmark detection system. Details of the system are given in [18,19]. In brief, the detection system has two steps. Firstly, each landmark is detected independently and in parallel by the corresponding landmark-dedicated detector. Each landmark detector outputs multiple candidate positions for its target landmark. Then, in the following final combinatorial optimization step, one landmark point set is determined for all landmarks using an interlandmark-distance-based spatial landmark distribution model. As the result, the most probable positions of 170 bony landmarks are outputted. All the landmarks are labeled (for example, as the spinal process of the 11th thoracic vertebra or as the right transverse foramen of the 5th cervical vertebra) by the detection system. The landmark detection system can handle CT volumes with various imaging ranges and field of view (FOV) sizes.



**Fig. 7.** Outline of the landmark detection system.

#### 4.2.2.2 Preprocessing

As a preprocess, every voxel value in the input CT volume is processed by a sigmoid function to emphasize the bony structure as follows:

$$I_{out} = \frac{1000}{1 + \exp\left(\frac{I_{in} - b}{a}\right)}. \quad (10)$$

In this study,  $a = 120$  and  $b = 300$  are used. These values were determined experimentally. Additionally, all volumes are rescaled so that the voxel size becomes 2 mm/voxel.

#### 4.2.2.3 Atlas selection

In the proposed method, atlas selection is performed using global image information before registration instead of local image information after registration. Both approaches have advantages and disadvantages; the latter may provide more precise segmentation results at the cost of heavier computational burden. In this study I perform global image information before registration because I have relatively large ( $\sim 40$ ) atlas datasets. Although my demons-based registration algorithm is reasonably fast, registering all the atlases is too costly in many practical applications. Global atlas selection can greatly reduce the computational cost (as described later).

In the atlas selection, the information on landmark positions is optimally used to extract feature values of each CT volume. Eight cubic regions are placed around each landmark, and the average CT values of these eight cubes are used as features. All the cubes have the same landmark point as one of their vertices. Each cube has a side of 20 mm. A total of  $8 \times 170 = 1360$  features are calculated for each CT volume and are used to compose the feature vector. Using this vector, the similarity between a given pair of CT volumes

is estimated using the simple squared Euclidean distance between the two vectors.

In actual atlas selection, the similarity between each atlas and the given unseen volume is calculated. Then, the atlases with the largest similarities are selected and used in the following multiatlas segmentation process. The number of atlases used  $n_{atlas}$  is a hyperparameter to be determined in advance, and is used to balance the computational cost and the segmentation accuracy.

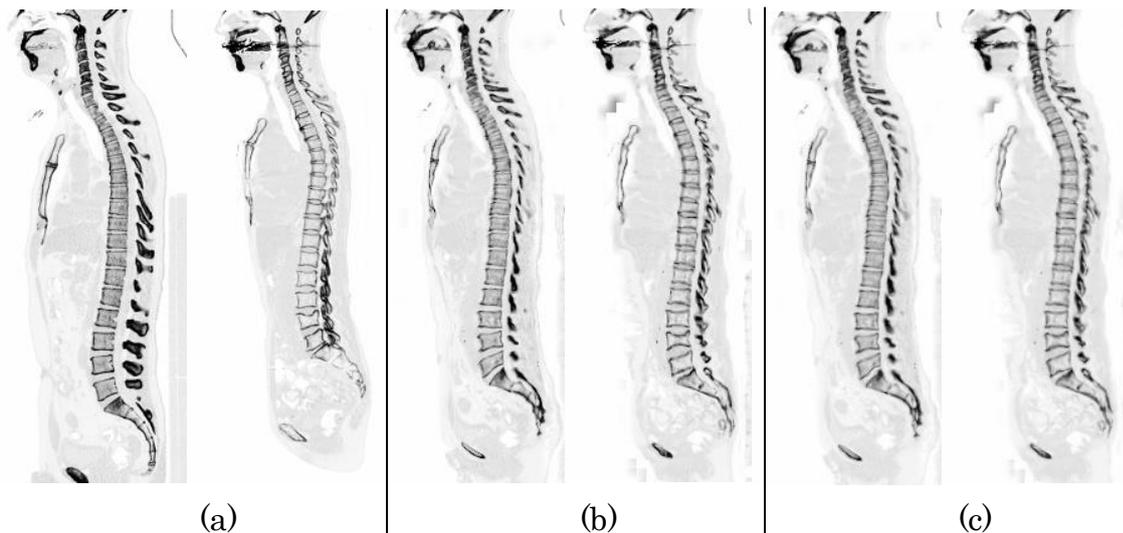
#### 4.2.2.4 Landmark-guided demons

After atlas selection, each atlas is registered to the given unseen volume. Firstly, the landmark-guided demons algorithm (described in 4.2.1) is performed using landmark information. All landmark positions in the atlas datasets are manually inputted by a radiologist in advance. Landmark positions in the unseen volume are determined by my automatic landmark detection system (as described in 4.2.2.1). Using two pairs comprising the landmark information and the grayscale volume, the landmark-guided demons algorithms is performed to register two bony structures.

I also utilize a multiresolutional approach. The CT volumes are resized with ratios of  $2^{-n}$ ,  $n = 3, 2, 1, 0$ . For each resolution, 20 cycles of the landmark-guided demons algorithm are performed. When the resolution is changed, the CT volume and speed image are rescaled. Furthermore, all the dislocation vectors (e.g.,  $\exp \frac{1}{2} \mathbf{v}$ ) must be scaled appropriately (i.e., multiplied by  $2^{-n}$ ) before every image warping. Note that the speed image  $\mathbf{v}$  cannot be multiplied simply; halving the speed of the flow does not result in halving the spatial scale.

#### 4.2.2.5 Local histogram matching

In practice, demons-based registration of the bony structure is affected by the variation of the bone mineral density. Osteoporosis, from which many old women suffer, causes very low CT values in the bony structure, particularly in vertebrae. Such a difference in CT values can cause large registration errors. To avoid such errors, a local histogram matching is performed on each local part of the given volumes, which is followed by an additional, precise demons registration. (Fig.8)



**Fig. 8.** Example of registration. (a) Original images after preprocessing. (b) Result of landmark-guided demons algorithm and histogram matching. (c) After performing detailed demons algorithm.

After the landmark-guided demons registration, the volumes are divided into cubes with a side length of 16 mm. The histogram of the voxel intensities of each cube in the two volumes is calculated. Then all the voxel values in the

same cube are matched between the two volumes, so that the histograms of the two cubes become identical. Let  $q_A(c)$  be the  $c$ th percentile of the image intensities of the given cube A. For example,  $q_A(50)$  is the median of the intensities of cube A. Suppose that the  $c$ th percentile of the intensities in the corresponding cube B in the other image is  $q_B(c)$ . Then, the intensity value  $z_A$  in cube A will be converted to a new value  $z_{A,\text{matched}}$  as follows:

$$q_{\text{mean}}(c) = \frac{q_A(c) + q_B(c)}{2}, \quad (11)$$

$$z_{A,\text{matched}} = q_{\text{mean}}(q_A^{-1}(z_A)),$$

where  $q_A^{-1}(\cdot)$  is the inverse function of  $q_A(\cdot)$ . Additionally, in implementation, the frequencies of histograms are spatially interpolated between cubes to avoid noncontinuity of the resulting volume.

#### 4.2.2.6 Detailed demons algorithm

After local histogram matching, precise registration is performed. In this phase, landmark information is not used to avoid the effects of small landmark detection errors. The precise registration is performed by the log-domain demons algorithm (as described in 2.1.1) without a multiresolutional approach. A total of 20 cycles of the demons algorithm are performed.

#### 4.2.2.7 Label propagation and voting

After all the atlases are registered to the unseen image, label propagation and voting are performed to compose the final segmentation result. Each atlas has a manually inputted label volume, which includes all 24 vertebral bones as well as the bilateral iliac bones and the sacrum. In the label propagation, each label

volume is deformed using the registration result so that it fits to the given unseen image. Then, voxelwise voting is performed using  $n_{atlas}$  registered labels and the majority label is chosen voxel by voxel.

### 4.2.3 Parameter optimization and evaluation

This study was approved by our institutional ethics review board. For this type of retrospective study, formal informed consent is not required.

I used a total of 50 whole-torso CT datasets without the administration of intravenous contrast agent. All subjects had no bone diseases other than osteopenia. The voxel size was  $0.977 \times 0.977 \times 1.250$  mm. Subjects with abnormal number of vertebrae was excluded from this study beforehand. For all the datasets, the answer label volumes and landmark positions were manually inputted. Among the 170 landmarks, 120 were spinal, 39 were pelvic and 11 were on other bony structures. Answer labels were inputted for 27 bones (the 24 vertebrae, the bilateral iliac bones and the sacrum). Note that landmark answers were used only for the atlases and were not used in the test phase (the results of automatic detection were used instead).

My method has many hyperparameters: the parameters of Gaussian filtering  $\sigma_{fluid;LM}, \sigma_{fluid;GS}, \sigma_{diff}$ , the weight coefficient  $\alpha_{LM}$ , the number of resolutions, the parameters of the sigmoid function, and so forth. I used 10 out of the 50 datasets for parameter optimization. The best combination of parameters was searched for by a grid search. A leave-one-out method was used, that is, all nine datasets other than the target case were used as atlases (without atlas selection). The criterion used was the Dice coefficient. In the result,  $\sigma_{fluid;LM} = 4.0, \sigma_{fluid;GS} = 1.0, \sigma_{diff} = 0.5, \alpha_{LM} = 1.0$  were

selected for the landmark-guided demons algorithm and  $\sigma_{fluid} = 1.0, \sigma_{diff} = 0.4$  were selected for the detailed demons algorithm.

For comparison, I also implemented another demons algorithm with landmark guidance but without diffeomorphism. It was simply performed by setting  $N = 1$  in Eqs. (7) and (8). In other words, the rough approximation  $\exp\frac{1}{2}\mathbf{v} \cong Id + \frac{1}{2}\mathbf{v}$  was used.

The experiment was performed using the other 40 datasets. A leave-one-out method was also used in the experiment. Variable numbers of atlases  $n_{atlas}$  were used to evaluate the effects of  $n_{atlas}$ . The segmentation result was evaluated by three criteria: the Dice coefficient, the (voxel-to-voxel) mean distance error and the Hausdorff distance error. The definitions of these criteria are as follows:

$$\text{Dice coefficient} = \frac{2|A \cap B|}{|A| + |B|}$$

Hausdorff distance

$$= \max \left\{ \max_{\mathbf{a} \in \partial A} \min_{\mathbf{b} \in \partial B} |\mathbf{a} - \mathbf{b}|, \max_{\mathbf{b} \in \partial B} \min_{\mathbf{a} \in \partial A} |\mathbf{a} - \mathbf{b}| \right\},$$

mean distance

$$= \frac{1}{2} \left\{ \frac{1}{|\partial A|} \sum_{\mathbf{a} \in \partial A} \min_{\mathbf{b} \in \partial B} |\mathbf{a} - \mathbf{b}| + \frac{1}{|\partial B|} \sum_{\mathbf{b} \in \partial B} \min_{\mathbf{a} \in \partial A} |\mathbf{a} - \mathbf{b}| \right\},$$

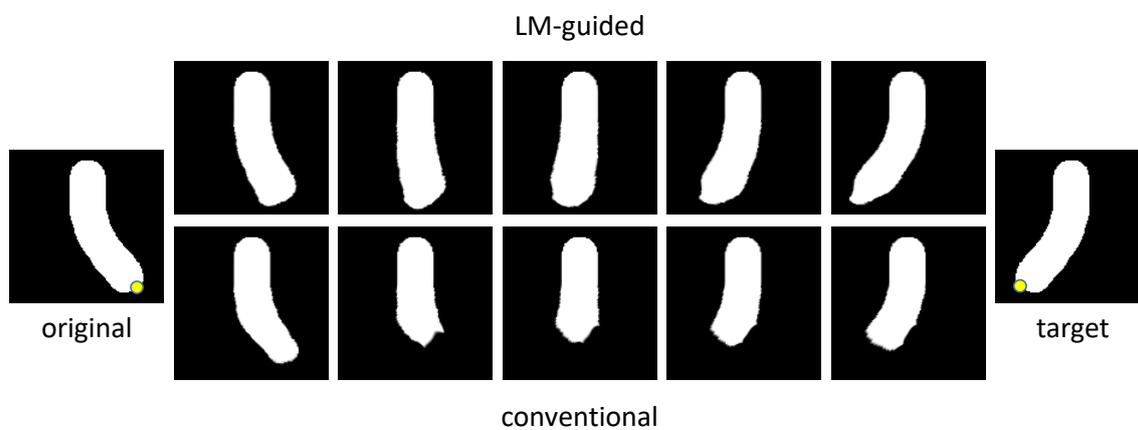
where  $A$  and  $B$  are the voxel sets of the answer and the computed label regions, respectively.  $\partial A$  denotes the border voxels of  $A$ .

(12)

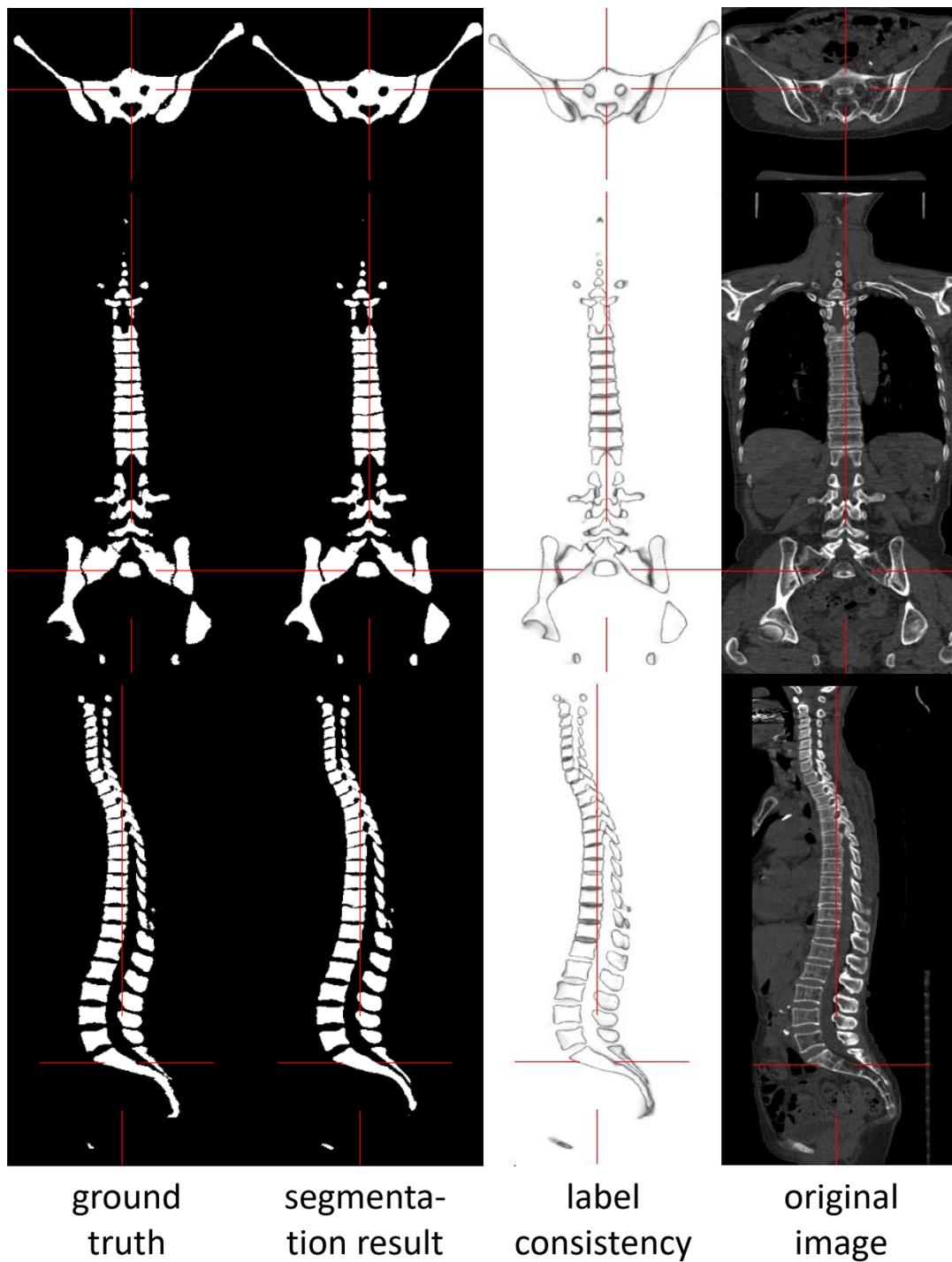
I also used the publicly available SpineWeb dataset (<http://spineweb.digitalimaginggroup.ca/>) [22] to evaluate the segmentation accuracy. The SpineWeb dataset include 20 thoracolumbar spinal small-field-of-view CT volumes, 10 for training (Dataset 2) and 10 for testing (Dataset 15). In the latter, five out of the 10 subjects have osteoporosis and compression fractures. In my experiment, an alternative hyperparameter tuning was performed in advance by a grid search using 10 training volumes (Dataset 2). The selected parameters were  $\sigma_{fluid;LM} = 5.0$ ,  $\sigma_{fluid;GS} = 1.5$ ,  $\sigma_{diff} = 1.0$ ,  $\alpha_{LM} = 1.0$  for the landmark-guided demons algorithm and  $\sigma_{fluid} = 0.5$ ,  $\sigma_{diff} = 0.4$  for the detailed demons algorithm. After hyperparameter optimization, the segmentation accuracies of thoracolumbar vertebrae were evaluated for each of the five healthy and five diseased spines in the test dataset.

### 3. Result

First, I tested the proposed method using a toy model. Figure 9 shows the registration results for the toy model comprised of a curved rod. Without landmark information, the curved rod shape firstly shrinks, and then elongates to fit the target image. On the other hand, using the proposed method, the shape of the rod is preserved during the deformation.



**Fig. 9.** Results for toy model using of the proposed and conventional demon methods. The yellow circle represents a landmark used in the proposed method.



**Fig. 10.** Example of segmentation result for the spine and pelvis.

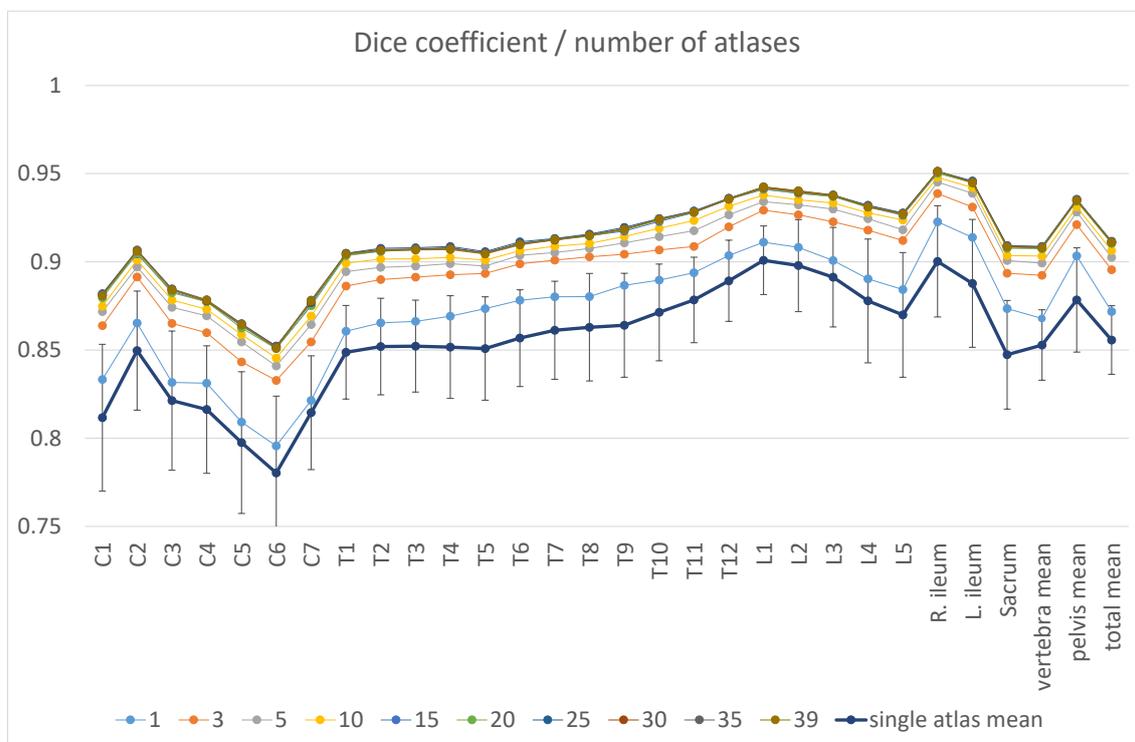


**Fig. 11.** Example of segmentation result for the spine and pelvis.

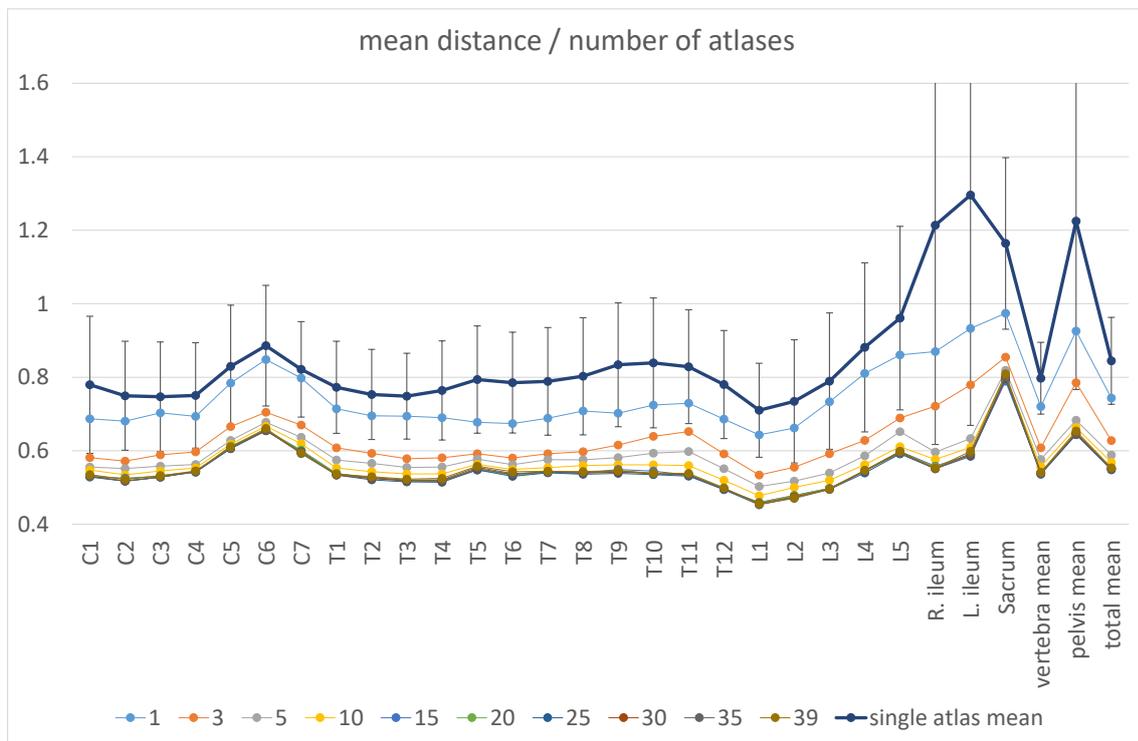
Figures 10 and 11 illustrate an example of the segmentation result for the proposed segmentation method. Note that all the vertebral processes and sacral foramina are correctly segmented. I consider that this correctness is mainly due to the landmark-preserving nature of the proposed method.

Figures 12-15 show the Dice coefficients, mean distances and Hausdorff distances of the proposed segmentation method. As shown, the segmentation accuracy improves as the number of atlases increases. However, this increase is almost saturated when the number of atlases reaches 10. Furthermore, the segmentation results with only one atlas (selected by the atlas selection algorithm) are significantly better than the average accuracies with a single atlas without atlas selection. Figure 16 also shows that the segmentation result with the proposed atlas selection surpassed the result with randomly selected atlases in all 40 cases (a total of 50 random tests were performed). Therefore, I concluded that my atlas selection method works well and is effective for reducing the number of required atlases. Even when only five atlases were used, a mean Dice coefficient of  $0.90 \pm 0.02$ , a mean distance error of  $0.59 \pm 0.14$  mm and a Hausdorff distance of  $5.30 \pm 2.14$  mm were achieved. For vertebral region only, a Dice coefficient of  $0.90 \pm 0.03$ , a mean error of  $0.59 \pm 0.14$  mm and a Hausdorff distance of  $4.93 \pm 2.01$  mm were achieved. Therefore, considering the computational cost, I consider that five atlases is sufficient for most applications. It took approximately 15 min for one segmentation task using five atlases, whereas it took approximately 110 min when 39 atlases were used. Therefore, the computational cost was

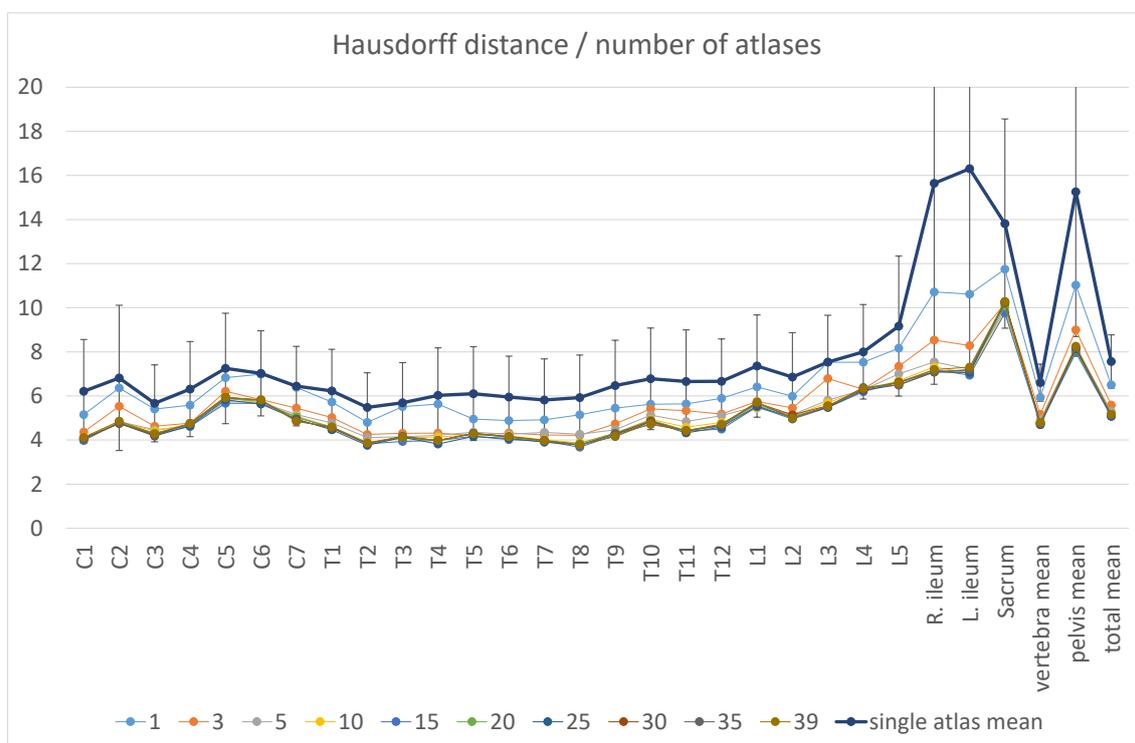
approximately proportional to the number of atlases used.



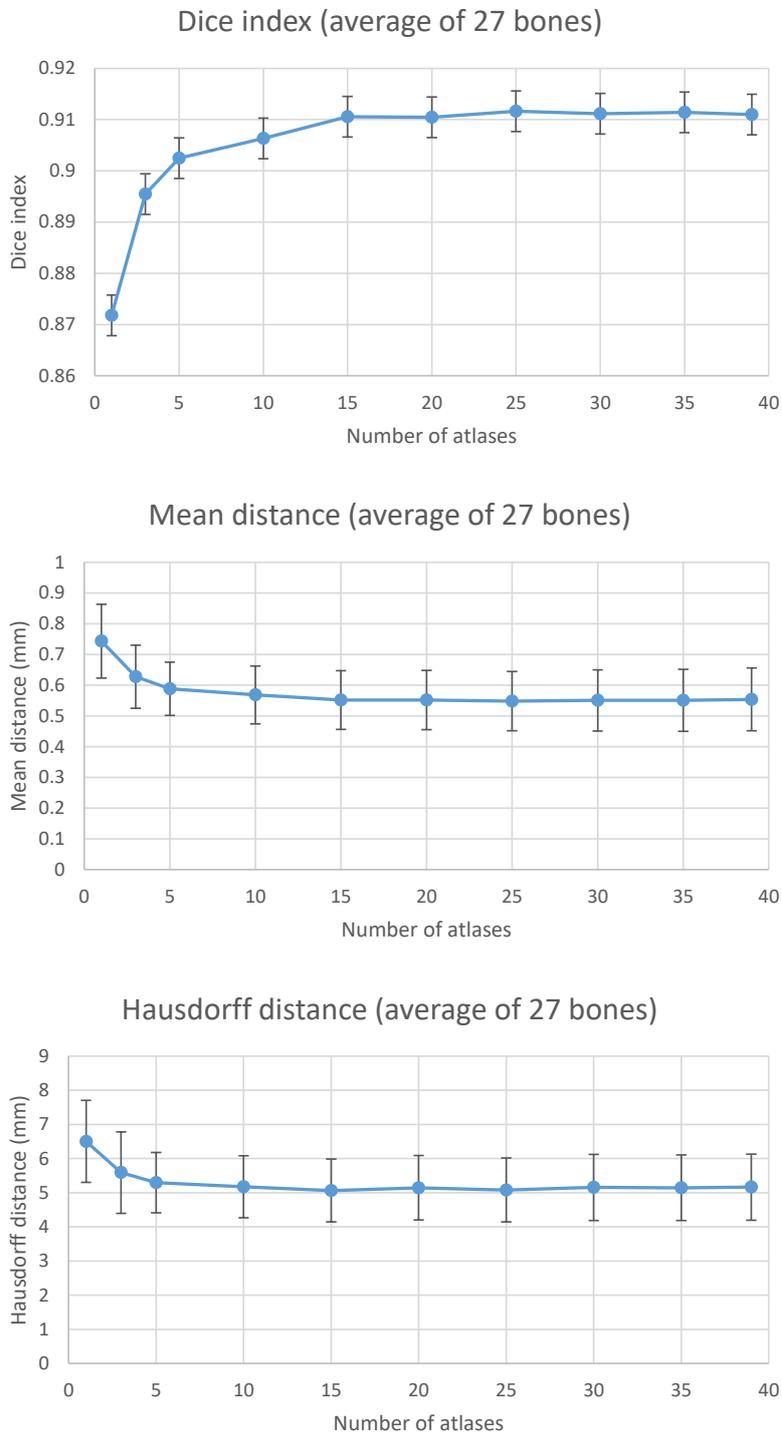
**Fig. 12.** Dice coefficients of the multiatlas and single-atlas segmentation results. Standard deviations of all single-atlas segmentation results (calculated from a total of  $40 \times 39 = 1760$  segmentations) are also shown.



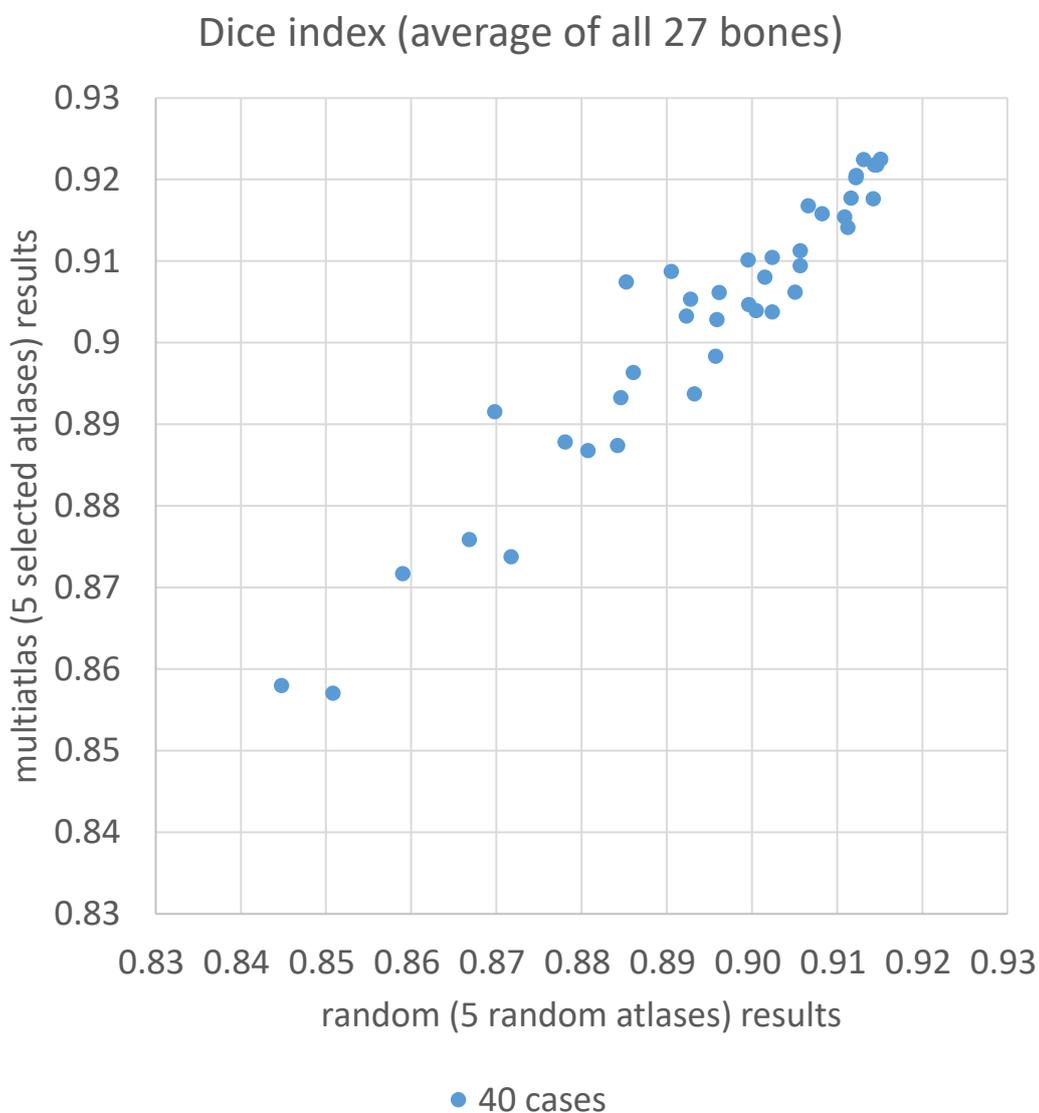
**Fig. 13.** Mean distances of the multiatlas and single-atlas segmentation results.



**Fig. 14.** Hausdorff (maximal) distances of the multiatlas and single-atlas segmentation results.

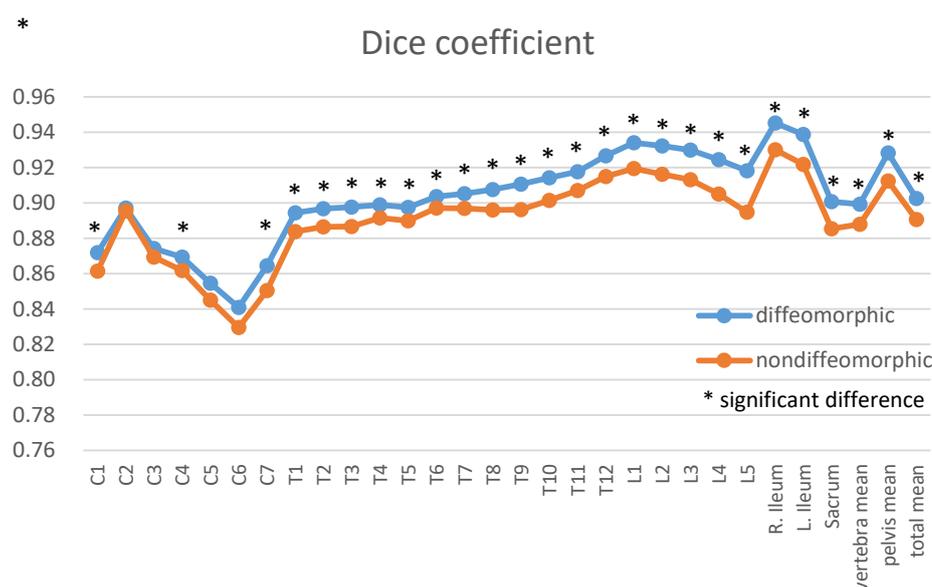


**Fig. 15.** Dice index, mean distance and Hausdorff distance for several numbers of atlases.



**Fig. 16.** Scatter plot of the segmentation accuracies of the proposed atlas selection vs random atlas selection is shown. The accuracy criterion is the average of the Dice coefficients among all 27 bone structures. Each dot represents one test case.

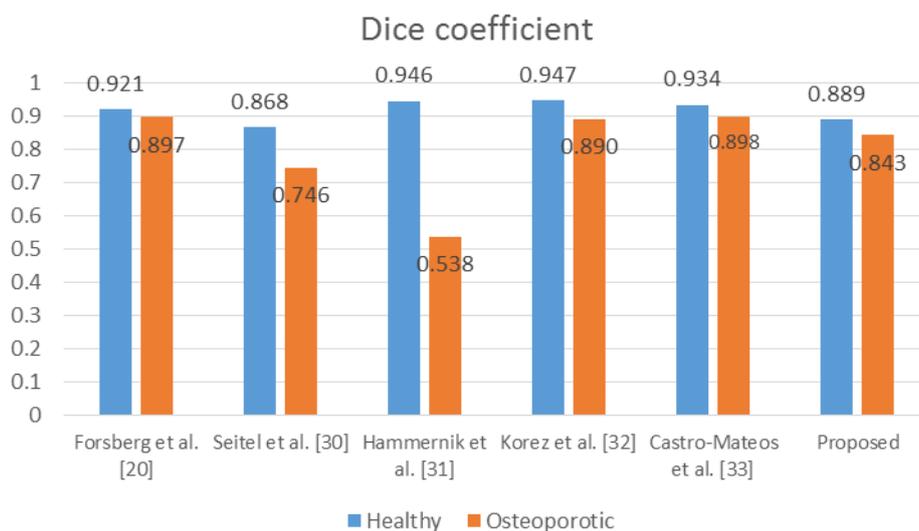
Figure 17 shows the Dice coefficients of the segmentation results obtained with the diffeomorphic demons and simulated non-diffeomorphic demons algorithms when five atlases were used. As shown, the segmentation accuracies were significantly better with the diffeomorphic demons algorithm for most of the vertebral and pelvic bones.



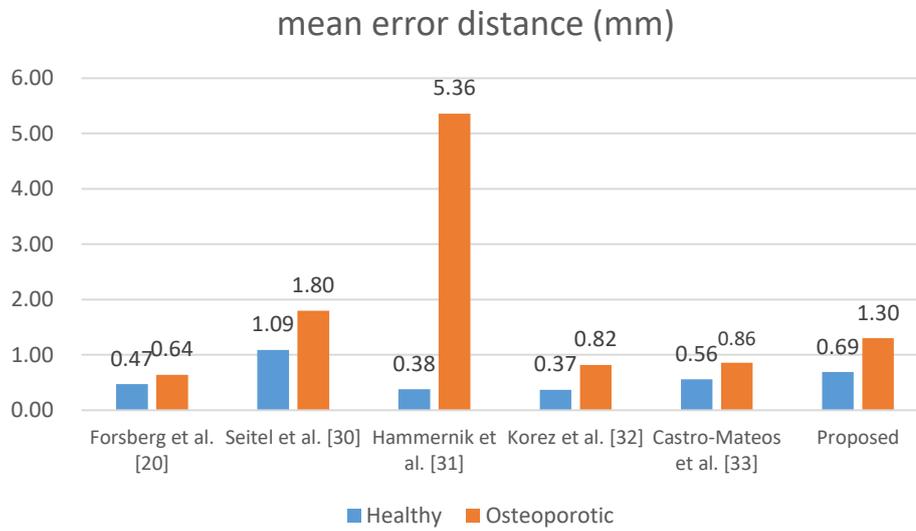
**Fig. 17.** Dice coefficients for diffeomorphic vs non-diffeomorphic demons algorithms (with five atlases). Asterisks indicate a significant difference ( $p < 0.05$ ) between the two algorithms (by paired t-test).

I also used the SpineWeb dataset [22] to evaluate the segmentation accuracy. Because the volumes in the dataset only include thoracolumbar spines, the Dice indices and mean error distances for only thoracic and lumbar vertebrae were evaluated. Figures 18 and 19 show my segmentation accuracies for healthy cases and diseased cases with compressed fractures, respectively, compared with the results of other state-of-the-art methods [20] [30-33]. Figure 20

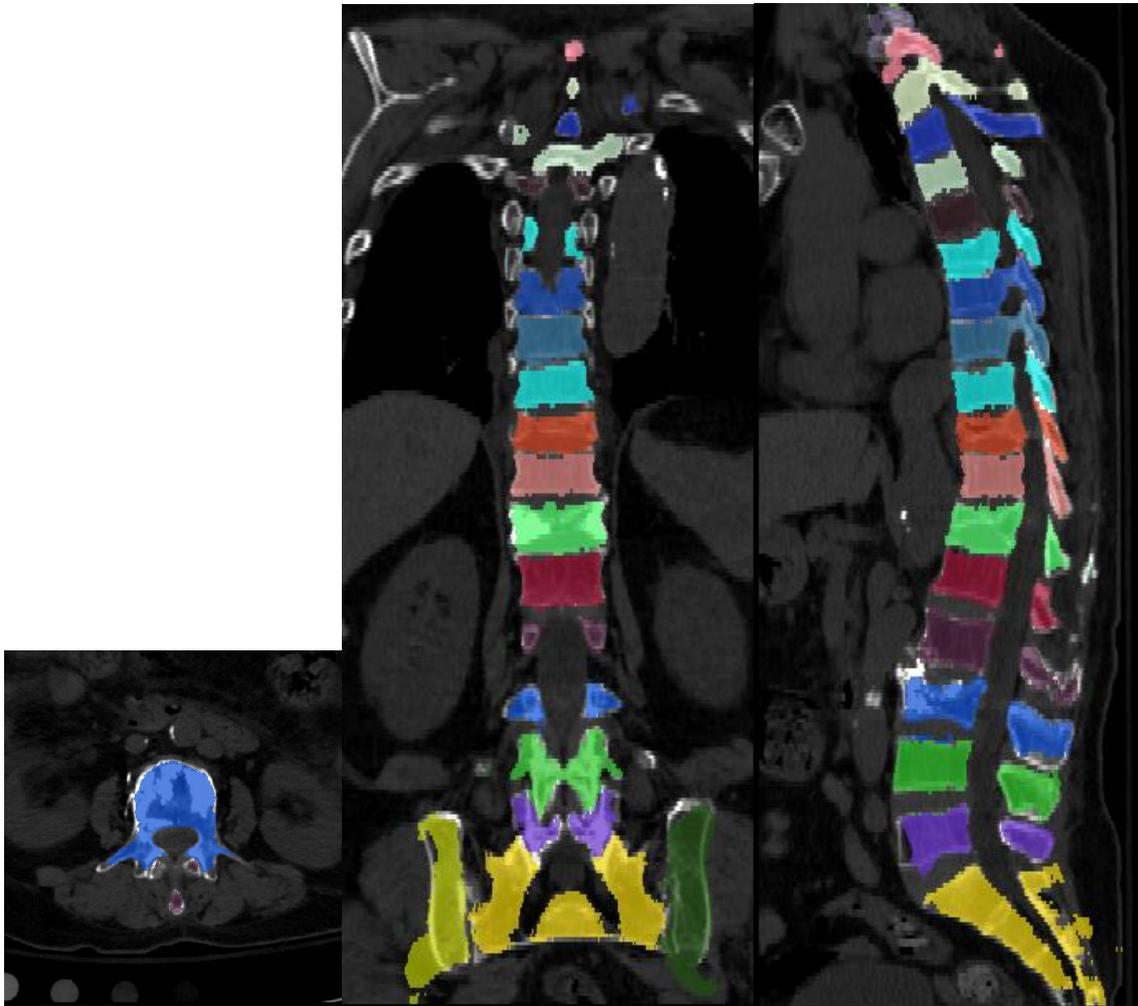
illustrates an example of a segmentation result (a case with multiple compression fractures). As shown, my method resulted in a mean Dice index of 0.889 and a mean error distance of 0.69 mm for the healthy subjects. On the other hand, the performance was slightly degraded for diseased subjects to a mean Dice index of 0.843 and a mean error distance of 1.30 mm. My method showed comparable accuracy to the other state-of-the-art methods.



**Fig. 18.** Mean Dice indices of the SpineWeb dataset and results of other state-of-the-art methods.



**Fig. 19.** Mean distance errors of the SpineWeb dataset and results of other state-of-the-art methods.



**Fig. 20.** Segmentation result of a SpineWeb subject with multiple compression fractures.

## 4. Discussion

A novel method for segmenting the whole spine and the pelvis was presented. In the evaluation, a mean distance error of  $0.59 \pm 0.14$  mm was achieved for the whole spine and the pelvic bones. Comparing my results with those in previous studies, Klinder et al. reported a mean distance error of  $1.12 \pm 1.04$  mm [6] for healthy or diseased whole-spine images using their method.

Recently, Korez et al. [9] reported a method for thoracolumbar vertebrae whose mean distance error was 0.3 mm and Dice coefficient was 0.946. In another method for thoracolumbar vertebrae reported by Castro-Mateos et al. [8], the mean distance error was 0.58 mm. Thus, my method showed comparable accuracy to these state-of-the-art methods. Additionally, a direct comparison using the SpineWeb dataset was performed for thoracolumbar vertebrae. In the comparison, the proposed method also showed comparable performance for both healthy and diseased thoracolumbar spines. Therefore, I believe that the accuracy and stability of my method have been confirmed.

The proposed method used a novel landmark-guided log-domain demons algorithm for registration. One of the advantages of this algorithm is the diffeomorphism/invertibility of the deformation field. The invertibility is required if it is necessary to warp both image(s) and landmark(s). The deformation field for warping images is not the same as, and is the inverted version of, the field for warping landmarks. Additionally, owing to the log-domain mechanism and the explicit speed vector field  $\mathbf{v}$ , not only the position of each landmark but also its trajectory can be traced and fitted. In my experience, this mechanism greatly improves the stability of my landmark-guided demons registration, especially when the landmark positions in the two volumes are distant in the initial setting. I assume that this trajectory-fitting strategy is effective in my voxelwise registration algorithm with an extremely high number of DOFs. I believe that, owing to both the high number of DOFs and the stable landmark fitting, the accurate registration of complex shapes of vertebrae around landmarks (e.g., vertebral processes and holes) was achieved. My method showed comparable, but not superior, segmentation performance to that of Forsberg et al. [20] for thoracolumbar vertebrae. In fact, their

approach and mine are similar in many ways. Both of them utilize a multiatlas-based approach. Their algorithm estimates the positions and poses of all vertebrae prior to the main multiatlas segmentation in order to determine the initial condition of the registration problems, which corresponds to my landmark detection phase. In the registration, their morphon-based method calculates a dense displacement field at every cycle and uses a multiresolutional approach, in accordance with [34]. However, the deformation field of their registration method does not have diffeomorphism. Although the benefit of diffeomorphism has been shown to be significant in my study, it is possible that diffeomorphism itself is not an indispensable property in registering the spinal structure.

My multiatlas-based segmentation method has several advantages and disadvantages compared with the other SSM-based methods. Firstly, it can simultaneously segment not only the spine but also the pelvis with little additional computational cost, since in my current implementation whole bodies are always registered. It is possible that, if correctly initialized using additional landmarks, my method can segment other body trunk bones such as the ribs and sternum. On the other hand, in the SSM-based approach, individual SSMs will be needed to segment these structures. Secondly, it can accurately segment fine protuberances or holelike structures such as vertebral processes and transverse foramina particularly in the cervical region, owing to high number of DOFs of the demons algorithm. For example, Klinder et al. [6] reported mean distance errors for the segmentation of cervical vertebrae of 0.81 to 1.13 mm, whereas my mean distance errors for cervical vertebrae ranged from 0.55 to 0.67 mm. Thirdly, because my method segments the spinal-pelvic structure as a whole, it is free from “model instance collision,”

that is, the overlap of segmented areas between two adjacent bones. This is a major problem that is due to the complicated shapes of intervertebral and sacroiliac joints. Many SSM-based methods have a mechanism dedicated to avoiding such a collision problem [6] [8]. On the other hand, my multiatlas-based method is essentially collision-free.

Among the disadvantages of my method compared with SSM-based methods, seemingly the most problematic one is the high computational cost. Although I attempted to reduce the computational cost by utilizing preregistration atlas selection, it took about 15 min for one subject excluding landmark detection, which took another 17 min. Further reduction of the computational time will be one of my future challenges. Secondly, my method relies on predetected landmark position information and is strongly affected by the landmark detection accuracy. However, this problem is not specific to my method because most other SSM-based methods require a preceding vertebral identification method that can be regarded as vertebral “landmark” detection method. For example, Korez et al. [9] used an interpolation-based vertebra identification method, whereas Klinder et al. [6] used another appearance-model-based identification method. And my vertebra identification accuracy [28] is comparable to those of other state-of-the-art vertebral body/disc identification methods and I was able to correctly detect most of the vertebrae in my 40 test cases.

This study has some limitations. My test dataset only includes healthy spines or those with osteoporosis. Spines with other diseases such as scoliosis, lordosis, postsurgical changes or bone metastasis were not included. Another problem is that spines with abnormal numbers of vertebrae were excluded from

my dataset. Such anatomical anomalies are quite common, for example, an occurrence of 10.8% was reported in [35]. I am now attempting to identify abnormal numbers of vertebrae by virtually converting the landmark configurations of anomalous spines to those of normal ones [36]. Note that, once the identification of abnormal vertebrae has been achieved, all vertebrae can be segmented by the same methodology using “abnormal atlas” datasets that consist of cases with a specific abnormality. Because my segmentation results are reasonable even when only one atlas is used, I expect that the use of only one atlas for each abnormality will be sufficient to correctly segment abnormal spine.

## 5. Conclusion

A method for segmentation of the spine and pelvis was presented. The experimental results showed high accuracy for cervical, thoracic and lumbar vertebrae as well as for pelvic bones. My future works include evaluation using datasets of spines with diseases, addressing the problem of abnormal numbers of vertebrae, and the simultaneous segmentation of other bony structures such as the rib cage and sternum.

## References

1. Yao J, O'Connor SD, Summers R (2006) Computer aided lytic bone metastasis detection using regular CT images. In: Medical Imaging, 2006. International Society for Optics and Photonics, pp 614459-614459-614459
2. Létourneau D, Kaus M, Wong R, Vloet A, Fitzpatrick DA, Gospodarowicz M, Jaffray DA (2008) Semiautomatic vertebrae visualization, detection, and identification for online palliative radiotherapy of bone metastases of the spine). *Medical Physics* 35 (1):367-376
3. Ferrari V, Parchi P, Condino S, Carbone M, Baluganti A, Ferrari M, Mosca F, Lisanti M (2013) An optimal design for patient - specific templates for pedicle spine screws placement. *The International Journal of Medical Robotics and Computer Assisted Surgery* 9 (3):298-304
4. Pereañez M, Lekadir K, Hoogendoorn C, Castro-Mateos I, Frangi A (2015) Detailed vertebral segmentation using part-based decomposition and conditional shape models. In: *Recent Advances in Computational Methods and Clinical Applications for Spine Imaging*. Springer, pp 95-103
5. Kadoury S, Labelle H, Paragios N (2013) Spine segmentation in medical images using manifold embeddings and higher-order MRFs. *IEEE Transactions on Medical Imaging* 32 (7):1227-1238
6. Klinder T, Ostermann J, Ehm M, Franz A, Kneser R, Lorenz C (2009) Automated model-based vertebra detection, identification, and segmentation in CT images. *Medical Image Analysis* 13 (3):471-482
7. Rasouljan A, Rohling R, Abolmaesumi P (2013) Lumbar spine segmentation using a statistical multi-vertebrae anatomical shape+pose model. *IEEE*

Transactions on Medical Imaging 32 (10):1890-1900

8. Castro-Mateos I, Pozo JM, Pereanez M, Lekadir K, Lazary A, Frangi AF (2015) Statistical interspace models (SIMs): application to robust 3D spine segmentation. *IEEE Transactions on Medical Imaging* 34 (8):1663-1675
9. Korez R, Ibragimov B, Likar B, Pernus F, Vrtovec T (2015) A framework for automated spine and vertebrae interpolation-based detection and model-based segmentation. *IEEE Transactions on Medical Imaging* 34 (8):1649-1662
10. Iglesias JE, Sabuncu MR (2015) Multi-atlas segmentation of biomedical images: A survey. *Medical Image Analysis* 24 (1):205-219
11. Aljabar P, Heckemann RA, Hammers A, Hajnal JV, Rueckert D (2009) Multi-atlas based segmentation of brain images: atlas selection and its effect on accuracy. *Neuroimage* 46 (3):726-738
12. van Rikxoort E, Arzhaeva Y, van Ginneken B (2007) Automatic segmentation of the liver in computed tomography scans with voxel classification and atlas matching. In: *Proceedings of the MICCAI Workshop, 3D Segmentation In The Clinic: A Grand Challenge*, 2007, pp. 101-108.
13. Rohlfing T, Brandt R, Menzel R, Maurer Jr CR (2004) Evaluation of atlas selection strategies for atlas-based image segmentation with application to confocal microscopy images of bee brains. *NeuroImage* 21 (4):1428-1442. doi: <http://dx.doi.org/10.1016/j.neuroimage.2003.11.010>
14. Heckemann RA, Hajnal JV, Aljabar P, Rueckert D, Hammers A (2006) Automatic anatomical brain MRI segmentation combining label propagation and decision fusion. *NeuroImage* 33 (1):115-126. doi:<http://dx.doi.org/10.1016/j.neuroimage.2006.05.061>

15. Sanroma G, Wu G, Gao Y, Shen D (2014) Learning to rank atlases for multiple-atlas segmentation. *IEEE Transactions on Medical Imaging* 33: 1939-1953.
16. McIntosh C, Purdie TG (2016) Contextual atlas regression forests: Multiple-atlas-based automated dose prediction in radiation therapy. *IEEE Transactions on Medical Imaging* 35: 1000-1012.
17. Konukoglu E, Glocker B, Zikic D, Criminisi A (2012) Neighbourhood approximation forests, In: Ayache N, Delingette H, Golland P, Mori K (Eds.), *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2012: 15th International Conference, Nice, France, October 1-5, 2012, Proceedings, Part III*. Springer, Berlin Heidelberg, pp. 75-82.
18. Thirion J-P (1998) Image matching as a diffusion process: an analogy with Maxwell's demons. *Medical image analysis* 2 (3):243-260
19. Vercauteren T, Pennec X, Perchant A, Ayache N (2008) Symmetric log-domain diffeomorphic registration: A demons-based approach. In: *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2008*. Springer, pp 754-761
20. Forsberg D (2014) Atlas-based segmentation of the thoracic and lumbar vertebrae, In: Yao J, Glocker B, Klinder T, Li S (Eds.), *Recent Advances in Computational Methods and Clinical Applications for Spine Imaging*. Springer International Publishing, Cham, pp. 215-220.
21. Wang Z, Zhen X, Tay K, Osman S, Romano W, Li S (2015) Regression segmentation for spinal images. *IEEE Transactions on Medical Imaging* 34 (8):1640-1648

22. Yao J, Burns JE, Forsberg D, Seitel A, Rasoulia A, Abolmaesumi P, Hammernik K, Urschler M, Ibragimov B, Korez R, Vrtovec T, Castro-Mateos I, Pozo JM, Frangi AF, Summers RM, Li S (2016) A multi-center milestone study of clinical vertebral CT segmentation. *Computerized Medical Imaging and Graphics* 49: 16-28.
23. Kurtek S, Srivastava A, Klassen E, Laga H (2013) Landmark-Guided Elastic Shape Analysis of Spherically-Parameterized Surfaces. *Computer Graphics Forum* 32: 429-438.
24. Kearney V, Chen S, Gu X, Chiu T, Liu H, Jiang L, Wang J, Yordy J, Nedzi L, Mao W (2015) Automated landmark-guided deformable image registration. *Physics in Medicine and Biology* 60: 101.
25. Xie Q, Kurtek S, Klassen E, Christensen GE, Srivastava A (2014) Metric-Based Pairwise and Multiple Image Registration, In: Fleet, D., Pajdla, T., Schiele, B., Tuytelaars, T. (Eds.), *Computer Vision – ECCV 2014: 13th European Conference, Zurich, Switzerland, September 6-12, 2014, Proceedings, Part II*. Springer International Publishing, Cham, pp. 236-250.
26. Lam KC, Gu X, Lui LM (2015) Landmark constrained genus-one surface Teichmüller map applied to surface registration in medical imaging. *Medical Image Analysis* 25: 45-55.
27. Nemoto M, Masutani Y, Hanaoka S, Nomura Y, Yoshikawa T, Hayashi N, Yoshioka N, Ohtomo K (2011) A unified framework for concurrent detection of anatomical landmarks for medical image understanding. In: *SPIE Medical Imaging 2011*. 7962, 03/14/2011 2011. pp 79623E-79623E-79613. doi:10.1117/12.878327

28. Hanaoka S, Shimizu A, Nemoto M, Nomura Y, Miki S, Yoshikawa T, Hayashi N, Ohtomo K, Masutani Y (2016) Automatic detection of over 100 anatomical landmarks in medical CT images: a framework with independent detectors and combinatorial optimization. *Medical Image Analysis* (in press)
29. Vercauteren T, Pennec X, Perchant A, Ayache N (2007) Non-parametric diffeomorphic image registration with the demons algorithm. In: *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2007*. Springer, pp 319-326
30. Seitel A, Rasouljan A, Rohling R, Abolmaesumi P (2015) Lumbar and thoracic spine segmentation using a statistical multi-object shape + pose model, In: Yao J, Glocker B, Klinder T, Li S (Eds.), *Recent Advances in Computational Methods and Clinical Applications for Spine Imaging*. Springer International Publishing, Cham, pp. 221-225.
31. Hammernik K, Ebner T, Stern D, Urschler M, Pock T (2015) Vertebrae segmentation in 3D CT images based on a variational framework, In: Yao J, Glocker B, Klinder T, Li S (Eds.), *Recent Advances in Computational Methods and Clinical Applications for Spine Imaging*. Springer International Publishing, Cham, pp. 227-233.
32. Korez R, Ibragimov B, Likar B, Pernuš F, Vrtovec T (2015) Interpolation-based shape-constrained deformable model approach for segmentation of vertebrae from CT spine images, In: Yao J, Glocker B, Klinder T, Li S (Eds.), *Recent Advances in Computational Methods and Clinical Applications for Spine Imaging*. Springer International Publishing, Cham, pp. 235-240.
33. Castro-Mateos I, Pozo JM, Lazary A, Frangi A (2015) 3D vertebra segmentation by feature selection active shape model, In: Yao J, Glocker B,

Klinder T, Li S (Eds.), *Recent Advances in Computational Methods and Clinical Applications for Spine Imaging*. Springer International Publishing, Cham, pp. 241-245.

34. Knutsson H, Andersson M (2005) Morphons: segmentation using elastic canvas and paint on priors. In: *IEEE International Conference on Image Processing 2005*, 11-14 Sept. 2005. pp II-1226-1229. doi:10.1109/ICIP.2005.1530283

35. Paik NC, Lim CS, Jang HS (2013) Numeric and morphological verification of lumbosacral segments in 8280 consecutive patients. *Spine* 38 (10):E573-578. doi:10.1097/BRS.0b013e31828b7195

36. Hanaoka S, Masutani Y, Nemoto M, Nomura Y, Yoshikawa T, Hayashi N, Yoshioka N, Ohtomo K (2011) Probabilistic modeling of landmark distances and structure for anomaly-proof landmark detection. In: *Proceedings of the Third International Workshop on Mathematical Foundations of Computational Anatomy*, 2011. pp 159-169



# Chapter 5 HoTPiG: A novel geometrical feature for vessel morphometry and its application to cerebral aneurysm detection

A novel feature set for medical image analysis, named HoTPiG (Histogram of Triangular Paths in Graph), is presented. The feature set is designed to detect morphologically abnormal lesions in branching tree-like structures such as vessels. Given a graph structure extracted from a binarized volume, the proposed feature extraction algorithm can effectively encode both the morphological characteristics and the local branching pattern of the structure around each graph node (e.g., each voxel in the vessel). The features are derived from a 3-D histogram whose bins represent a triplet of shortest path distances between the target node and all possible node pairs near the target

node. The extracted feature set is a vector with a fixed length and is readily applicable to state-of-the-art machine learning methods. Furthermore, since my method can handle vessel-like structures without thinning or centerline extraction processes, it is free from the “short-hair” problem and local features of vessels such as caliper changes and bumps are also encoded as a whole.

Using the proposed feature set, a cerebral aneurysm detection application for clinical magnetic resonance angiography (MRA) images was implemented. In an evaluation with 300 datasets, the sensitivities of aneurysm detection were 81.8% and 89.2% when the numbers of false positives were 3 and 10 per case, respectively, thus validating the effectiveness of the proposed feature set.

## 5.1 Introduction

A branching treelike structure is one of the major types of structure in the human body. For example, a wide variety of vessels (blood vessels, bronchi, bile ducts, etc.) have a treelike structure. Quite a large number of diseases affect these vascular structures and cause pathological shape changes including narrowing, occlusion, and dilation. Vascular diseases, including cerebral infarction and coronary occlusive disease, are one of the major causes of death in advanced nations. Since precise evaluation of the shape of vessels is essential in diagnosing these diseases, computer-assisted detection/diagnosis (CAD) of these treelike structures is required.

Among the vascular diseases, cerebral aneurysm has been one of the targets of CAD applications [1-3]. Although unruptured cerebral aneurysms are generally asymptomatic, they rupture in approximately 1% of patients per year,

leading to high rates of mortality and disability [2]. This is why the early detection of cerebral aneurysms is needed. In clinical practice, noninvasive magnetic resonance arteriography (MRA) examination is most frequently used for screening, in which diagnostic radiologists search for abnormal structures (i.e., saccular protuberances and fusiform dilation). However, it is known that a normal arterial system may include pseudo-lesions such as infundibular dilatations. CAD applications for detecting cerebral aneurysm also have to distinguish abnormal aneurysmal structures from normal ones, including branching sites of small cerebral arteries and tightly curving carotid siphons.

In previous studies, two approaches to searching for aneurysms have generally been used: (1) voxel-by-voxel evaluation using Hessian matrix-derived features, and (2) three-dimensional (3-D) thinning of a presegmented arterial region and branching pattern analysis. In the first approach, a Hessian-based filter emphasizes spherical structures with various sizes. For example, Arimura et al. [3] used a dot enhancement filter that outputs a high value when all three eigenvalues of the Hessian matrix have large negative values. Nomura et al. [2] used a similarity index that can distinguish spherelike aneurisms from ridgelike vessels. Although their approach usually works well, the detected candidates inevitably include a large number of false positives, especially at vessel bifurcations. Therefore, subsequent processes to eliminate false positives are required, greatly affecting the overall performance.

The other approach is to find an abnormal arterial branching pattern from the graph structure of an extracted artery. After segmentation of the artery voxels, a 3-D thinning algorithm is applied to extract the centerlines. Then the

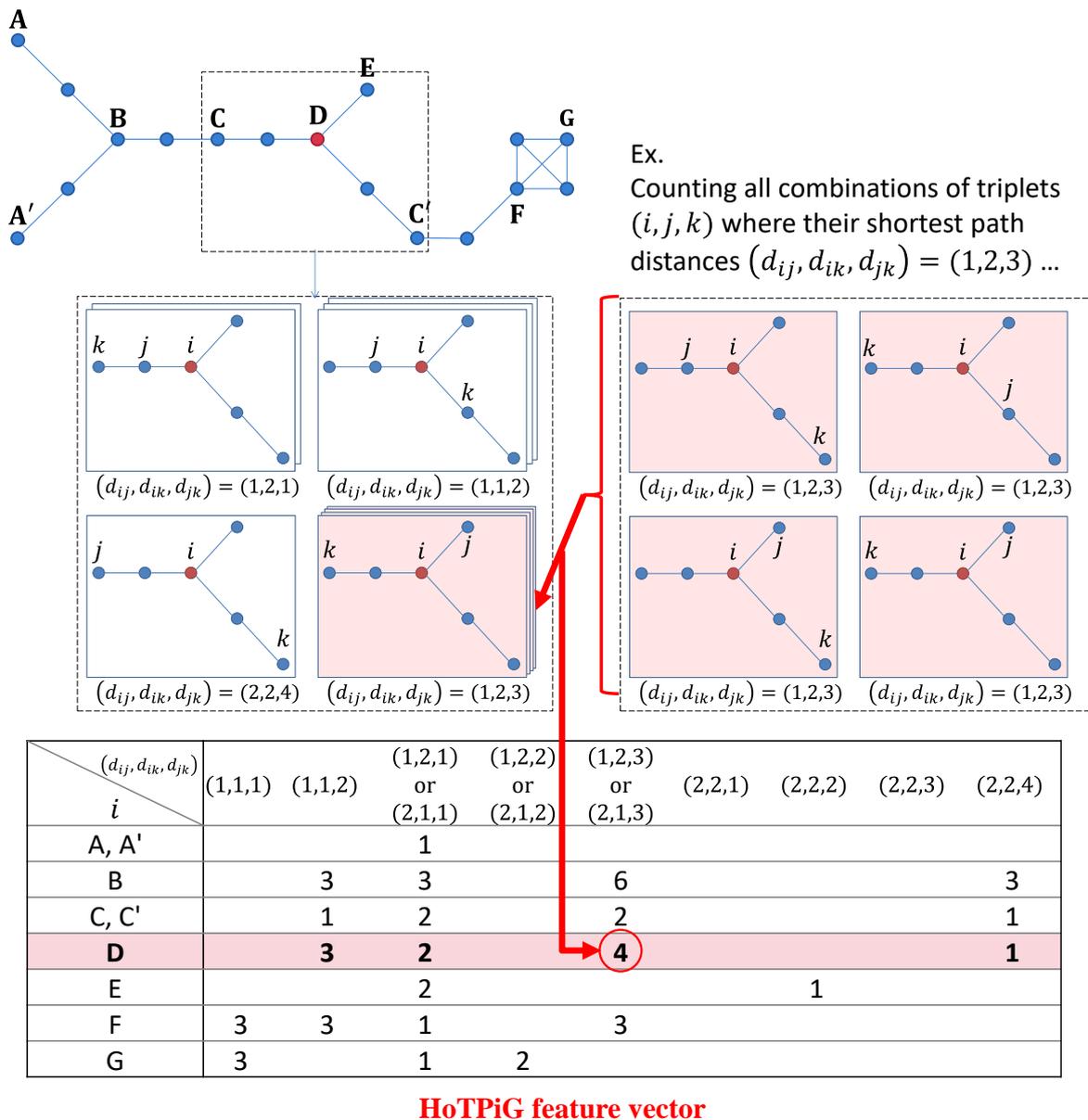
centerlines are analyzed to find any suspicious points, such as end points of centerlines [4], short branches, or points with a locally maximal vascular radius [1]. In contrast to the curvature approach, the centerline approach can utilize branching pattern information to discriminate aneurysms from bifurcations. On the other hand, the 3-D thinning process has the “short hair” problem, i.e., a large number of false short branches where the arterial wall has small “lumps.” Therefore, a postprocess to remove [5] or classify [1, 3] short hairs is indispensable. Another problem is how to represent local morphological and topological changes in the graph in the context of machine learning. A large number of studies on the analysis of whole graph structures have been conducted in which the graph structure is embedded into a vector field (graph embedding) or evaluated by kernel methods (graph kernel) [6].

In this study, I propose a novel feature set named HoTPiG (Histogram of Triangular Paths in Graph). It is defined at each node in a given graph based on a 3-D histogram of shortest path distances between the node of interest and each of its neighboring node pairs. The feature vector efficiently encodes the local graph network pattern around the node. The graph structure can be determined directly from a binary label volume. Since the thickness of the vessel is naturally encoded without any centerline extraction process, the “short hair” problem caused by the thinning algorithm does not occur. Furthermore, the proposed feature is essentially robust to nonrigid deformations under the assumption that the graph structure extracted from the original image is not significantly changed by the deformation. The proposed feature set is sufficiently effective for aneurysms in MRA images to be accurately classified by a support vector machine (SVM) without any complicated pre- or postprocesses. The contributions of this study are as follows: (1) A novel

vector representation of a local graph structure for detecting abnormalities is presented, (2) a CAD application for detecting aneurysms in MRA images is implemented using the proposed graph feature and a state-of-the-art SVM classifier with explicit feature mapping, and (3) the usefulness of the proposed method is experimentally validated using a large dataset with 300 clinical MRA images, and high performance comparable to that of other state-of-the-art methods is demonstrated.

## 5.2 HoTPiG

The proposed HoTPiG feature is defined for any arbitrary undirected graph based on bin counts of a 3-D histogram of shortest path lengths (Fig. 1). One feature vector is determined for each node in the graph and can be readily used to classify the corresponding node as positive (e.g., aneurysm) or negative. The 3-D histogram accumulates counts of each triplet of distances between the target node and its two neighbor nodes as well as between the two neighbor nodes.



**Fig. 2.** Example of calculation of HoTPiG features (with  $d_{max} = 2$ ).

Suppose that the graph includes  $|U|$  nodes, and each node in the graph has an integer index  $l \in U = \{1, 2, 3, \dots, |U|\}$ . Also suppose that the feature vector of node  $i$  is to be calculated. First, the shortest path distances from  $i$  to all other nodes are calculated (by a breadth first search). Here, the shortest

path distance is the number of steps (edges) along the shortest path between the pair of nodes. Let the distance between nodes  $a$  and  $b$  be  $dist(a,b)$ . I define the *neighborhood* of  $i$ ,  $N_i$ , as the set of nodes whose distances from  $i$  are no more than a predefined integer  $d_{max}$ . That means  $N_i = \{l \in U | 0 < dist(i,l) \leq d_{max}\}$ .

Then, for any triplet of distances  $(d_{ij}, d_{ik}, d_{jk})$ , the value of the 3-D histogram  $H_i(d_{ij}, d_{ik}, d_{jk})$  is defined as the number of node pairs  $(j,k)$  that satisfy the following conditions

$$j \in N_i, k \in N_i, dist(i,j) = d_{ij}, dist(i,k) = d_{ik}, dist(j,k) = d_{jk}. \quad (1)$$

In practice, the two bins  $(d_{ij}, d_{ik}, d_{jk})$  and  $(d_{ik}, d_{ij}, d_{jk})$  are simply those with neighbor nodes  $j$  and  $k$  swapped. Thus, these two bins are considered to be the same and only one count is incremented for such pairs of distance triplets.

The counts of bins in histogram  $H_i$  are used as the feature vector of node  $i$ . As shown in Fig. 1, the feature vector tends to vary widely among different nodes and is sensitive to topological changes in the local graph structure. Note that the extent of the locality can be controlled via the parameter  $d_{max}$ .

The calculation cost for the proposed method is estimated as follows. The breadth-first search algorithm can calculate all the shortest path distances by performing  $O(|U| \cdot E(|N_i|))$  calculations, where  $E(|N_i|)$  is the mean size of the neighborhoods. On the other hand, the histogram counting requires  $O(|U| \cdot E(|N_i|)^2)$  count increment calculations. Therefore, most of the calculation cost is for histogram counting.

## 5.3 Computer-assisted detection of aneurysms

As an application of the proposed HoTPiG feature, I have developed CAD software for aneurysm detection in MRA images. The proposed CAD application is composed of four steps: (1) extraction of the binary label volume of arteries from MRA images, (2) calculation of graph structure features, (3) voxel-based classification by SVM, and (4) a thresholding and labeling process.

### 5.3.1 Artery region extraction and HoTPiG feature calculation

Firstly, the artery region is extracted by a conventional region growing method. The average  $\bar{I}$  and standard deviation  $\sigma_I$  of the brain region are estimated by sampling voxel values from a predefined mid-central subregion, that is, a horizontal rectangular plane with half the width and height placed at the center of the volume. Then, the initial artery region is extracted by region growing, where the seed threshold and growing threshold are  $> \bar{I} + 3\sigma_I$  and  $> \bar{I} + 2.5\sigma_I$ , respectively.

After the artery region is extracted, an undirected graph is composed. I choose a simple graph structure whose nodes are all foreground (i.e., intra-arterial) voxels, and the edges connect all 18-neighbor voxel pairs (Fig. 2). Here, an 18-neighborhood is chosen because it is more similar to the Euclidean distance than 6- and 26-neighborhoods.

Using this graph, the HoTPiG feature is calculated at each foreground voxel. The maximum distance used is  $d_{max} = 11$ , considering the balance between the performance and the calculation cost. Two modifications are applied to the method described in Section 2. Firstly, a 1-D histogram with  $d_{max} = 11$

bins whose distances are  $\{1, 2, 3, \dots, 11\}$  may be too sparse when it is used as part of a 3-D histogram. To cope with this, some distances are grouped into one bin and only six bins  $\{1, 2, 3, [4, 5], [6, 7], [8, 11]\}$  are used. These bins are determined so that their upper bounds are the geometric series  $1.5^n, n = 1, 2, 3, 4, 5, 6$ . Applying this bin set to each of three distances  $(d_{ij}, d_{ik}, d_{jk})$ , the entire 3-D histogram has  $6 \cdot {}_6C_2 = 126$  bins. However, some bins never have a count because the corresponding distance triplet does not satisfy the triangle inequality. After removing such bins, a total of 85 bins are included in the 3-D histogram in this study.

Additionally, a multidimensional approach is added to analyze gross vascular structures. After downsampling the artery binary volume to half and a quarter of its original size, the graph structure features are extracted in the same manner. After feature extraction, each feature is upsampled by nearest neighbor interpolation and all the features of the three scales are merged voxel by voxel. Therefore, a total of  $85 \times 3 = 255$  features are calculated for each voxel.

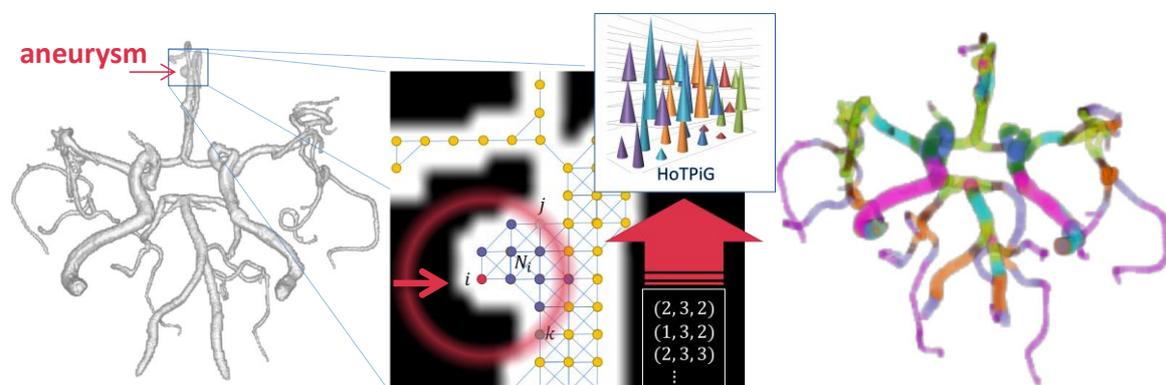
Prior to the classification process, each feature is normalized by dividing by the standard deviation estimated from training datasets.

### 5.3.2 Voxel-based classification by SVM

Using the extracted features, each voxel is classified as positive (aneurysm) or negative (normal artery) by a an SVM classifier [7]. The exponential- $\chi^2$  kernel  $K(\mathbf{x}, \mathbf{y}) = \exp\left(-\frac{1}{2\sigma^2} \cdot \frac{1}{2} \sum_l \frac{(x_l - y_l)^2}{x_l + y_l}\right)$  [8], which is designed specifically for histogram comparison, is used in this study. The classifier is trained using manually inputted aneurysm voxels in the training datasets as positive samples and other arterial voxels as negative samples. Here, one of

the difficulties is the huge number ( $> 10^8$ ) of training samples, because one MRA volume has approximately  $10^6$  artery voxels. It is known that the computational cost of kernel SVM is of order  $O(dM^2) \sim O(dM^3)$ , where  $d$  and  $M$  are the data dimensionality and number of samples, respectively. To solve this problem, I utilized a feature map of the exponential- $\chi^2$  kernel [8] to reduce the original problem to a linear SVM whose computational cost is  $O(dM)$ . The feature map is a function that *explicitly* maps the original feature vector of each sample to a higher-dimensional space, in contrast to the conventional kernel method, in which a vector is *implicitly* mapped to a higher-dimensional space. Using this feature map and the random reduction of negative samples (to 3% of the original number), the training task was calculated in approximately 20 min for  $2 \times 10^8$  original samples.

In addition to the classifier with the HoTPiG feature set only, another classifier is also trained by adding two sets of Hessian-derived features (the dot enhancement filter [3] and shape index [2]) to evaluate the cooperativity of both types of features. The two Hessian-derived features are calculated with six different scales; thus, a total of 12 features are added to the HoTPiG features.



**Fig. 3.** (Left) Example of cerebral arteries in a volume. (Middle) HoTPiG feature calculation for a voxel in an aneurysm. (Right) Result of voxelwise

clustering of HoTPiG features into 20 clusters (displayed by their colors) by a k-means method. Note that the vessel thickness and branching pattern can be clearly distinguished. Furthermore, the mirror symmetry of the clustering result implies its robustness against local deformations and sensitivity to caliper changes.

In this study, the parameters of feature mapping are set to  $m = 5000$ ,  $n = 2$ , and  $L = 0.6$ , referring to [8]. The parameters of the kernel  $\sigma$  and the linear SVM  $C$  are experimentally optimized (as described later in the next section).

Using the output values of the SVM, candidate aneurysm lesions and their lesionwise likelihoods are determined as follows. Firstly, the SVM outputs are thresholded by zero and all nonpositive voxels are discarded. Then, the positive voxels are labeled by connected component analysis and all connected components are outputted as candidate lesions. The likelihood of each lesion is determined as the maximal value of SVM-derived likelihoods of the voxels in the lesion. The representative point of each lesion is defined as this maximal value point.

## 5.4 Experimental results

This study was approved by the ethical review board of our institution. A total of 300 time-of-flight cerebral MRA volumes with 333 aneurysms were used in the experiment. The voxel size was  $0.469 \times 0.469 \times 0.6$  mm. Two board-certified radiologists diagnosed all images and manually inputted aneurysm regions.

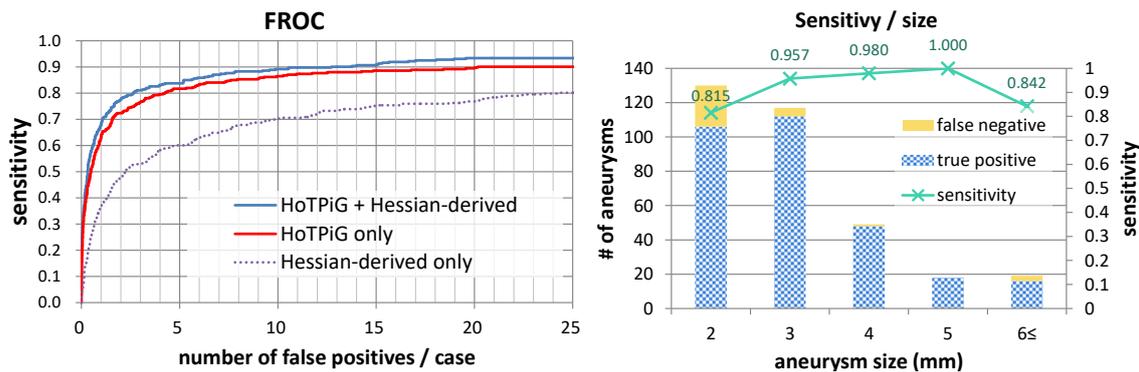
The proposed method was evaluated using 3-fold cross-validation. Before the actual training, a hyperparameter optimization was performed in each fold using another nested 3-fold cross validation. The optimal values of the two parameters  $\sigma$  and  $C$  were searched for from the search space  $\sigma \in \{20,30,40\}$  and  $C \in \{10,20,30\}$  by performing a grid search. After optimization, the actual training was performed using all training datasets. The calculation of HoTPiG features took approximately 3 min per case using a workstation with  $2 \times 6$  core Intel Xeon processors and 72 GB memory.

The overall performance of the method was evaluated using the free receiver operating characteristic (FROC) curve. Each outputted lesion was determined as a successful detection if the representative point of the lesion was no more than 3 mm from the center of gravity of the ground truth region.

Figure 3 shows the FROC curves of the proposed method with and without additional Hessian features, as well as the one with Hessian features only. The sensitivities with only HoTPiG features were 76.6% and 86.5% when the numbers of false positives (FPs) were 3 and 10 per case, respectively. When combined with Hessian features, the sensitivities increased to 81.8% and 89.2% for 3 and 10 FPs/case, respectively. Although a strict comparison cannot be made owing to the different datasets used, the sensitivity of 81.8% is superior to that reported by Yang et al. [1], whose detection sensitivity was 80% for 3 FPs/case. On the other hand, Nomura et al. [2] reported sensitivities of 81.5% and 89.5% when the training dataset sizes were 181 and 500 (also for 3 FPs/case), respectively. Since I used 200 datasets to train each SVM, I conclude that the performance of my CAD is comparable to that of Nomura et al. when the dataset size is equal. Note that Nomura et al. did not use any

objective criterion (e.g., maximum acceptable error distance) to judge lesions outputted by CAD as true positives or false positives; instead, radiologists subjectively decided whether or not each CAD-outputted lesion corresponded to an aneurysm.

Figure 3 also shows the sensitivity for each aneurysm size (maximized by using 21 FPs/case). Most detection failures occurred when the size of the aneurysms was less than 4 mm. On the other hand, my method failed to detect two large aneurysms whose sizes were 6 mm and 13 mm. This was very likely to have been due to a shortage of large aneurysms (only 19 with sizes  $\geq 6$  mm) in my dataset.



**Fig. 4.** (Left) FROC curves for the proposed method with and without additional Hessian-derived features. (Right) Sensitivities (with Hessian features) for each aneurysm size.

## 5.5 Discussion

Among the various vascular diseases that involve the human body, aneurysms are characterized by their particular protuberant shape. This study was inspired by the fact that many radiologists rely on 3-D reconstructed

vascular images to find aneurysms and other diseases in daily image interpretation. This implies that only the shape of the tissue can be sufficient to detect such abnormalities. The HoTPiG feature is designed to evaluate only the shape of the tissue and discard all image intensity information including image gradations and textures. This can be both a disadvantage and an advantage of HoTPiG. On the one hand, it can only utilize a small part of the information provided by the original image. On the other hand, HoTPiG can reveal image characteristics very different from those collected by most other image features based on image intensity information. Indeed, HoTPiG showed cooperativity with existing Hessian-based features which has a weakness at branching sites of vessels. The effectiveness of HoTPiG shown in this study may also be owing to the robustness of HoTPiG against local deformations. Therefore, I believe that HoTPiG will be a powerful alternative tool for vectorizing shape characteristics of vessel-like organs.

## 5.6 Conclusion

A novel HoTPiG feature set for evaluating vessel-like shapes was presented. It showed high performance for detecting cerebral aneurysms and cooperativity with existing image features. My future works may include the application of HoTPiG to other applications such as lung nodule detection, in which discrimination between lesions and vascular bifurcations has similar importance.

## References

Chapter 5 HoTPiG: A novel geometrical feature for vessel morphometry and its application to cerebral aneurysm detection

1. Yang, X., Blezek, D.J., Cheng, L.T., Ryan, W.J., Kallmes, D.F., Erickson, B.J.: Computer-aided detection of intracranial aneurysms in MR angiography. *Journal of digital imaging* 24, 86-95 (2011)
2. Nomura, Y., Masutani, Y., Miki, S., Nemoto, M., Hanaoka, S., Yoshikawa, T., Hayashi, N., Ohtomo, K.: Performance improvement in computerized detection of cerebral aneurysms by retraining classifier using feedback data collected in routine reading environment. *Journal of Biomedical Graphics and Computing* 4, p12 (2014)
3. Arimura, H., Katsuragawa, S., Suzuki, K., Li, F., Shiraishi, J., Sone, S., Doi, K.: Computerized scheme for automated detection of lung nodules in low-dose computed tomography images for lung cancer screening 1. *Academic Radiology* 11, 617-629 (2004)
4. Suniaga, S., Werner, R., Kemmling, A., Groth, M., Fiehler, J., Forkert, N.D.: Computer-aided detection of aneurysms in 3D time-of-flight MRA datasets. *Machine Learning in Medical Imaging*, pp. 63-69. Springer (2012)
5. Kobashi, S., Kondo, K., Yutaka, H.: Computer-aided diagnosis of intracranial aneurysms in MRA images with case-based reasoning. *IEICE transactions on information and systems* 89, 340-350 (2006)
6. Suard, F., Guigue, V., Rakotomamonjy, A., Benschrair, A.: Pedestrian detection using stereo-vision and graph kernels. In: *Intelligent Vehicles Symposium, 2005. Proceedings. IEEE*, pp. 267-272. IEEE, (2005)
7. Fan, R.-E., Chang, K.-W., Hsieh, C.-J., Wang, X.-R., Lin, C.-J.: LIBLINEAR: A library for large linear classification. *The Journal of Machine Learning Research* 9, 1871-1874 (2008)

Chapter 5 HoTPiG: A novel geometrical feature for vessel morphometry and its application to cerebral aneurysm detection

8. Sreekanth, V., Vedaldi, A., Zisserman, A., Jawahar, C.: Generalized RBF feature maps for efficient detection. In: Proceedings of the British Machine Vision Conference (BMVC), 31 Aug - 3 September 2010, Aberystwyth. 1-11 (2010)

## Chapter 6 Conclusion

Anatomical landmarks are one of the most primitive representations of anatomical knowledge. As known, a lot of anatomically salient points have been defined as landmarks and given their own names in anatomy. They are used by physicians in their clinical daily works, as well as by many medical image processing applications. Detecting, defining and using landmarks are the theme of this thesis.

In Chapter 2, I introduced a framework to detect over 100 landmarks simultaneously. Since landmark detection process is usually used as preprocessing, its accuracy and robustness are important. The proposed framework uses an L-PDM which is a statistical model of spatial distribution of landmarks. Therefore, the framework can accurately detect many landmarks which are difficult to be detected individually. Moreover, the framework can estimate positions of landmarks which are out of the imaging

range or undetected by the corresponding detector. Therefore, I believe the proposed method has sufficient reliability for most of medical image analysis applications.

In Chapter 3 I attempted to define landmarks automatically. The criterion named TCC was introduced, under an assumption that anatomical landmark points must be registered correctly and consistently in most of volume triplets. The experimental result showed that the proposed method can determine anatomically salient points in the human body. This is a feasibility study and it requires many future works including validation with other image modalities, automatic detection of newly defined landmarks, and applying to medical image analysis applications.

Since landmarks are one of the most primitive representations of anatomical knowledge, it has a wide variety of applications in medical image analysis. In Chapter 4 I introduced an application which performs registration-based segmentation (a multiatlas method) of the bone. The registration method used is a combination of the diffeomorphic demons algorithm and landmark trajectory-based guidance. The method was applied to the bony structure of the spine and the pelvis and the performance was comparable to other state-of-the-art methods.

In Chapter 5, I introduced a new image feature set named HoTPiG. Although HoTPiG is not directly relevant to landmarks, it is very likely that HoTPiG can be used to detect landmarks defined on vessel-like structures, such as branching points of arteries or bronchi. Because a lot of clinically important landmarks are defined on vessel-like structures, HoTPiG can be used widely in the landmark detection. Combining HoTPiG and landmark

detection will be one of my future works.

Among other future works, I would like to attempt to detect anatomical variant automatically. Especially I am now focusing upon detecting vertebral number anomalies. It is a very challenging problem to detect a series of landmarks where the number of landmarks can vary among subjects. Another challenging future work is to detect vessel bifurcation landmarks (e.g. airways and blood vessels) where the bifurcation pattern can vary.

In conclusion, methods for detection, definition and application of landmarks were developed and discussed in this thesis. It is obvious that both the number of landmarks and reliability of their detection are not fully satisfactory, thus more research is desired in this field. It can be a goal of this research field that the human body is filled with automatically detectable landmarks.



## Appendix A: List of Publications

### Papers used in this thesis:

#### International journal paper

- 1) Hanaoka S, Shimizu A, Nemoto M, Nomura Y, Miki S, Yoshikawa T, Hayashi N, Ohtomo K, Masutani Y. Automatic detection of over 100 anatomical landmarks in medical CT images: a framework with independent detectors and combinatorial optimization. *Medical Image Analysis* (impact factor 4.565), Volume 35 , 192 – 214
- 2) Hanaoka S, Masutani Y, Nemoto M, Nomura Y, Miki S, Yoshikawa T, Hayashi N, Ohtomo K, Shimizu A. Landmark-guided diffeomorphic demons algorithm and its application to automatic segmentation of the whole spine and pelvis in CT images. *International journal of computer*

assisted radiology and surgery (2016) 1-18.

### International conference papers

- 1) Hanaoka S, Nomura Y, Nemoto M, Miki S, Yoshikawa T, Hayashi N, Ohtomo K, Masutani Y, Shimizu A. HoTPiG: A novel geometrical feature for vessel morphometry and its application to cerebral aneurysm detection. Medical Image Computing and Computer-Assisted Intervention--MICCAI 2015. Springer International Publishing, 2015. 103-110.
- 2) Hanaoka S, Nomura Y, Nemoto M, Miki S, Yoshikawa T, Hayashi N, Ohtomo K, Shimizu A. Fully automatic definition of anatomical landmarks in medical images: a feasibility study. Int J CARS 11 (Suppl 1):S166-167, 2016.

### Other publications:

#### International journal paper

- 1) Hanaoka S, Nakano Y, Nemoto M, Nomura Y, Takenaga T, Miki S, Yoshiwaka T, Hayashi N, Masutani Y, Shimizu A. Automatic detection of vertebral number abnormalities in body CT images. International Journal of Computer Assisted Radiology and Surgery (accepted).
- 2) Hanaoka S, Nomura Y, Nemoto M, Masutani Y, Maeda E, Yoshikawa T, Hayashi N, Yoshioka N, Ohtomo K. Automated segmentation method for

## Appendix A: List of Publications

- spinal column based on a dual elliptic column model and its application for virtual spinal straightening. *J Comput Assist Tomogr.* 2010 Jan;34(1):156-62.
- 3) Miki S, Hayashi N, Masutani Y, Nomura Y, Yoshikawa T, Hanaoka S, Nemoto M, Ohtomo K. Computer-assisted detection of cerebral aneurysms in MR angiography in a routine image-reading environment: Effects on diagnosis by radiologists. *American Journal of Neuroradiology*, vol.37, no.6, pp.1038-1043, 2016.
  - 4) Nomura Y, Masutani Y, Miki S, Nemoto M, Hanaoka S, Yoshikawa T, Hayashi N, Ohtomo K. Performance improvement in computerized detection of cerebral aneurysms by retraining classifier using feedback data collected in routine reading environment. *Journal of Biomedical Graphics and Computing* (2014) 4(4), 12-30.
  - 5) Yasaka K, Maeda E, Hanaoka S, Katsura M, Sato J, Ohtomo K. Single-energy metal artifact reduction for helical computed tomography of the pelvis in patients with metal hip prostheses. *Jpn J Radiol.* 2016 Sep;34(9):625-32.
  - 6) Yasaka K, Katsura M, Hanaoka S, Sato J, Ohtomo K. High-resolution CT with new model-based iterative reconstruction with resolution preference algorithm in evaluations of lung nodules: Comparison with conventional model-based iterative reconstruction and adaptive statistical iterative

- reconstruction. *Eur J Radiol.* 2016 Mar;85(3):599-606.
- 7) Nemoto M, Yeernuer T, Masutani Y, Nomura Y, Hanaoka S, Miki S, Yoshikawa T, Hayashi N, Ohtomo K. Development of automatic visceral fat volume calculation software for CT volume data. *J Obes.* 2014;2014:495084.
  - 8) Nomura Y, Hayashi N, Masutani Y, Yoshikawa T, Nemoto M, Hanaoka S, Miki S, Maeda E, Ohtomo K. CIRCUS: an MDA platform for clinical image analysis in hospitals. *Transactions on Mass-Data Analysis of Images and Signals*, vol.2, no.1, pp.112-127, 2010.
  - 9) Maeda E, Yoshikawa T, Hayashi N, Akai H, Hanaoka S, Sasaki H, Matsuda I, Yoshioka N, Ohtomo K. Radiology reading-caused fatigue and measurement of eye strain with critical flicker fusion frequency. *Jpn J Radiol.* 2011 Aug;29(7):483-7. Epub 2011 Sep 1.
  - 10) Furuta T, Maeda E, Akai H, Hanaoka S, Yoshioka N, Akahane M, Watadani T, Ohtomo K. Hepatic segments and vasculature: projecting CT anatomy onto angiograms. *Radiographics.* 2009 Nov;29(7):1-22. Epub 2009 Sep 4.
  - 11) Takayama Y, Kishimoto R, Hanaoka S, Nonaka H, Kandatsu S, Tsuji H, Tsujii H, Ikehira H, Obata T. ADC value and diffusion tensor imaging of prostate cancer: changes in carbon-ion radiotherapy. *J Magn Reson Imaging.* 2008 Jun;27(6):1331-5.

## Appendix A: List of Publications

- 12) Matsuda I, Hanaoka S, Akahane M, Sato S, Komatsu S, Inoh S, Kiryu S, Yoshioka N, Ino K, Ohtomo K. Adaptive statistical iterative reconstruction for volume-rendered computed tomography portovenography: improvement of image quality. Japanese journal of radiology. 2010, 28(9), 700-706
- 13) Katsura M, Matsuda I, Akahane M, Yasaka K, Hanaoka S, Akai H, Sato J, Kunimatsu A, Ohtomo K. Model-based iterative reconstruction technique for ultralow-dose chest CT: comparison of pulmonary nodule detectability with the adaptive statistical iterative reconstruction technique. Invest Radiol. 2013 Apr;48(4):206-12
- 14) Amemiya S, Takao H, Hanaoka S, Ohtomo K. Global and structured waves of rs-fMRI signal identified as putative propagation of spontaneous neural activity. Neuroimage. 2016 Jun;133:331-40.

## International conference paper

- 1) Hanaoka S, Fritscher K, Welk M, Nemoto M, Masutani Y, Hayashi N, Ohtomo K, Schubert R. 3-D graph cut segmentation with Riemannian metrics to avoid the shrinking problem. Med Image Comput Comput Assist Interv. 2011;14(Pt 3):554-61.
- 2) Hanaoka S, Fritscher KD, Schuler B, Masutani Y, Hayashi N, Ohtomo K, Schubert R. Whole vertebral bone segmentation method with a statistical intensity-shape model based approach. Medical Imaging 2011: Image

- Processing. Proceedings of the SPIE, Volume 7962, pp. 796242-796242-14 (2011).
- 3) Nemoto M, Masutani Y, Hanaoka S, Nomura Y, Yoshikawa T, Hayashi N, Yoshioka N, Ohtomo K. A unified framework for concurrent detection of anatomical landmarks for medical image understanding Medical Imaging 2011: Image Processing. Proceedings of the SPIE, Volume 7962, pp. 79623E-79623E-13 (2011)
  - 4) Hanaoka S, Nomura Y, Nemoto M, Masutani Y, Yoshioka N, Yoshikawa T, Maeda E, Hayashi N, Ohtomo K. Automated segmentation method for spinal column based on parametric model and its application for curved MPR display. Present at Computer Assisted Radiology and Surgery, June 2008. Int J CARS (2008) 3 (Suppl 1):S398-399.
  - 5) Nomura Y, Nemoto M, Masutani Y, Maeda E, Ito H, Hanaoka S, Yoshikawa T, Hayashi N, Yoshioka N, Ohtomo K. Computerized detection of pulmonary nodules based on dual classifiers for juxta-pleural and non-pleural nodules. Int J CARS (2008) 3 (Suppl 1):S186-187.
  - 6) Nemoto M, Nomura Y, Masutani Y, Hanaoka S, Yoshioka N, Yoshikawa T, Hayashi N, Ohtomo K. Development of automated detection system for skin lesions in whole-body PET/CT images. Int J CARS (2008) 3 (Suppl 1):S433-434.
  - 7) Hanaoka S, Masutani Y, Nomura Y, Nemoto M, Yoshikawa T, Hayashi N,

## Appendix A: List of Publications

- Yoshioka N, Ohtomo K. Vertebral body segmentation algorithm for whole spine CT images with various pathological changes. *Int J CARS* (2010) 5 (Suppl 1):S84-86.
- 8) Hanaoka S, Masutani Y, Nemoto M, Nomura Y, Yoshikawa T, Hayashi N, Ohtomo K. An improved multiple anatomical landmark detection method with combinatorial optimization and Madaboost-based candidate likelihood determination. *Int J CARS* (2012) 7 (Suppl 1):S330-331
- 9) Hanaoka S, Masutani Y, Nemoto M, Nomura Y, Yoshikawa T, Hayashi N, Ohtomo K. Automatic Categorization of Anatomical Landmark-Local Appearances Based on Diffeomorphic Demons and Spectral Clustering for Constructing Detector Ensembles. *Med Image Comput Comput Assist Interv.* 2012; (Pt 2):106-113.
- 10) Inagaki S, Imiya A, Hontani H, Hanaoka S, Masutani Y. Scale Space and Variational Methods in Computer Vision. *Proc. of SSVM 2013, LNCS 7893*, pp.440-451 (2013)
- 11) Hanaoka S, Nomura Y, Nemoto M, Miki S, Yoshikawa T, Hayashi N, Ohtomo K, Shimizu A. Fully automatic definition of anatomical landmarks in medical images: a feasibility study. *Int J CARS* 11 (Suppl 1):S166-167, 2016
- 12) Hanaoka S, Masutani Y, Nemoto M, Nomura Y, Yoshikawa T, Hayashi N, Yoshioka N, Ohtomo K. Probabilistic Modeling of Landmark Distances

- and Structure for Anomaly-proof Landmark Detection. Proceedings of the Third International Workshop on Mathematical Foundations of Computational Anatomy - Geometrical and Statistical Methods for Modelling Biological Shape Variability (2011) 159-169
- 13) Nemoto M, Nomura Y, Hanaoka S, Masutani Y, Yoshikawa T, Hayashi N, Yoshioka N, Ohtomo K. Preliminary study on appearance-based detection of anatomical point landmarks in body trunk CT images. Machine Learning in Medical Imaging, Lecture Notes in Computer Science, Volume 6357. Springer-Verlag Berlin Heidelberg, 2010, p. 174
- 14) Nomura Y, Hayashi N, Masutani Y, Yoshikawa T, Nemoto M, Hanaoka S, Maeda E, Ohtomo K. An integrated platform for development and clinical use of CAD software: building and utilization in the clinical environment. Int J CARS, vol.4, suppl.1, pp.S161-S162, June 2009
- 15) Nomura Y, Hayashi N, Masutani Y, Yoshikawa T, Nemoto M, Ohtomo K, Hanaoka S, Maeda E. An integrated platform for clinical use of CAD software and feedback. Proc. of RSNA 2009, pp.919 (LL-IN2158-R01), Nov. 2009
- 16) Nomura Y, Hayashi N, Masutani Y, Yoshikawa T, Nemoto M, Hanaoka S, Miki S, Ohtomo K. Development and long-term validation of web-based CAD server for synergic performance improvement of CAD software and radiologists. ECR 2011 Book of Abstracts, pp.S190, Mar. 2011

## Appendix A: List of Publications

- 17) Nomura Y, Masutani Y, Hayashi N, Miki S, Nemoto M, Hanaoka S, Yoshikawa T, Ohtomo K. Additional learning of CAD software based on multicenter trial in teleradiology environment. Int J CARS (2012) 7 (Suppl 1):S270-271
- 18) Miki S, Nomura Y, Hayashi N, Masutani Y, Yoshikawa T, Nemoto M, Hanaoka S, Ohtomo K. Do radiologists have confidence in CAD for cerebral aneurysms? Evaluation in routine reading environment. Proc. of RSNA 2011, SSC08-02, Nov. 2011
- 19) Hanaoka S, Masutani Y, Nemoto M, Nomura Y, Miki S, Yoshikawa T, Hayashi N, Yoshioka N, Ohtomo K. Sparse Gaussian graphical model estimation for spatial distribution of multiple anatomical landmarks in the human body - a GPGPU implementation of the graphical lasso algorithm and its application to automatic landmark detection system. Int J CARS (2013) 8 (Suppl 1):S288-289.
- 20) Hanaoka S, Masutani Y, Nemoto M, Nomura Y, Miki S, Yoshikawa T, Hayashi N, Yoshioka N, Ohtomo K. Sparse Gaussian graphical model on spatial distribution of multiple anatomical landmarks – model construction from training datasets with insufficient imaging ranges. Proceedings of the Third International Workshop on Mathematical Foundations of Computational Anatomy (2013) 107-116.

## Domestic conference papers:

- 1) 天野 敏之 , 玉木 徹 , 村上 剛基 , 青木 康昌 , 尾藤 幸司 , 秋間 雄太 , 荒木 孝介 , 佐古 祐 , 花岡 昇平. 2007年 PRMUアルゴリズムコンテスト : 「ジグソーパズルを完成させよう!-マッチングによる画像復元-」 実施報告と受賞アルゴリズム紹介. 電子情報通信学会技術研究報告.PRMU, パターン認識・メディア理解 107(384), 125-136, 2007-12-06.
- 2) Hanaoka S, Kishimoto R, Obata T, Yanagi T, Tsuji H, Ikehira H, Tanada S, Tsujii H. Diffusion tensor imaging of body trunk: preliminary study with botanic phantom, healthy individual and prostate cancer patients. Present at The 63th Annual Meeting of Japan Radiological Society, April 2004 (in Japanese)
- 3) Hanaoka S, Nomura Y, Nemoto M, Masutani Y, Yoshioka N, Yoshikawa T, Maeda E, Hayashi N, Ohtomo K. Spine and spinal canal segmentation in body trunk CT images with elliptic column model. Presented in Annual Meeting of Japan Society of Computer Aided Diagnosis of Medical Images, October 2007 (in Japanese)
- 4) 花岡 昇平, 増谷 佳孝, 野村 行弘, 根本 充貴, 前田 恵理子, 吉川 健啓, 林 直人, 吉岡 直紀, 大友 邦. 側彎に対応した全脊椎の自動CTセグメンテーション. 第27回日本医用画像工学会大会. 2008.8.5-6. 法政大学小金井キャンパス : 東京都小金井市梶野町 3-7-2

## Appendix A: List of Publications

- 5) 花岡 昇平, 増谷 佳孝, 野村 行弘, 根本 充貴, 前田 恵理子, 吉川 健啓, 林 直人, 吉岡 直紀, 大友 邦. 体幹部 CT 画像における椎体骨転移のコンピュータ支援画像診断の初期的検討. 第 28 回 日本医用画像工学会大会. 2009.8.4-5. 中京大学名古屋キャンパス名古屋市昭和区八事本町 101-2
- 6) 花岡 昇平, 増谷 佳孝, 野村 行弘, 根本 充貴, 前田 恵理子, 吉川 健啓, 林 直人, 吉岡 直紀, 大友 邦. 多数のランドマーク位置候補検出結果からの最適組み合わせ検索アルゴリズム. 電子情報通信学会技術研究報告.MI, 医用画像 110(195), 67-74, 2010-08-27.
- 7) 花岡 昇平, 増谷 佳孝, 野村 行弘, 根本 充貴, 前田 恵理子, 吉川 健啓, 林 直人, 吉岡 直紀, 大友 邦. 一般化円筒モデルに基づく脊椎椎体の海綿骨, 皮質骨の分離と標準化骨濃度マップの作成. 電子情報通信学会技術研究報告.MI, 医用画像 109(127), 33-38, 2009-07-08
- 8) 花岡 昇平, 野村 行弘, 根本 充貴, 増谷 佳孝, 前田 恵理子, 吉川 健啓, 林 直人, 吉岡 直紀, 大友 邦. 後方要素に対応した全椎骨のセグメンテーションアルゴリズム. 電子情報通信学会技術研究報告.MI, 医用画像 108(271), 21-26, 2008-10-23
- 9) 花岡 昇平, 増谷 佳孝, 野村 行弘, 根本 充貴, 前田 恵理子, 吉川 健啓, 林 直人, 大友 邦. 脊柱ランドマーク検出結果を利用した全椎骨セグメンテーションの精度改善の試み. 電子情報通信学

会技術研究報告.MI, 医用画像 111(199), 11-16, 2011-08-30

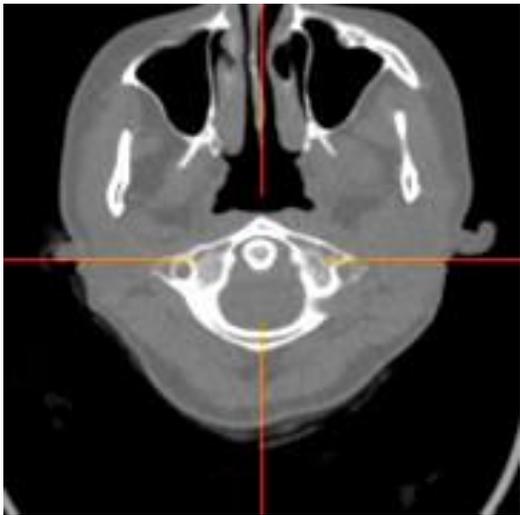
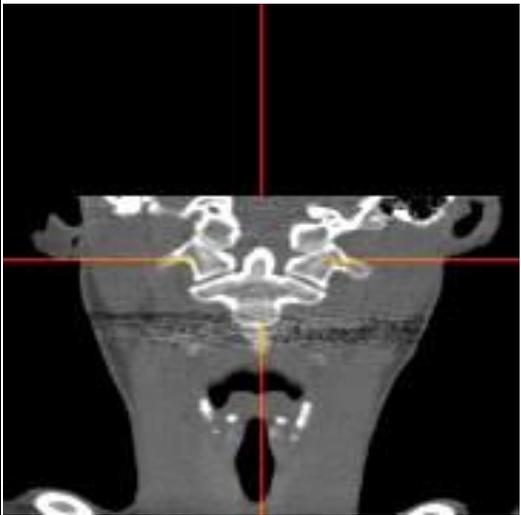
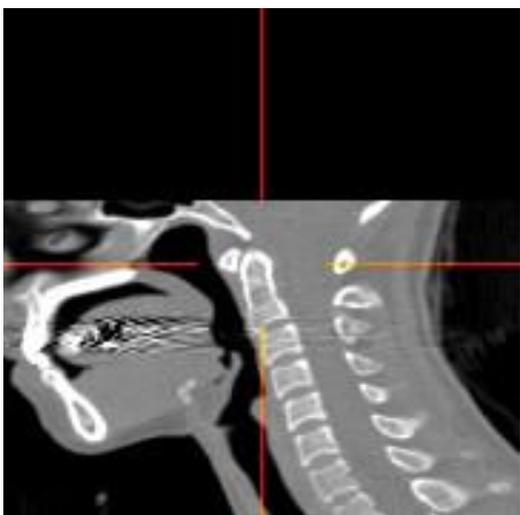
- 10) 花岡 昇平, 増谷 佳孝, 根本 充貴, 野村 行弘, 三木 聡一郎, 吉川 健啓, 林 直人, 大友 邦. 撮像範囲外ランドマーク位置の高速推定アルゴリズムを用いた LM 検出システムの性能改善. 電子情報通信学会技術研究報告.MI, 医用画像 MI2011-115, pp.209-214
- 11) 花岡 昇平, 増谷 佳孝, 根本 充貴, 野村 行弘, 三木 聡一郎, 吉川 健啓, 林 直人, 大友 邦. 撮像範囲が部分的かつ不一致な医用画像群からの統計学的ランドマーク空間分布モデルの作成. 電子情報通信学会技術研究報告.MI, 医用画像 MI2011-67, pp.25-30

#### Thesis (PhD of Medicine)

- 1) 花岡 昇平. CT 画像における脊椎骨転移のコンピュータ支援検出に関する基礎的研究. 平成 22 年 1 月.

## Appendix B Anatomical landmark list

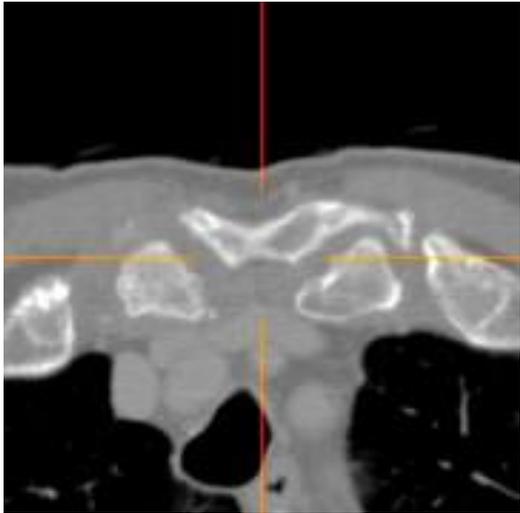
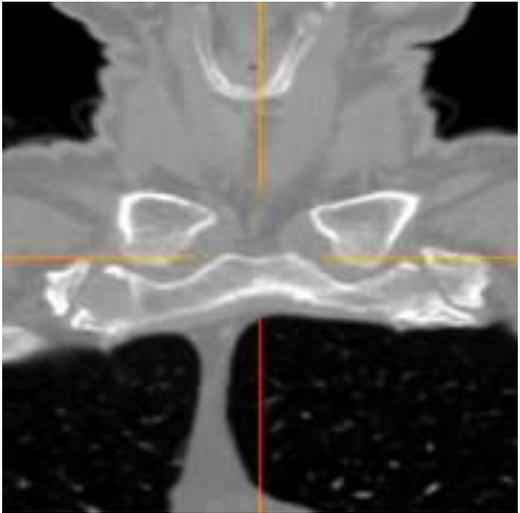
## 環軸関節歯突起中心 [Atlantoaxial\_J] (atlantoaxial joint, center of dens)

	AX	COR
AX		
SAG		

大体環椎の椎弓が見えるスライスで、歯突起の中心点。(骨のランドマークだが、例外的に骨皮質上に点をとらないことに注意)

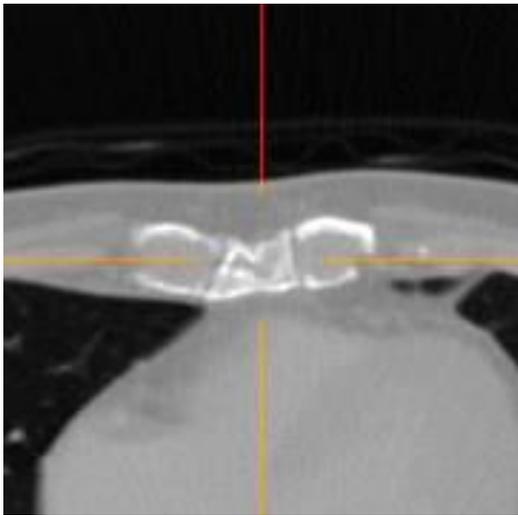
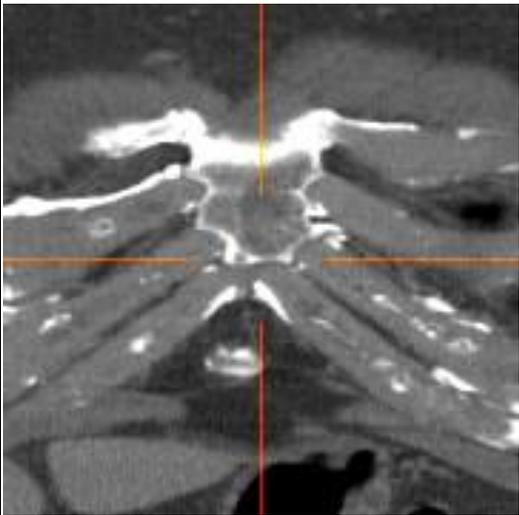
## Appendix B Anatomical landmark list

## 胸骨上端 [Sternum\_SupTip] (superior tip of sternum)

	AX	COR
AX		
SAG		

矢状断で、胸骨正中の上端。胸骨の厚みについてもできるだけ中央をとること。

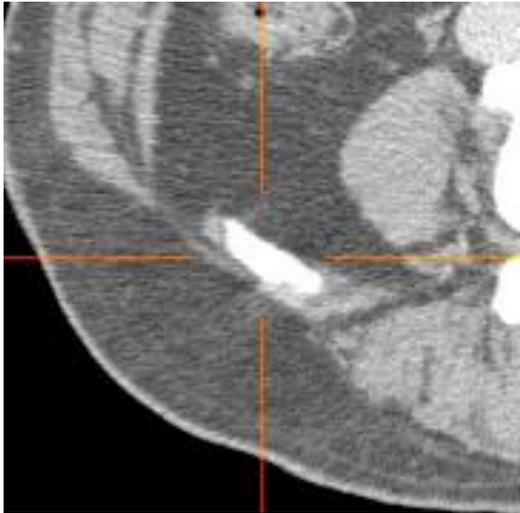
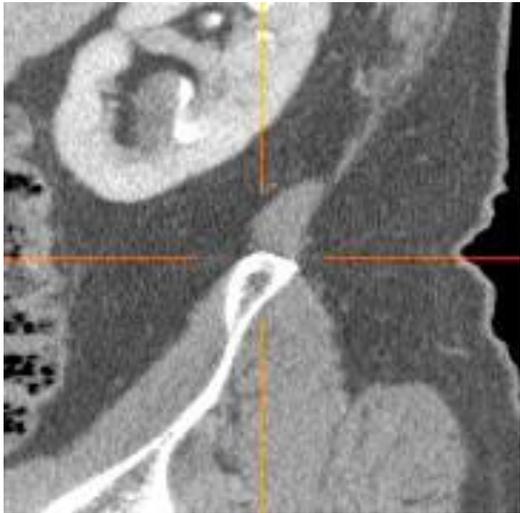
胸骨下端（剣状突起除く） [Sternum\_InfTip] (inferior tip of sternum)

	AX	COR
AX		
SAG		

矢状断で、(剣状突起は除いた)胸骨正中の下端。

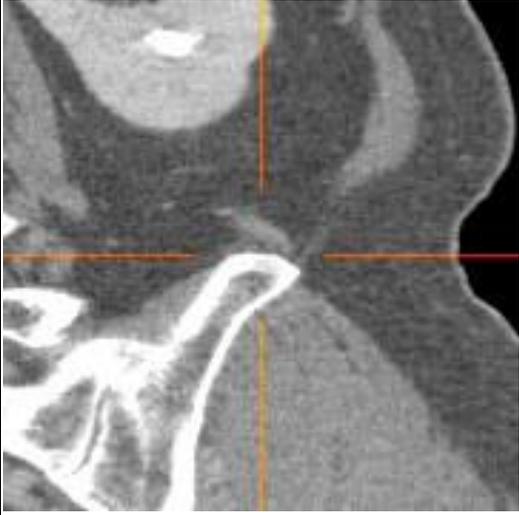
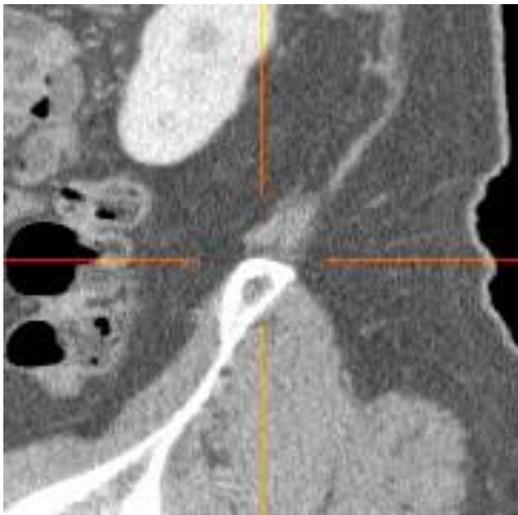
## Appendix B Anatomical landmark list

## 右腸骨稜上端 [R\_IliacCrest\_Sup] (rt. superior margin of iliac crest)

	AX	COR
AX		
SAG		

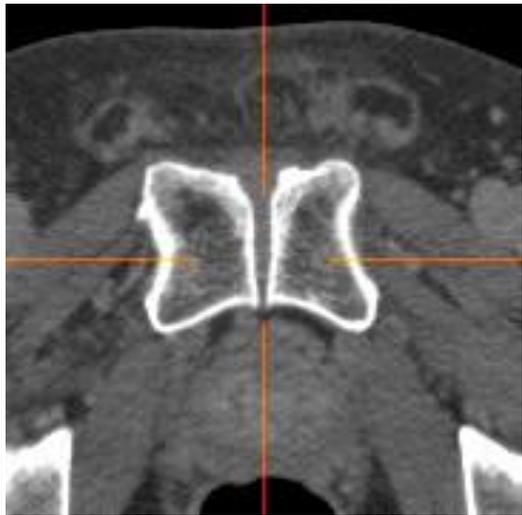
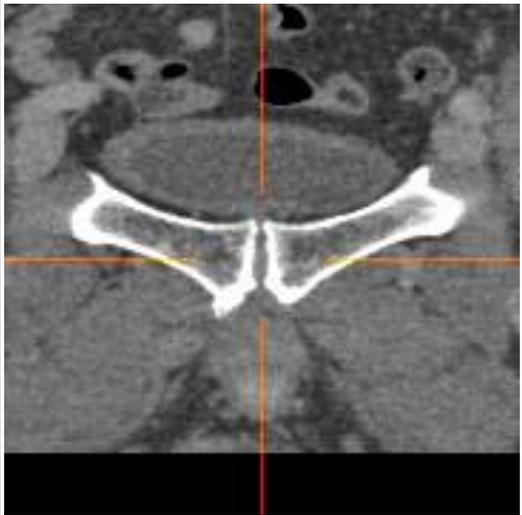
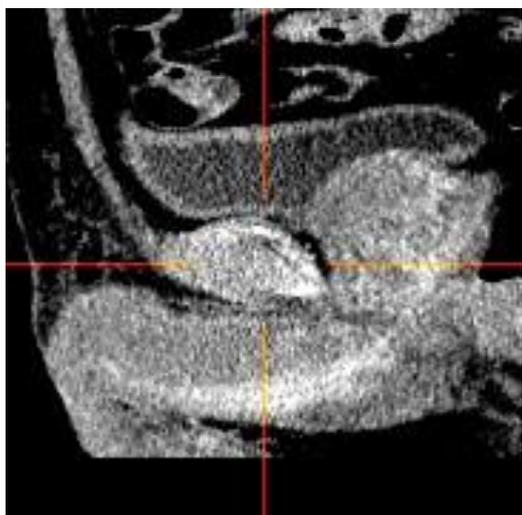
腸骨稜の、水平断で見たときの上縁のスライスで骨皮質上をとる。側彎がひどい場合は頭側方向を考えて適宜補正する。

## 左腸骨稜上端 [L\_IliacCrest\_Sup] (lt. superior margin of iliac crest)

	AX	COR
AX		
SAG		

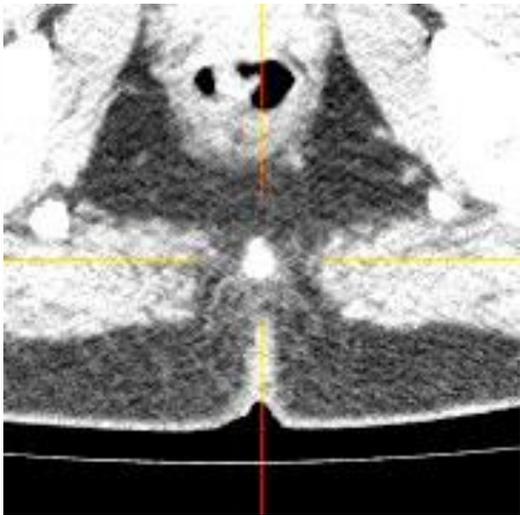
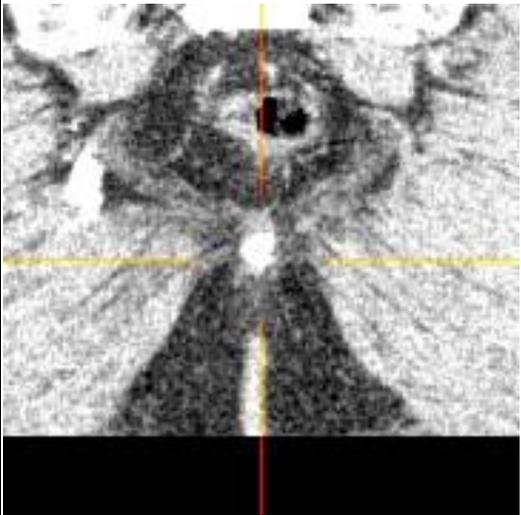
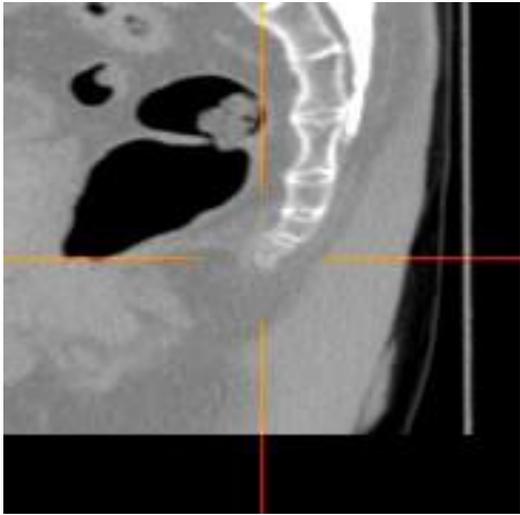
## Appendix B Anatomical landmark list

## 恥骨結合 [PubicSymphysis]

	AX	COR
AX		
SAG		

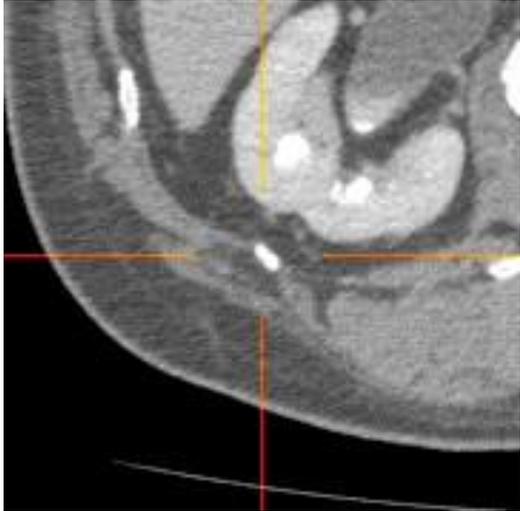
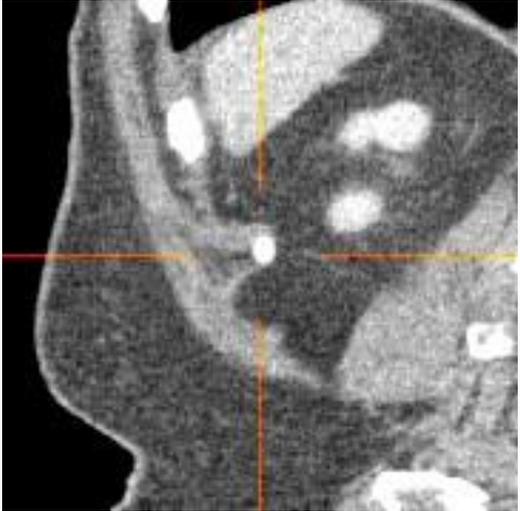
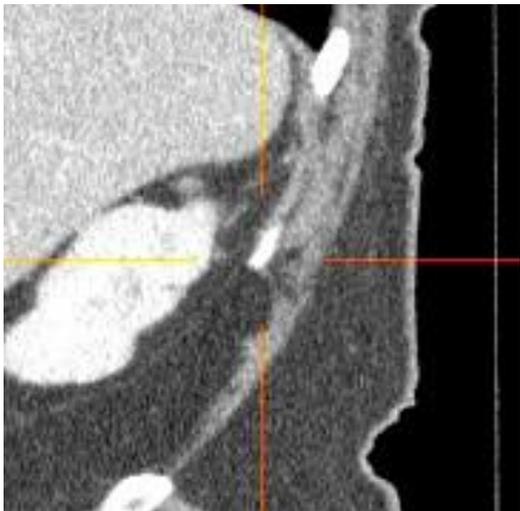
恥骨結合の骨の間隙で、(矢状断で見て)恥骨結合の真ん中の点。典型的には矢状断で骨の途切れるスライスで選ぶ。骨皮質のボクセルを選ばないように注意。

## 尾骨先端 [Coccyx\_Tip]

	AX	COR
AX		
SAG		

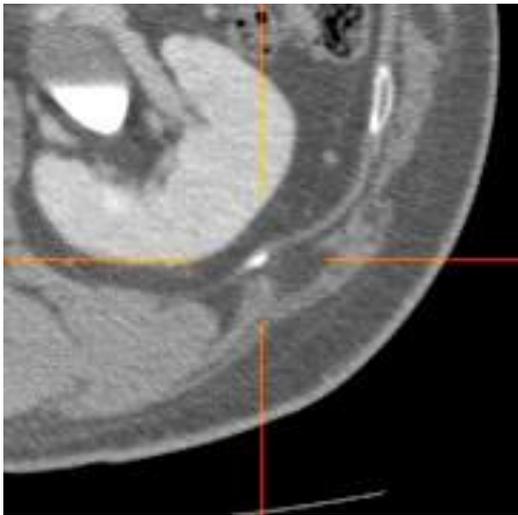
尾骨の、CTで見える限りもっとも尾側の骨の先端。

## 右最下位肋骨先端 [R\_12thRib\_Tip]

	AX	COR
AX		
SAG		

右最下位肋骨(肋骨の数の破格がある場合でも最下位)の先端の皮質上の点。断面方向では中心をとる。軟骨は(骨化があったとしても)入れないこと。

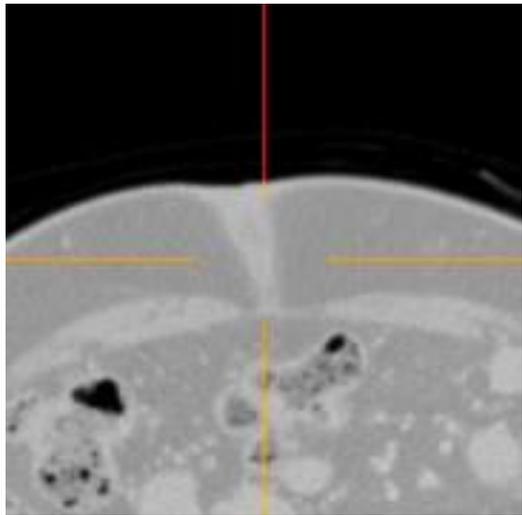
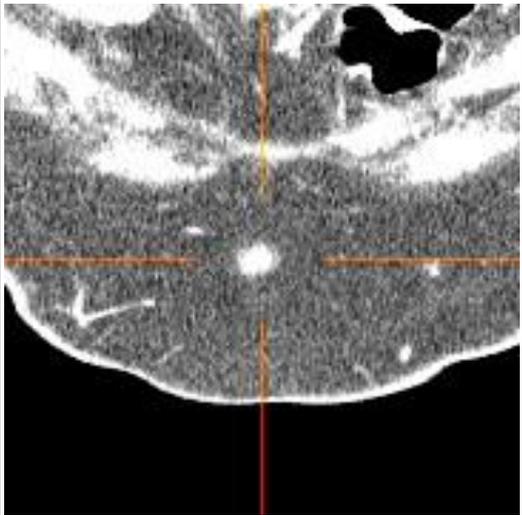
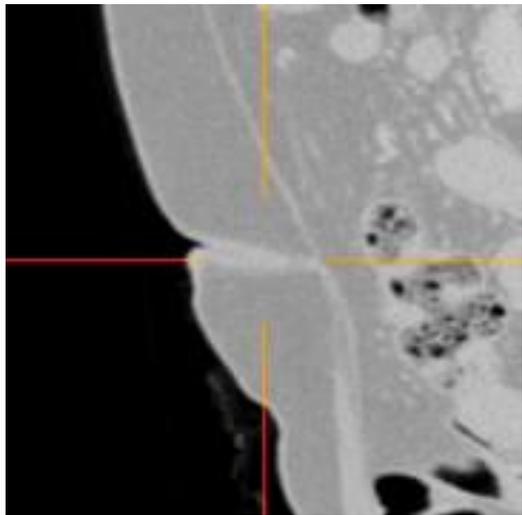
左最下位肋骨先端 [L\_12thRib\_Tip]

	AX	COR
AX		
SAG		

L\_Rib-1\_Tip

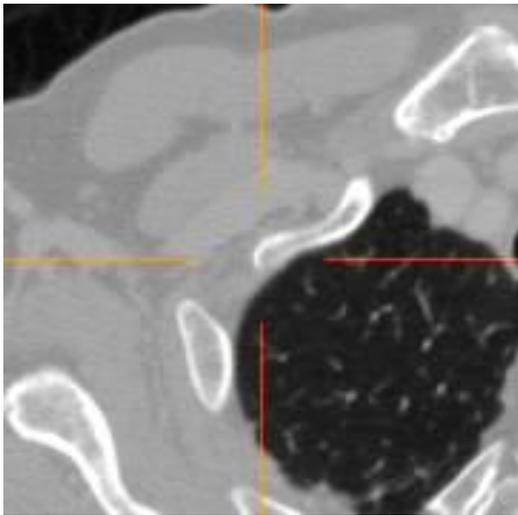
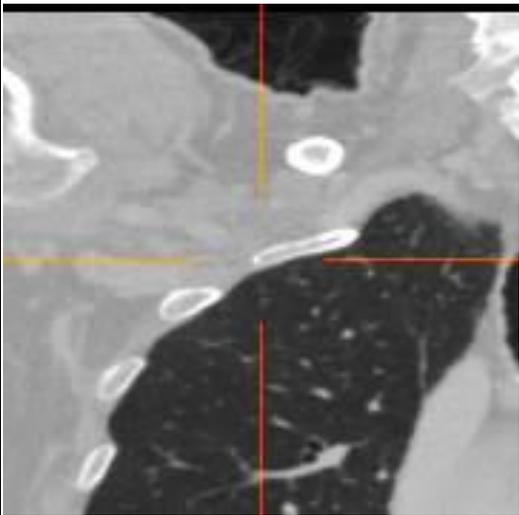
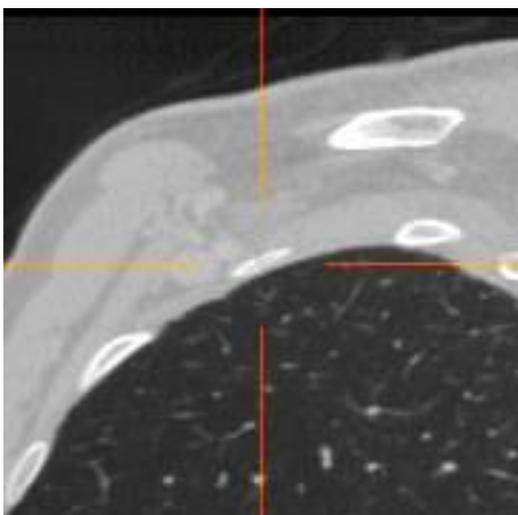
## Appendix B Anatomical landmark list

## 臍 [Umbilicus]

	AX	COR
AX		
SAG		

臍。皮下脂肪の厚みの中央の深さあたりをとること。ただし空気濃度のボクセルは避ける(へこんだ臍では、必要に応じてより深い点をとってよい)。

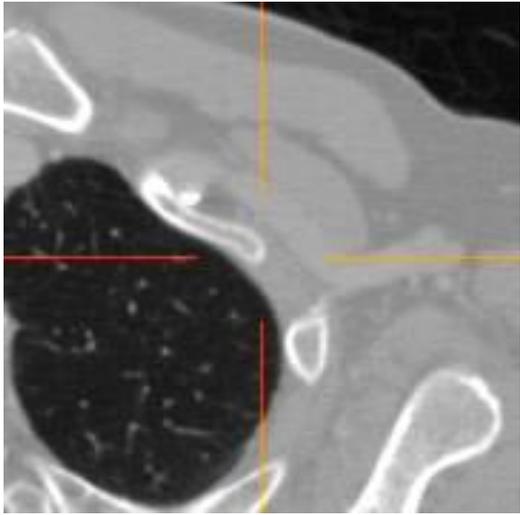
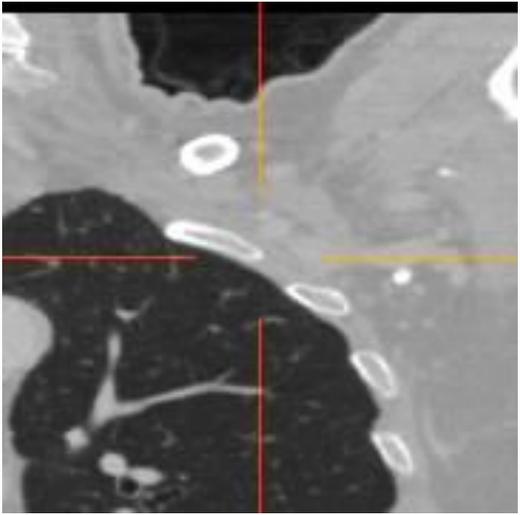
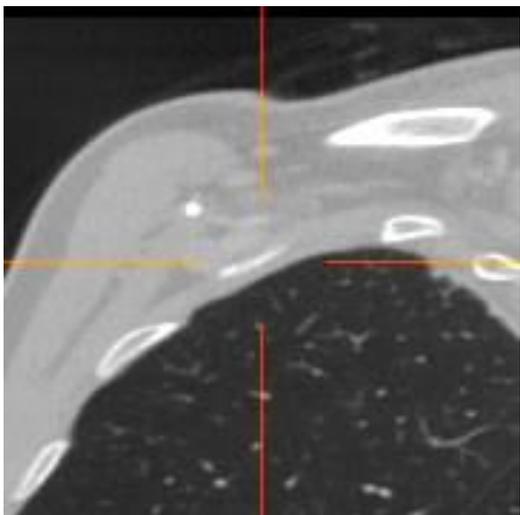
## 右第1肋骨外側端 [R\_1stRib\_LatMargin]

	AX	COR
AX		
SAG		

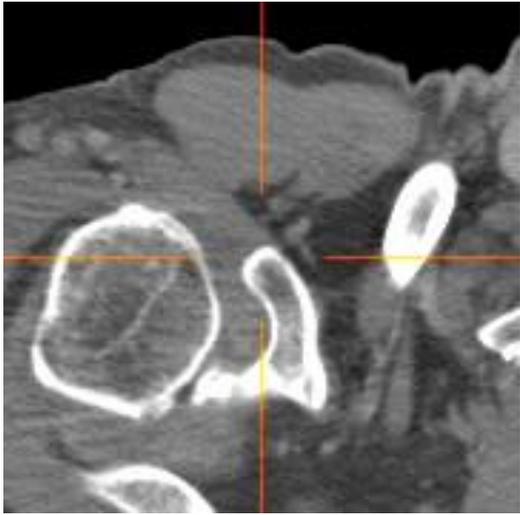
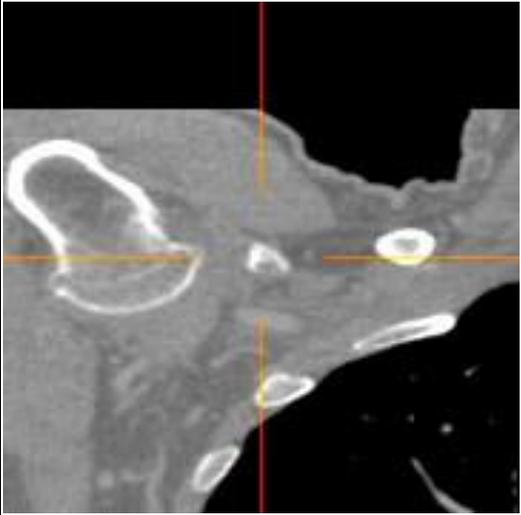
矢状断で外側に向かって見ていって、骨皮質のもっとも外側のところのやや腹側。第1肋骨が比較的急角に折れ曲がるあたり。

Appendix B Anatomical landmark list

左第 1 肋骨外側端 [L\_1stRib\_LatMargin]

	AX	COR
AX		
SAG		

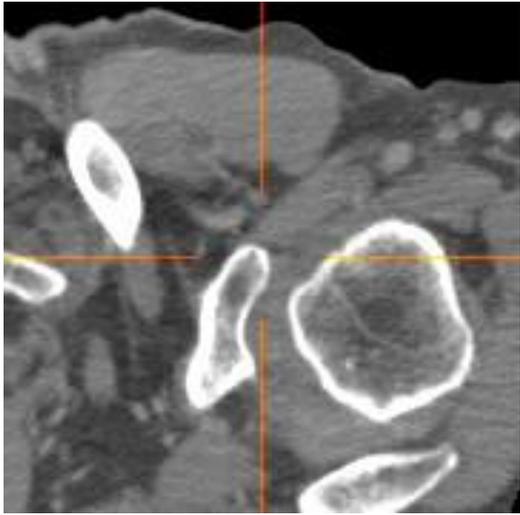
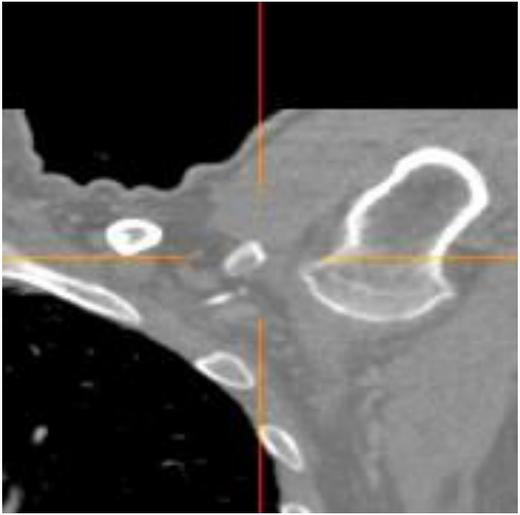
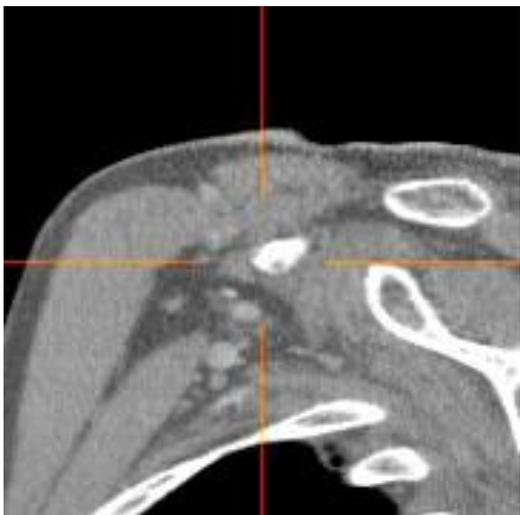
## 右鳥口突起 [R\_CoracoidP] (rt. coracoid process)

	AX	COR
AX		
SAG		

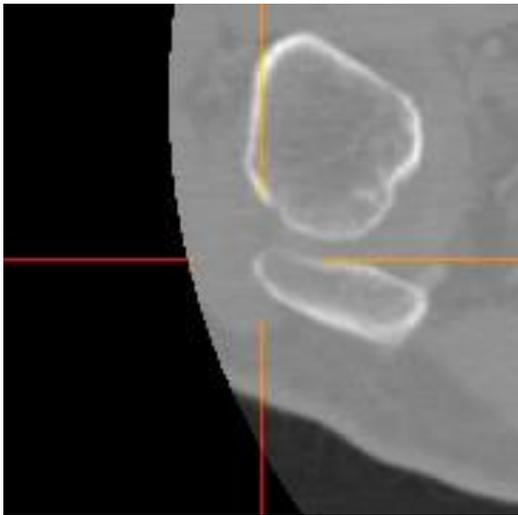
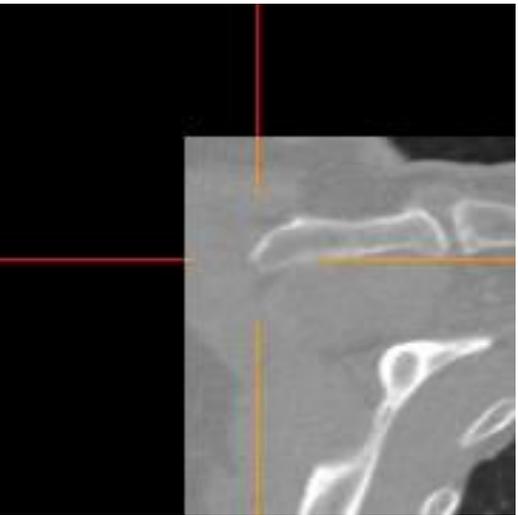
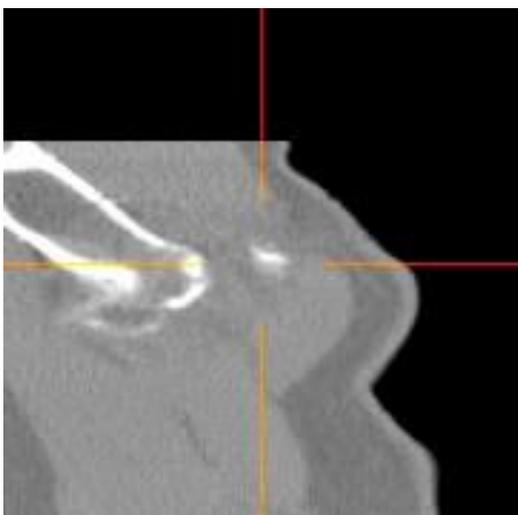
鳥口突起の先端で、突起の輪切り方向では中心となるあたりの骨皮質上。

## Appendix B Anatomical landmark list

## 左鳥口突起 [L\_CoracoidP] (lt. coracoid process)

	AX	COR
AX		
SAG		

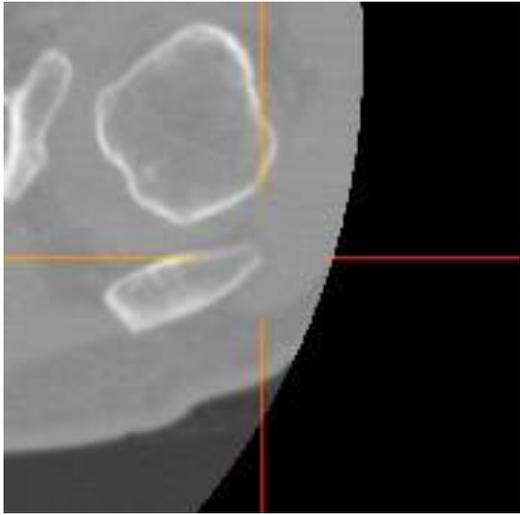
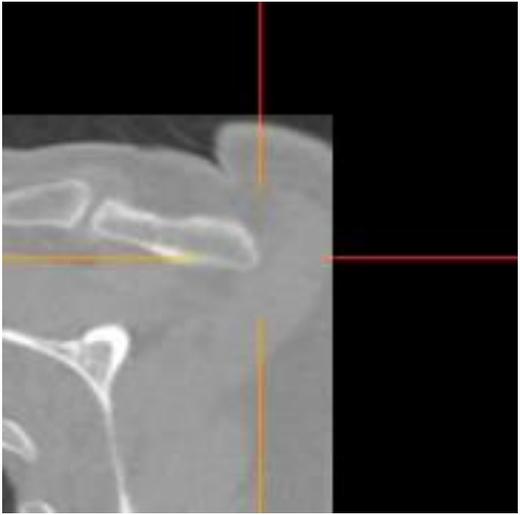
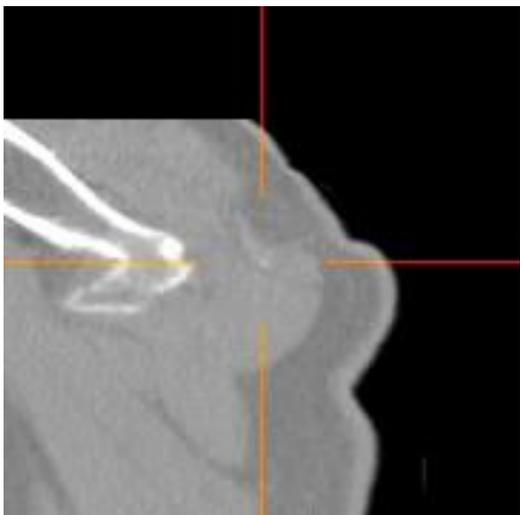
## 右肩峰 [R\_Acromion]

	AX	COR
AX		
SAG		

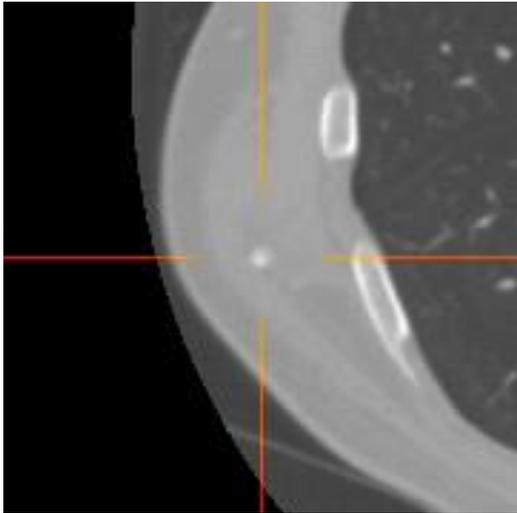
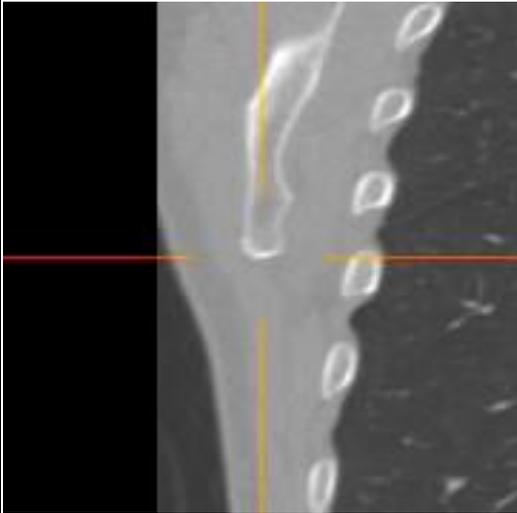
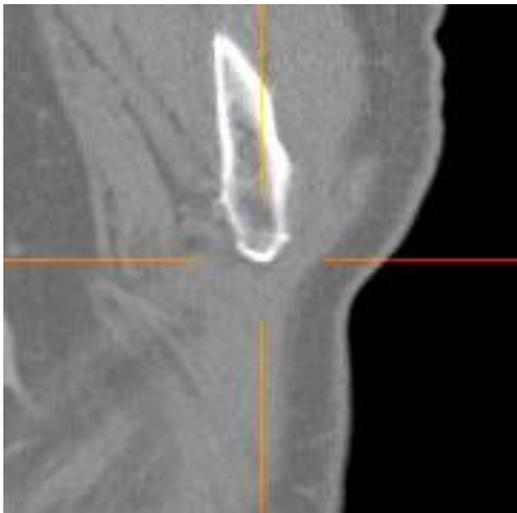
肩峰の先端。矢状断では(手下げ画像では)肩峰の幅のなかでやや前方よりとなる。

## Appendix B Anatomical landmark list

## 左肩峰 [L\_Acromion]

	AX	COR
AX		
SAG		

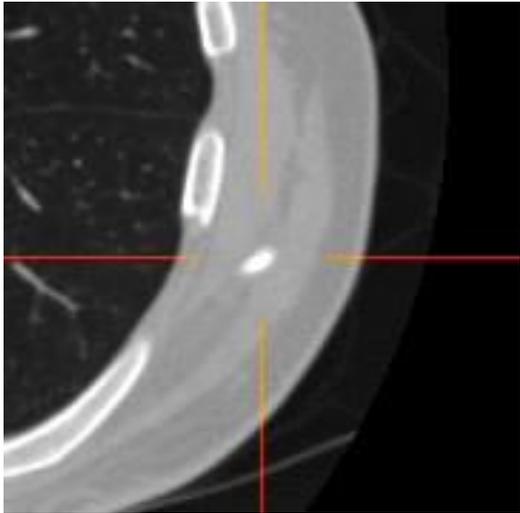
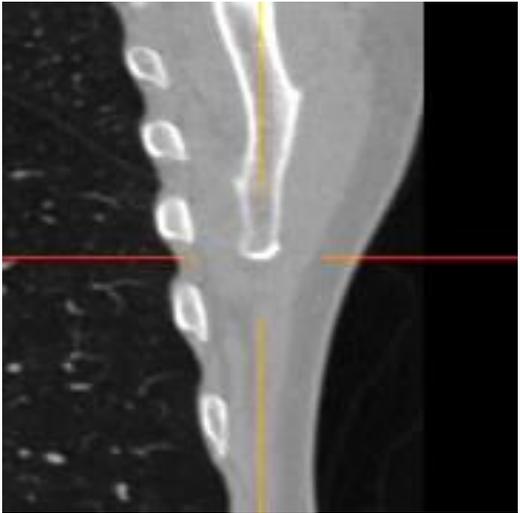
## 右肩甲骨下端 [R\_Scapula\_InfTip] (inferior tip of rt. scapula)

	AX	COR
AX		
SAG		

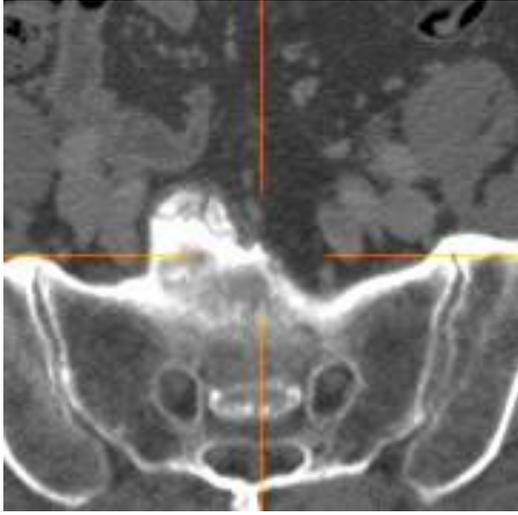
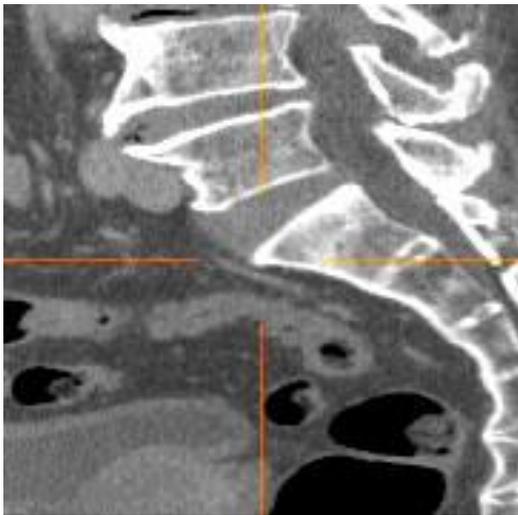
肩甲骨の板状部の下端。

## Appendix B Anatomical landmark list

左肩甲骨下端 [L\_Scapula\_InfTip] (inferior tip of lt. scapula)

	AX	COR
AX		
SAG		

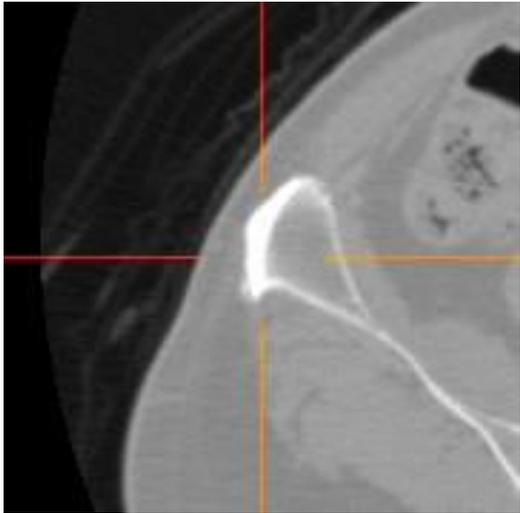
## 仙骨上前端 [Sacrum\_AntSupTip] (anteriosuperior tip of sacrum)

	AX	COR
AX		
SAG		

岬角正中。ひどい骨棘は無視してよいが、できるだけ骨皮質に一致するように。

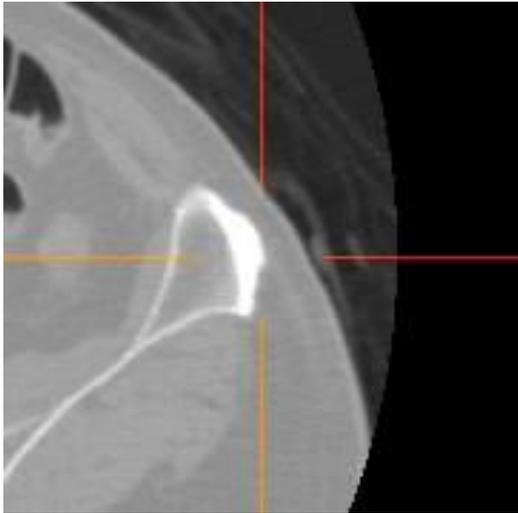
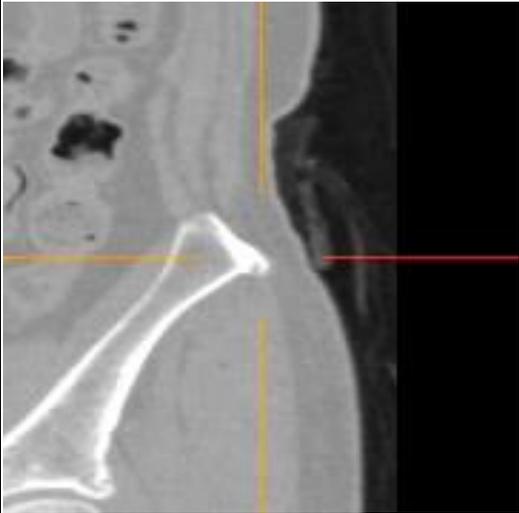
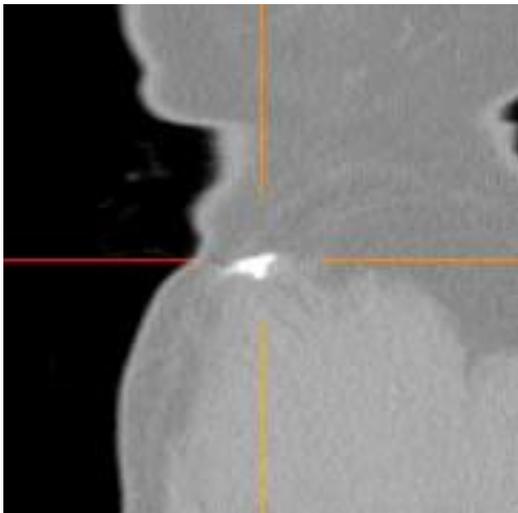
## Appendix B Anatomical landmark list

## 右腸骨稜外側縁 [R\_IliacCrest\_Lat] (lateral margin of rt. iliac crest)

	AX	COR
AX		
SAG		

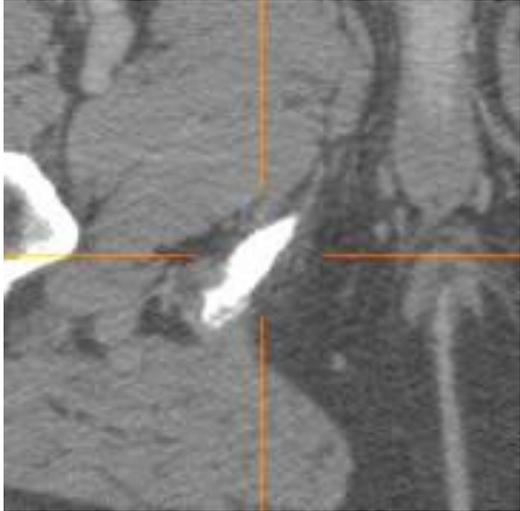
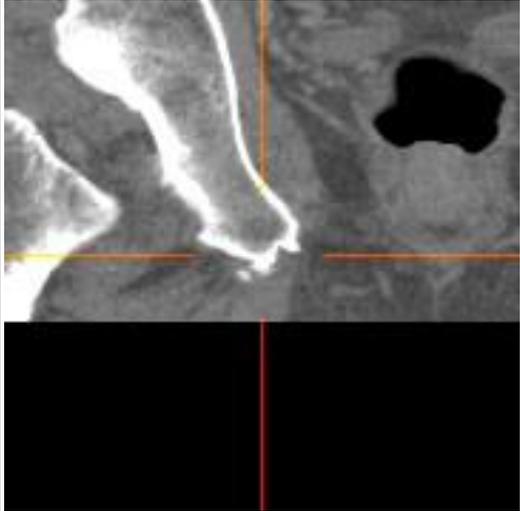
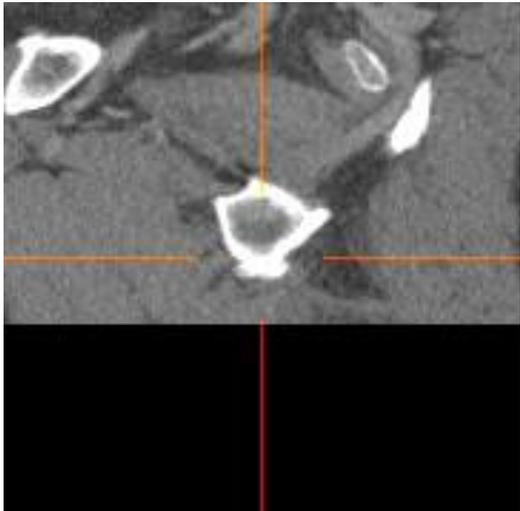
腸骨稜の外側縁で、縫工筋の起始付着部の骨皮質の厚くなっているところ。

## 左腸骨稜外側縁 [L\_IliacCrest\_Lat] (lateral margin of lt. iliac crest)

	AX	COR
AX		
SAG		

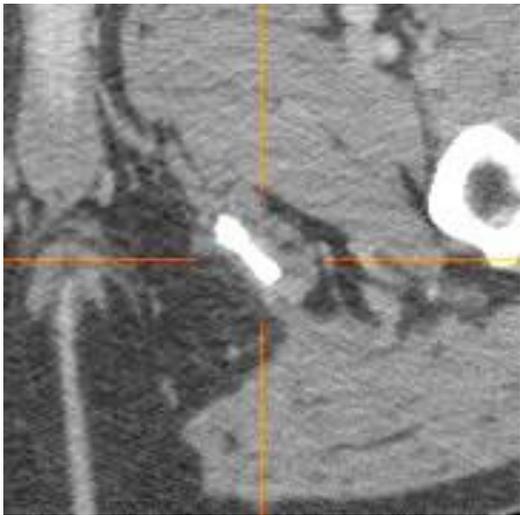
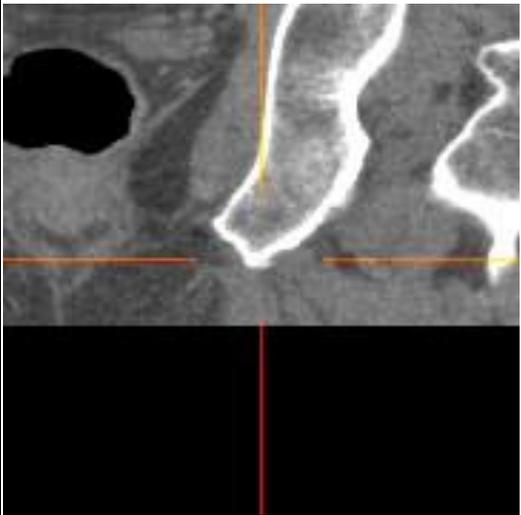
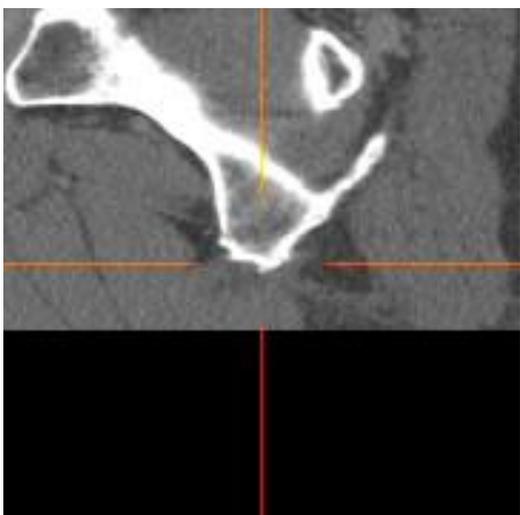
## Appendix B Anatomical landmark list

## 右坐骨下端 [R\_Ischium\_InfMargin] (inferior margin of rt. ischium)

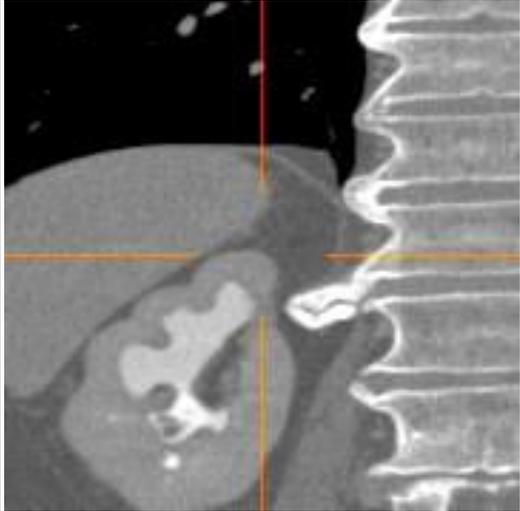
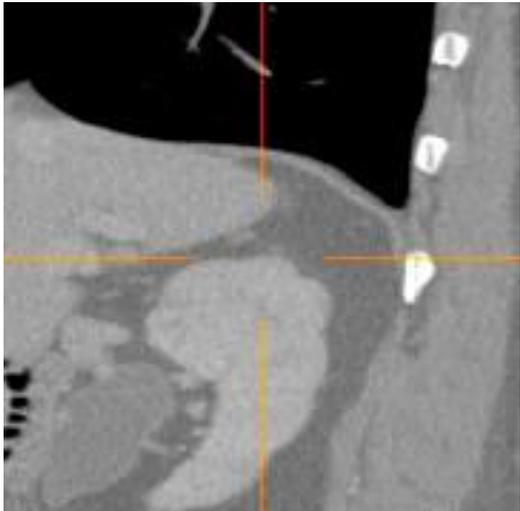
	AX	COR
AX		
SAG		

水平断でスクロールしたときの坐骨の下端。

左坐骨下端 [L\_Ischium\_InfMargin] (inferior margin of lt. ischium)

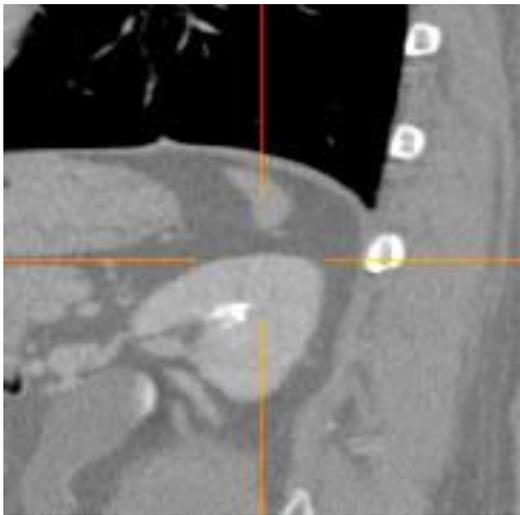
	AX	COR
AX		
SAG		

## 右腎上極 [R\_Kidney\_SupTip]

	AX	COR
AX		
SAG		

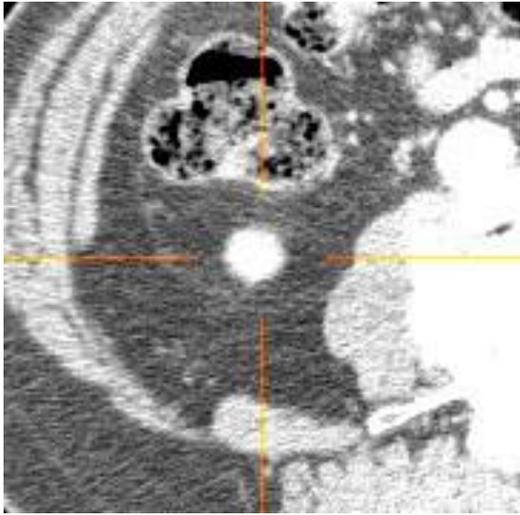
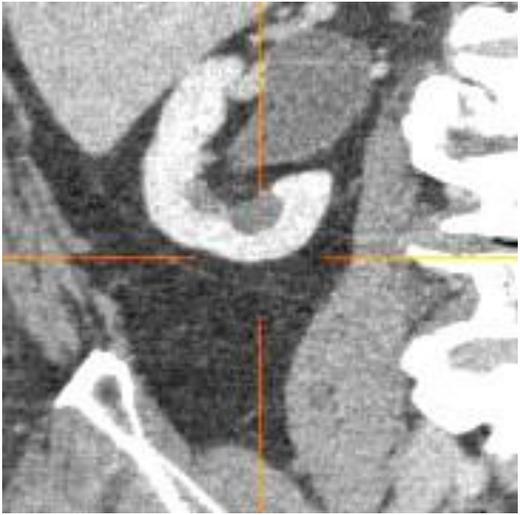
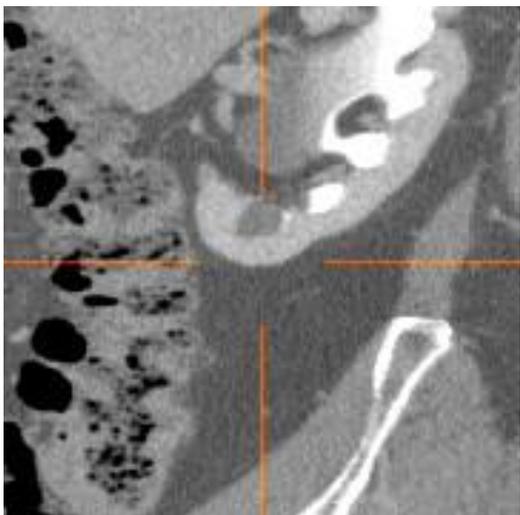
右腎の上極(もっとも頭側)の点。

左腎上極 [L\_Kidney\_SupTip]

	AX	COR
AX		
SAG		

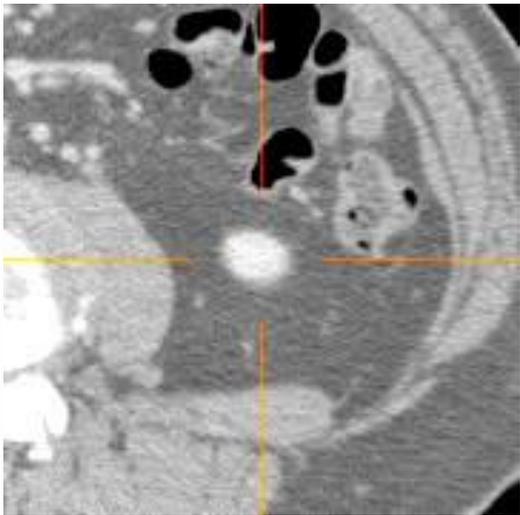
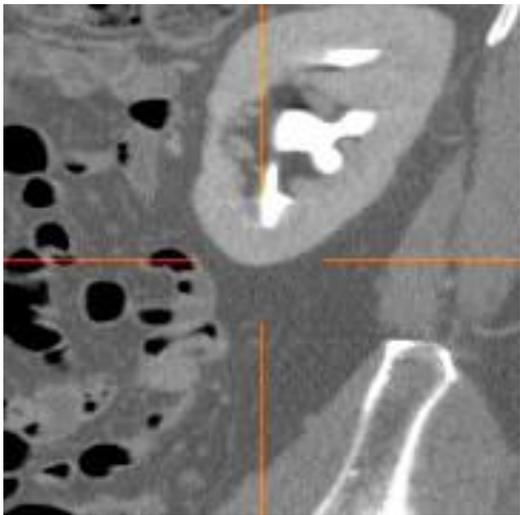
## Appendix B Anatomical landmark list

## 右腎下極 [R\_kidney\_InfTip]

	AX	COR
AX		
SAG		

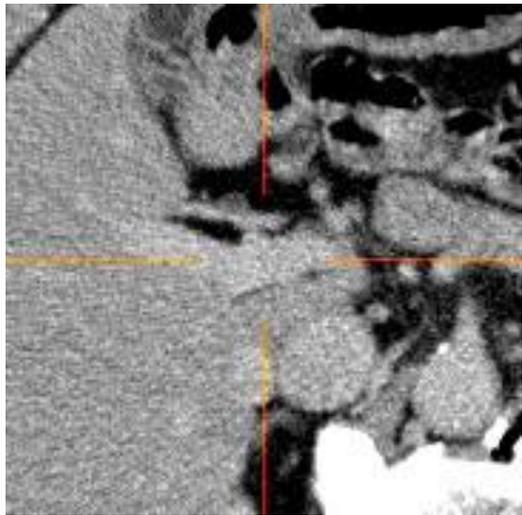
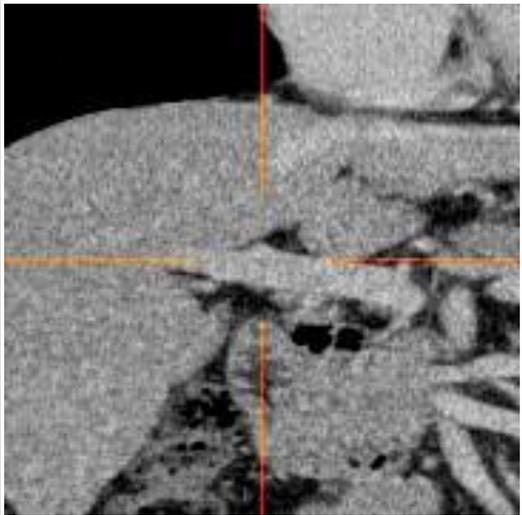
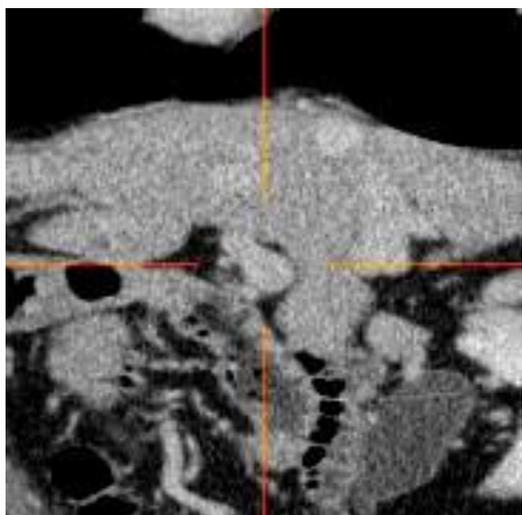
右腎の下極のボクセル。

左腎下極 [L\_Kidney\_InfTip]

	AX	COR
AX		
SAG		

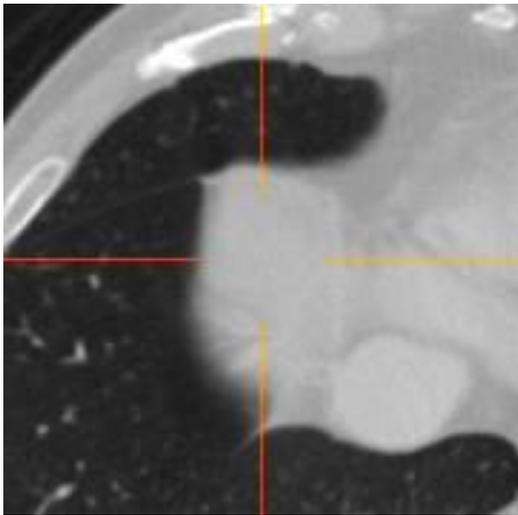
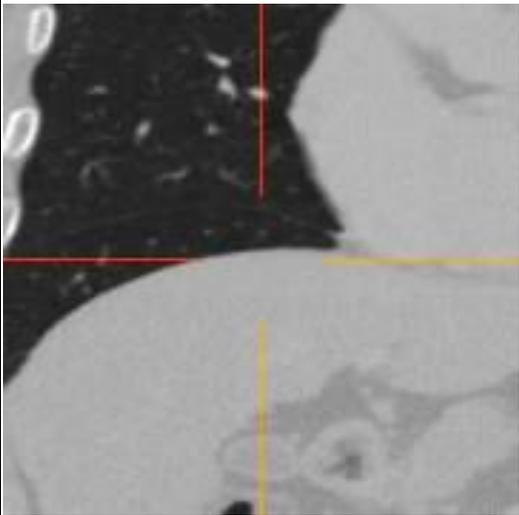
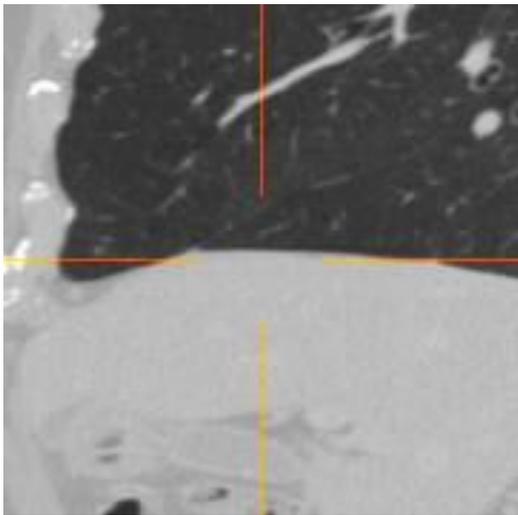
## Appendix B Anatomical landmark list

## 門脈左右枝分枝部 [IHPV\_Bifur] (bifurcation of intrahepatic portal vein)

	AX	COR
AX		
SAG		

肝内門脈の右枝と左枝の分枝部の又のところ。見えにくいときは心眼でとるか、相談を。

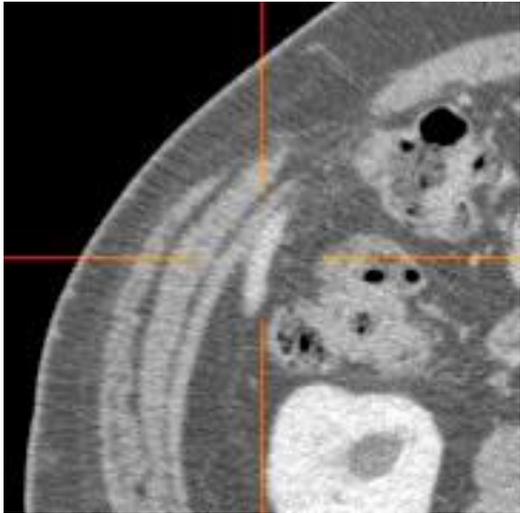
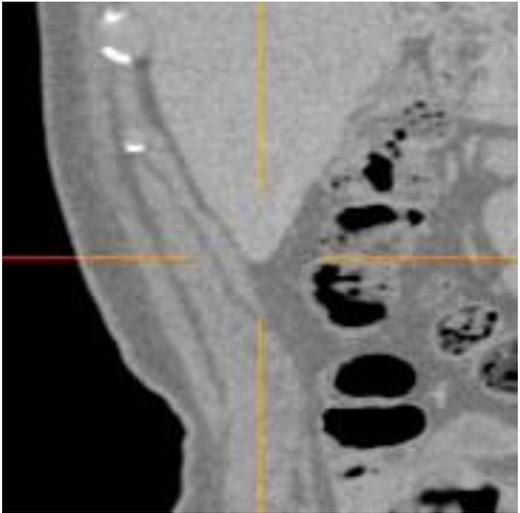
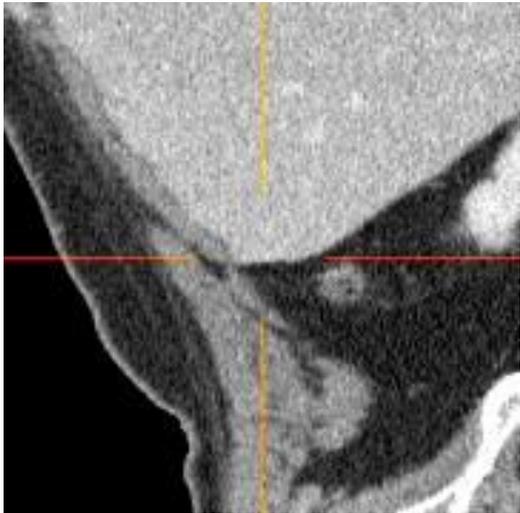
## 肝右葉上端 [Liver\_R\_Lobe\_Sup] (superior margin of rt. lobe of tliver)

	AX	COR
AX		
SAG		

肝右葉の最上端、横隔膜のドーム頂点。

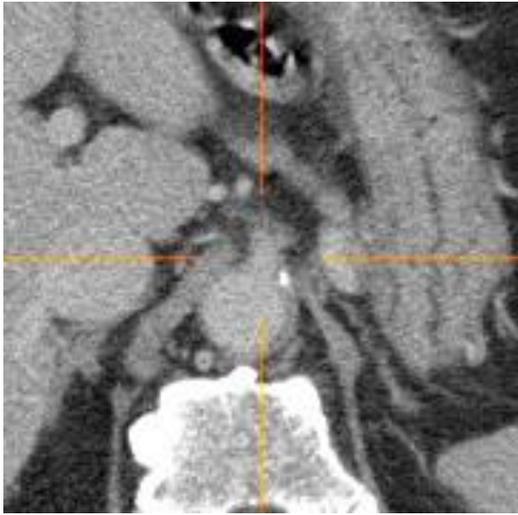
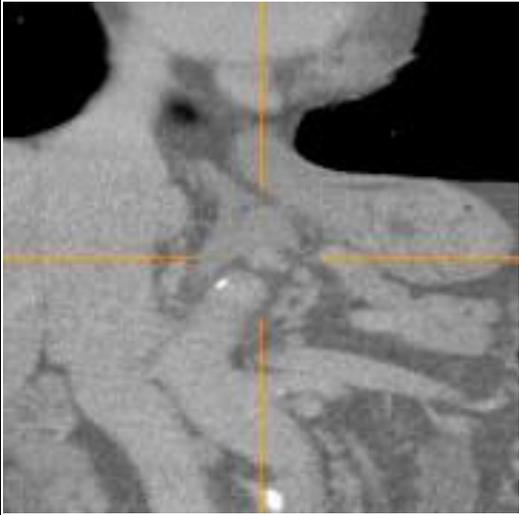
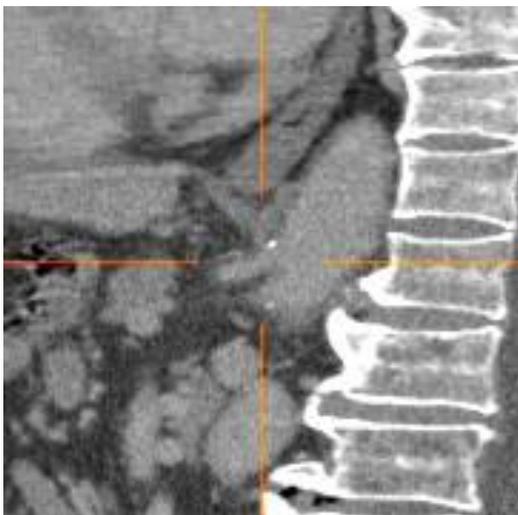
## Appendix B Anatomical landmark list

## 肝下端 [Liver\_InfTip]

	AX	COR
AX		
SAG		

肝右葉後区域の下端。

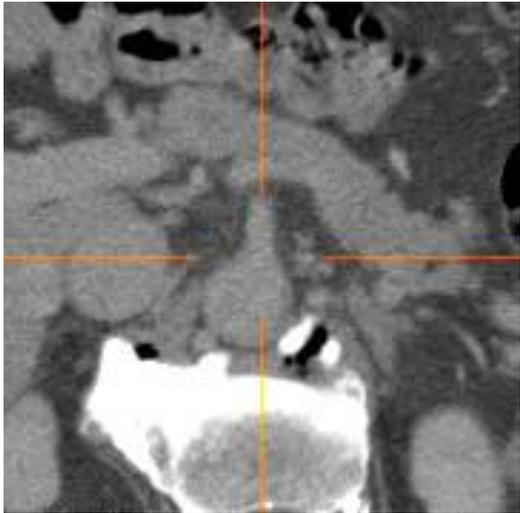
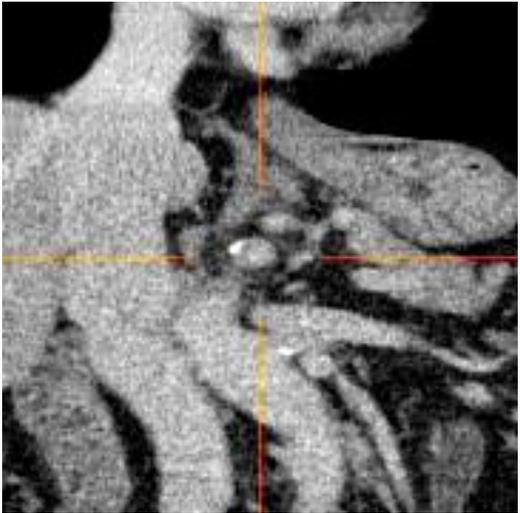
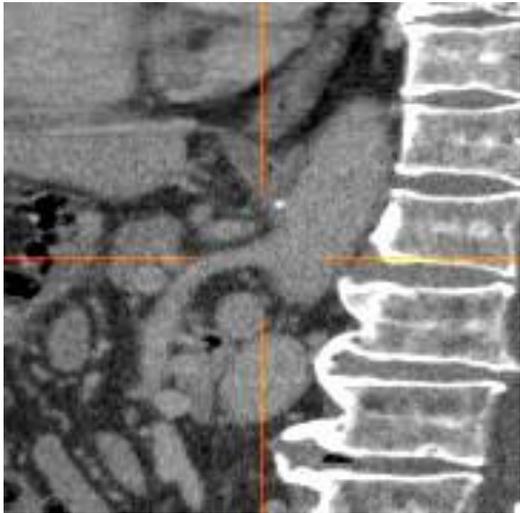
## 腹腔動脈起始部 [CA\_Root] (root of celiac artery)

	AX	COR
AX		
SAG		

腹腔動脈が大動脈から起始した直後のその中心。

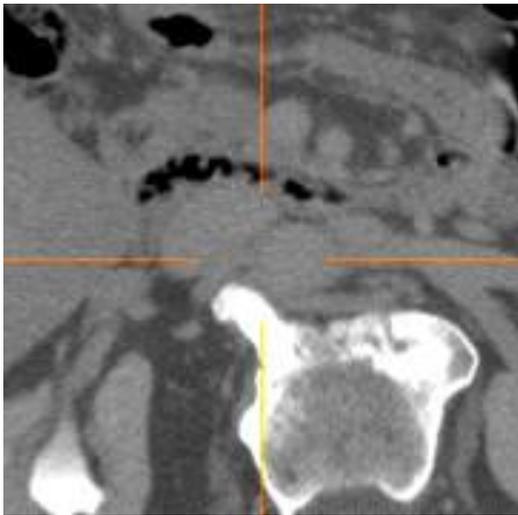
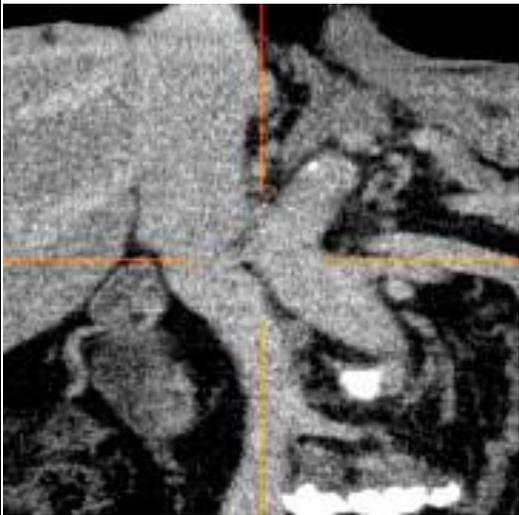
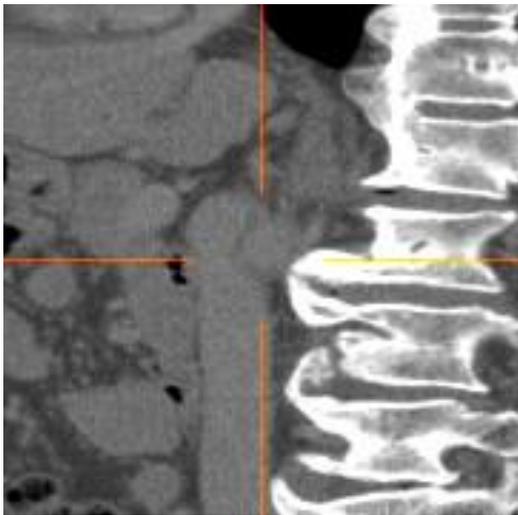
## Appendix B Anatomical landmark list

上腸間膜動脈（SMA）起始部 [SMA\_Root] (root of superior mesenteric a.)

	AX	COR
AX		
SAG		

上腸間膜動脈が大動脈から起始するところのその中心。

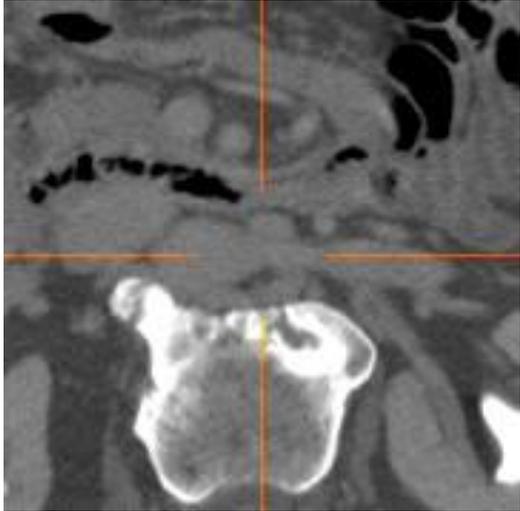
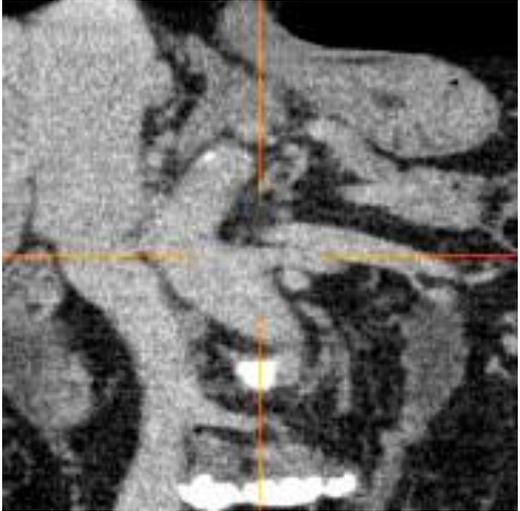
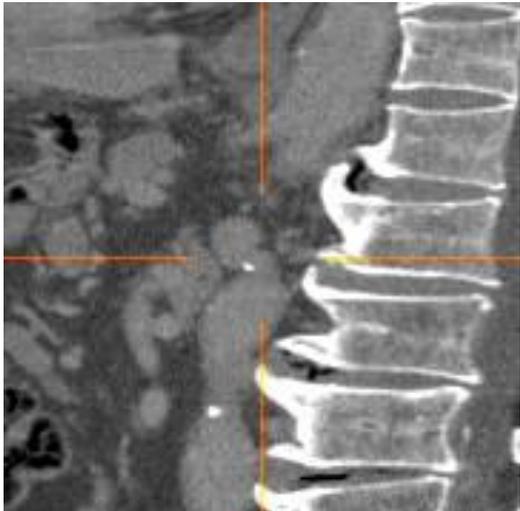
## 右腎動脈起始部 [R\_RenA\_Root] (root of rt. renal artery)

	AX	COR
AX		
SAG		

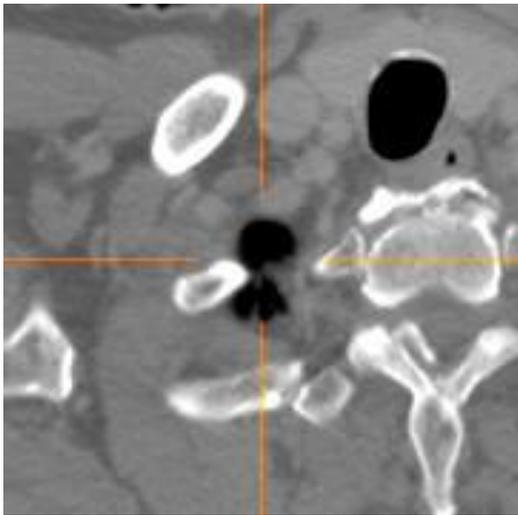
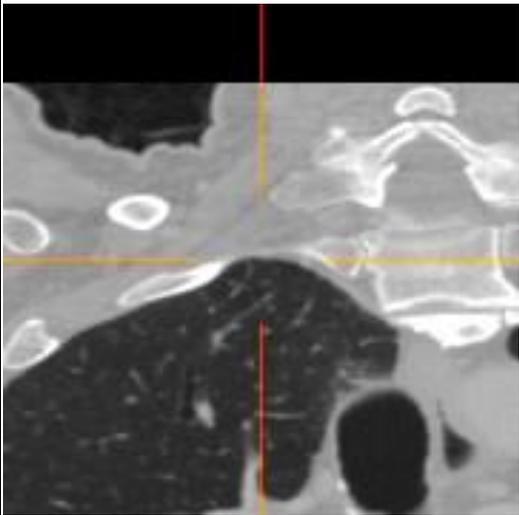
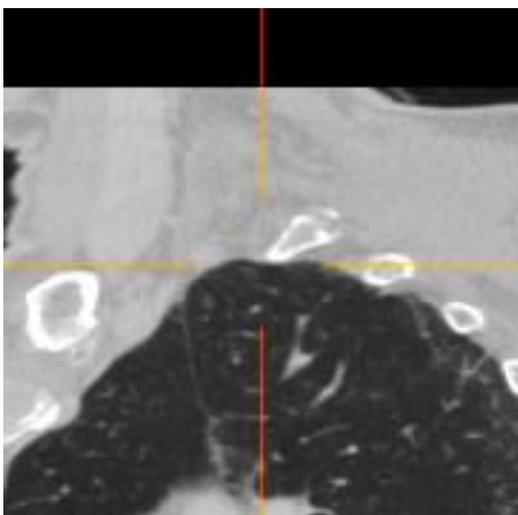
同定できないときは推測で大体のところを指定。

## Appendix B Anatomical landmark list

## 左腎動脈起始部 [L\_RenA\_Root] (root of lt. renal artery)

	AX	COR
AX		
SAG		

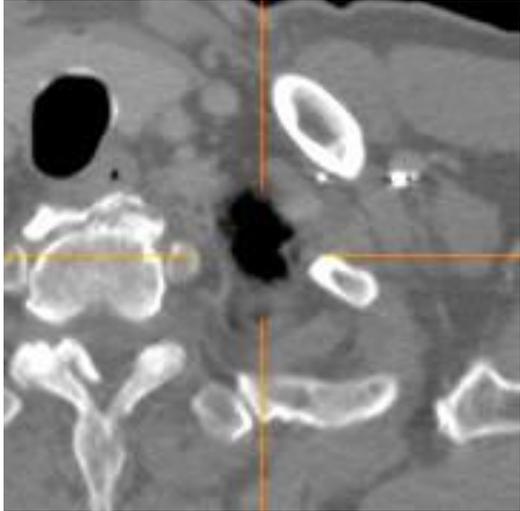
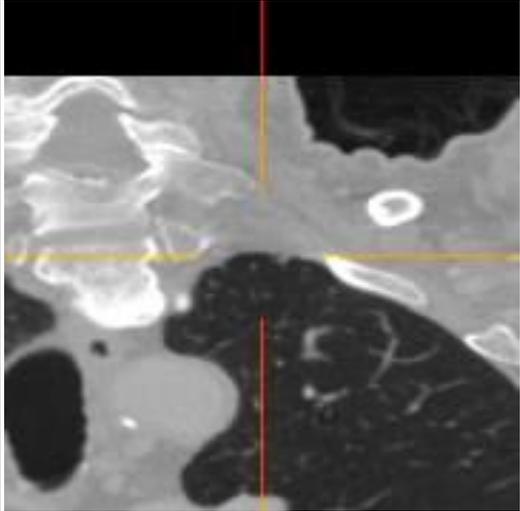
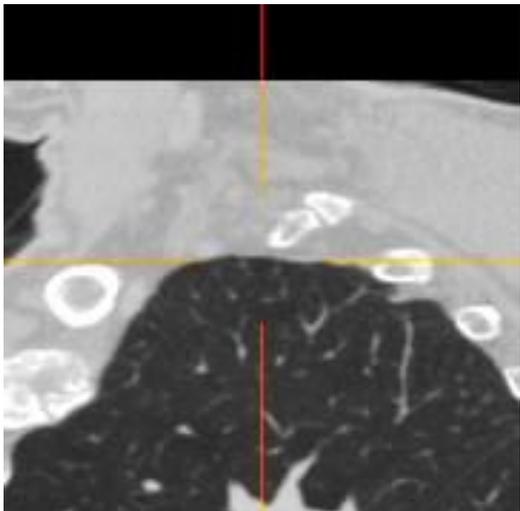
## 右肺尖 [R\_LungApex]

	AX	COR
AX		
SAG		

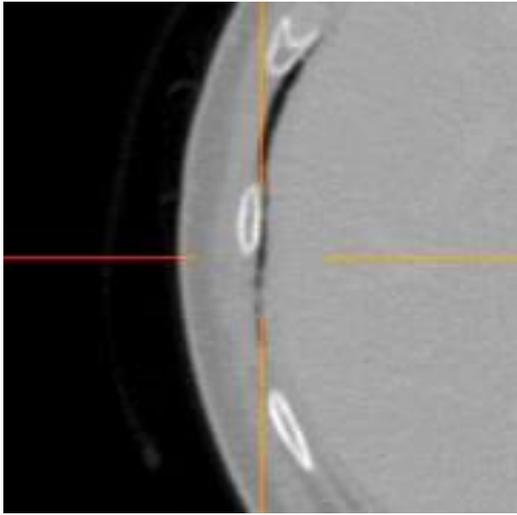
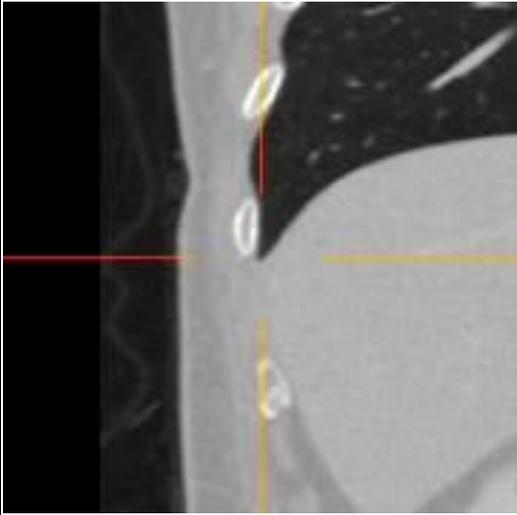
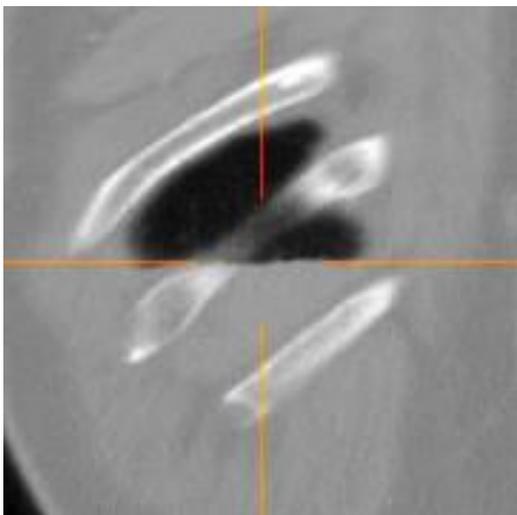
肺尖の上端。空気密度の点をとる。以下の肺の点も同じ。

## Appendix B Anatomical landmark list

## 左肺尖 [L\_LungApex]

	AX	COR
AX		
SAG		

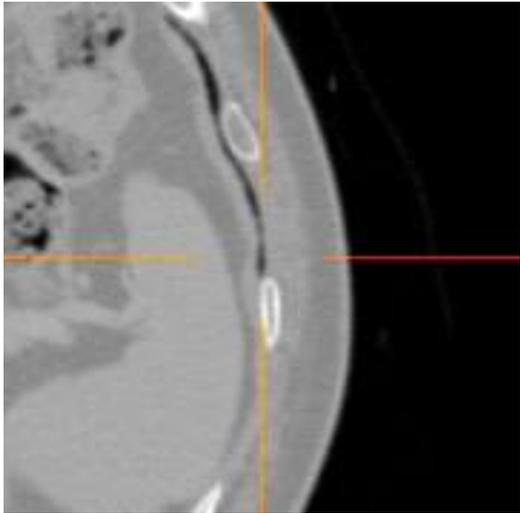
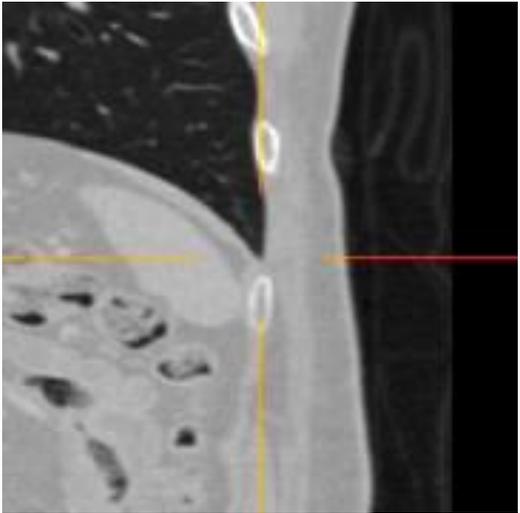
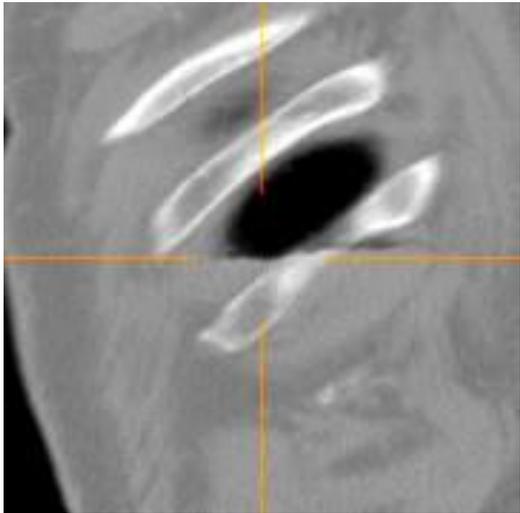
## 右肺底外側端 [R\_LungBase\_Lat]

	AX	COR
AX		
SAG		

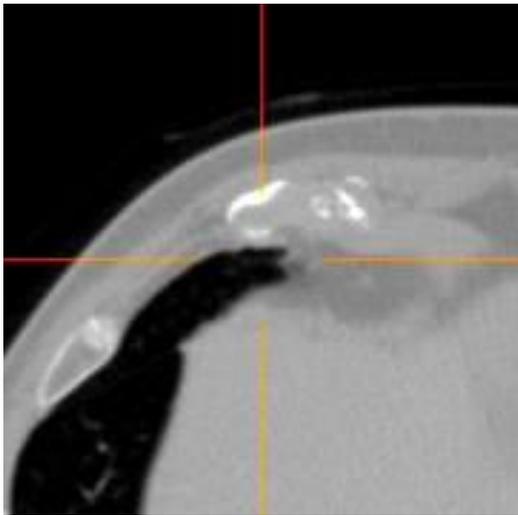
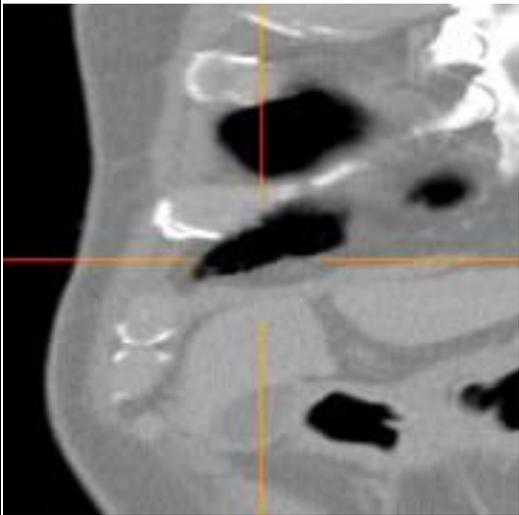
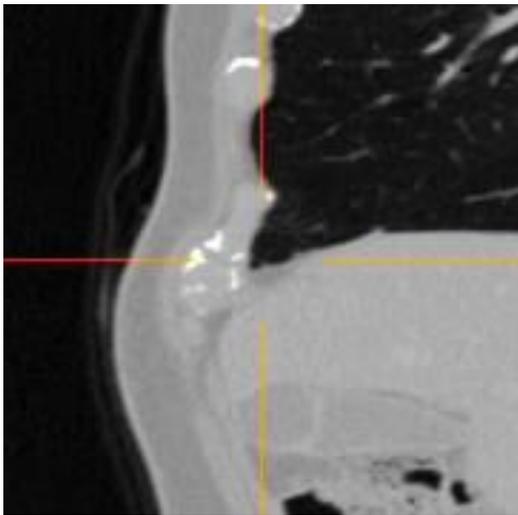
空気密度で、矢状断で最も外側尾側の肺の点をとる。

## Appendix B Anatomical landmark list

## 左肺底外側端 [L\_LungBase\_Lat]

	AX	COR
AX		
SAG		

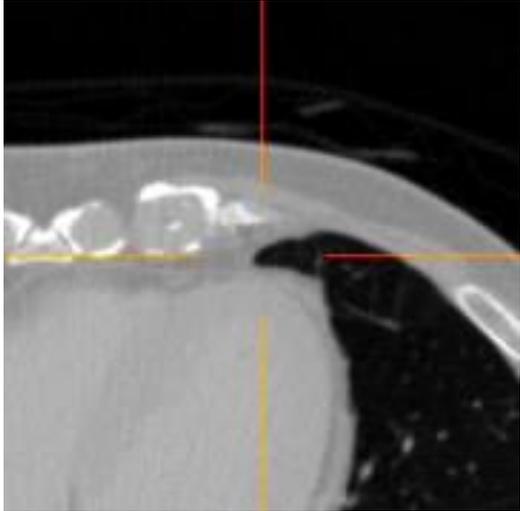
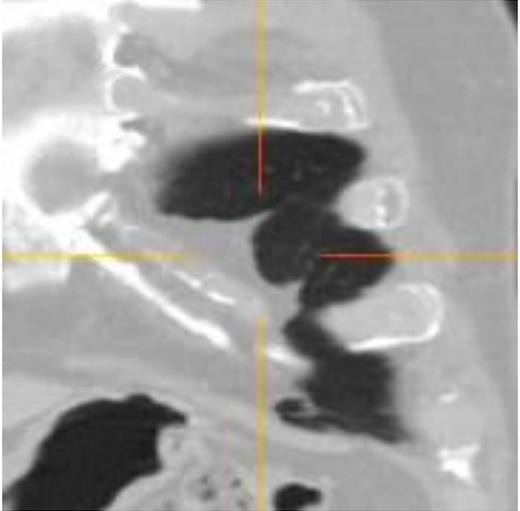
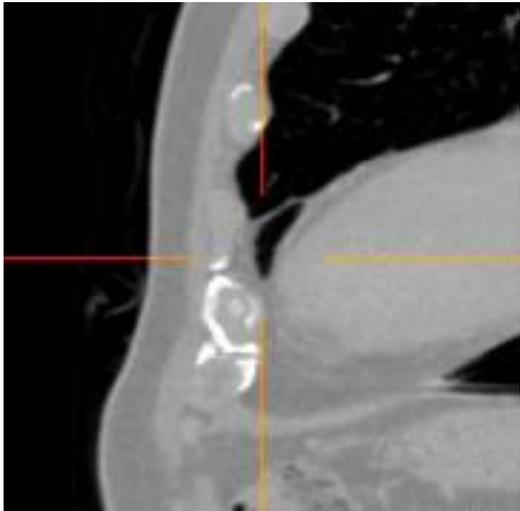
## 右肺底腹側端 [R\_LungBase\_Ant]

	AX	COR
AX		
SAG		

右肺前下縁の点。

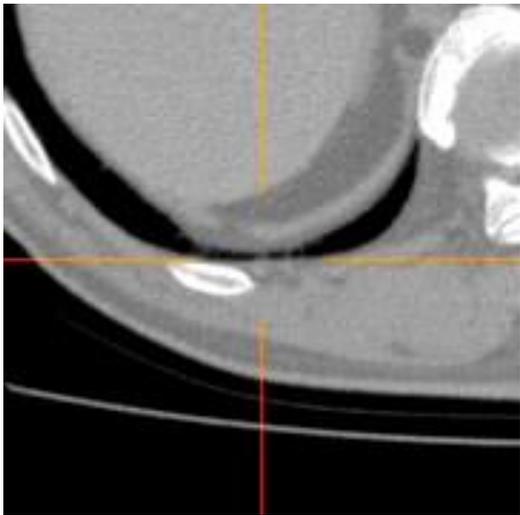
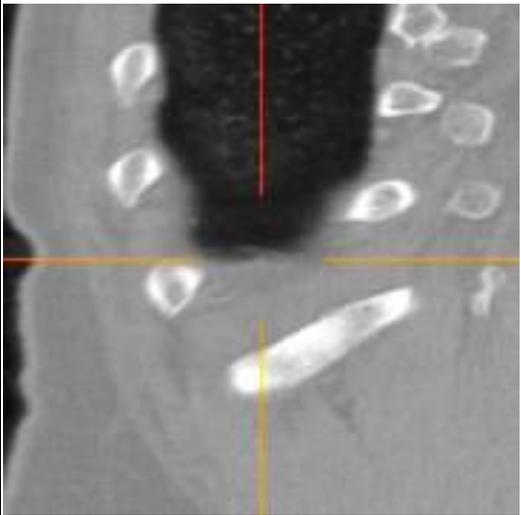
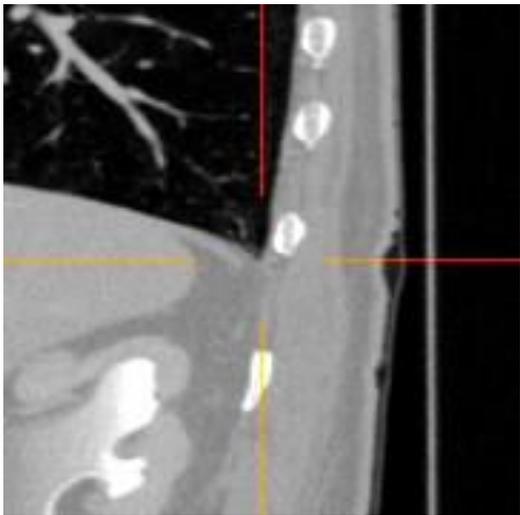
## Appendix B Anatomical landmark list

## 左肺底腹側端 [L\_LungBase\_Ant]

	AX	COR
AX		
SAG		

左の心横隔膜角で、左心前面の肺下端。一点に決めにくいですが、左右方向の位置に迷ったら、左心の中央になるくらいを目安に決める。

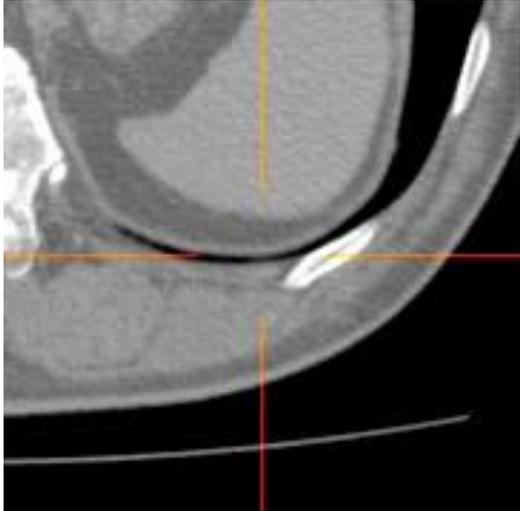
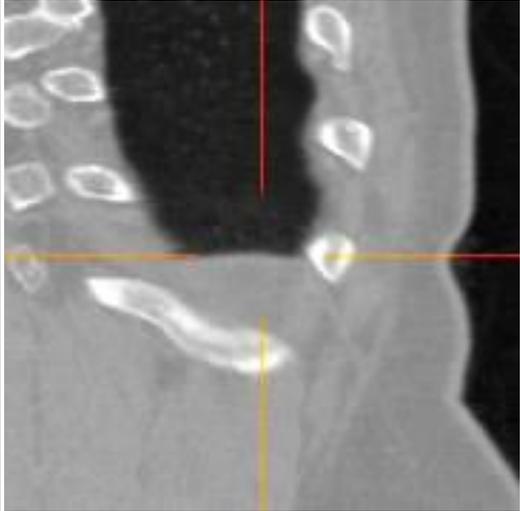
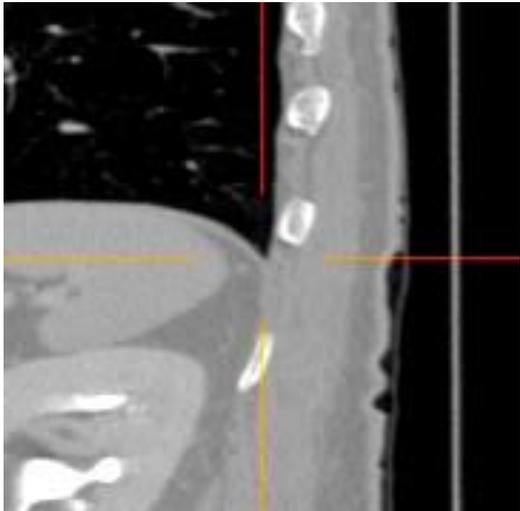
## 右肺底背側端 [R\_LungBase\_Post]

	AX	COR
AX		
SAG		

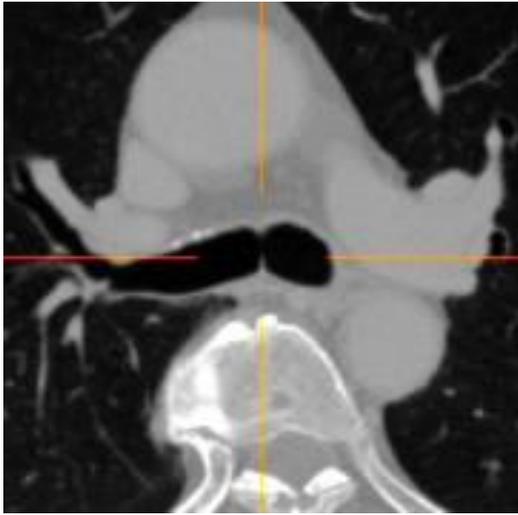
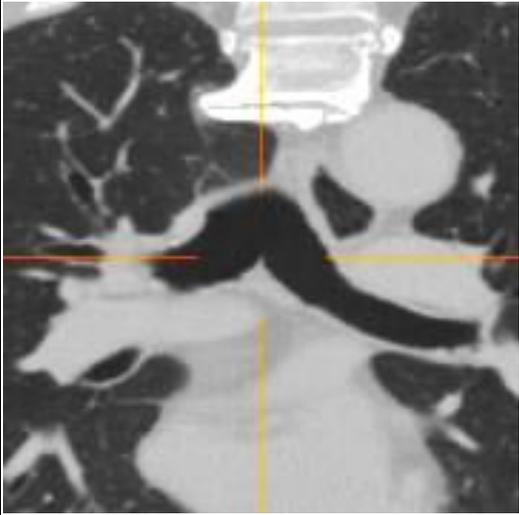
最も尾側の右肺の点。

## Appendix B Anatomical landmark list

## 左肺底背側端 [L\_LungBase\_Post]

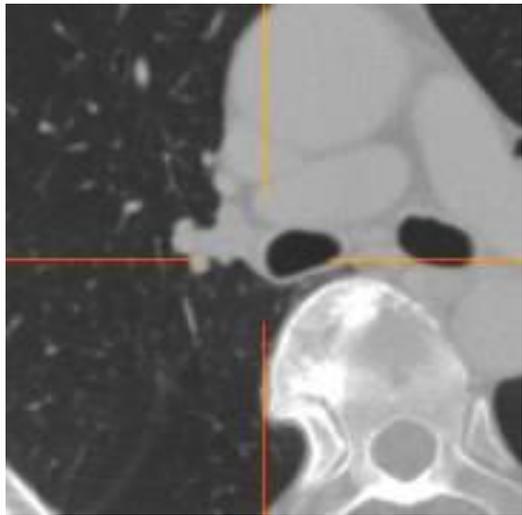
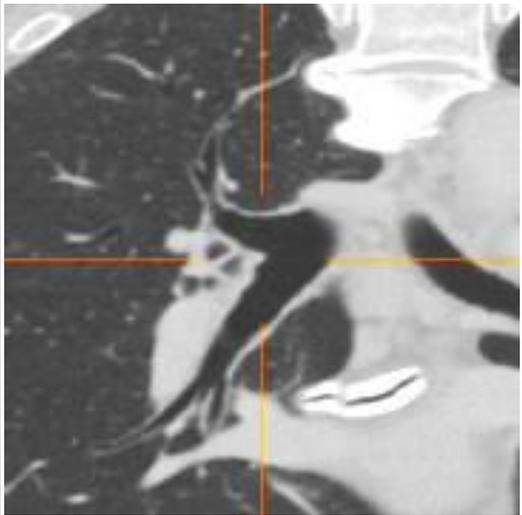
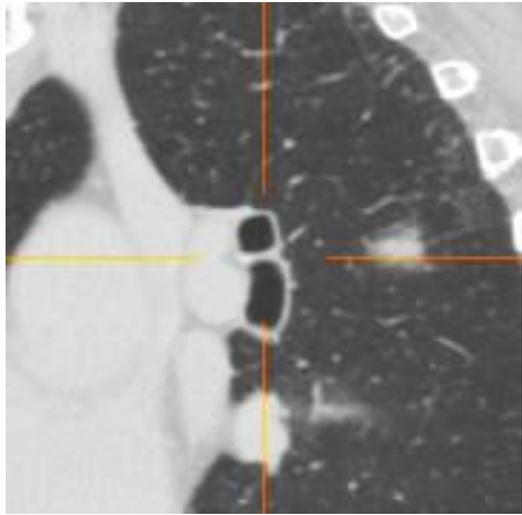
	AX	COR
AX		
SAG		

## 気管分枝部 [Trachea\_Bifur] (bifurcation of trachea)

	AX	COR
AX		
SAG		

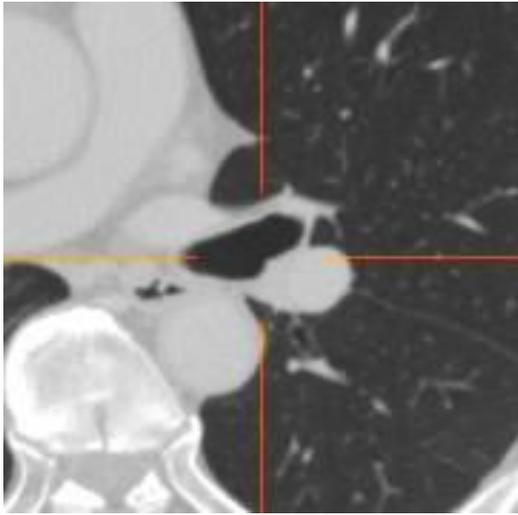
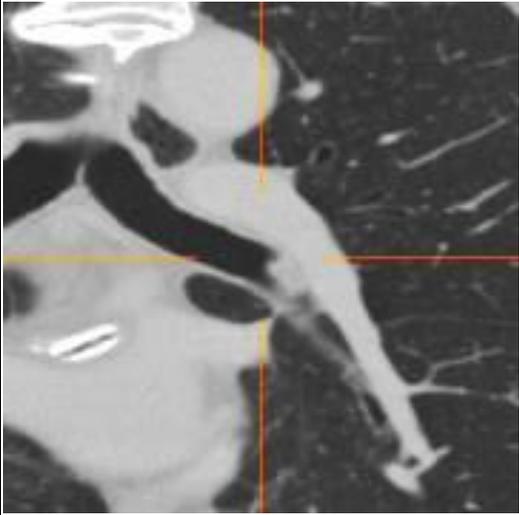
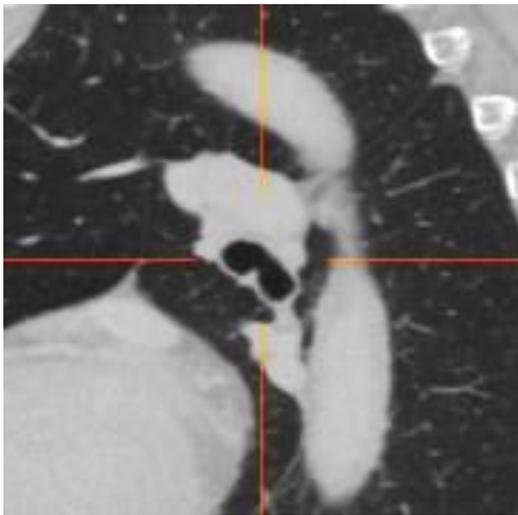
気管分枝部の股間にあたる点。冠状断でみて、逆 Y 字の Y のまたのところの鞍点をとる。空気濃度のボクセルは避ける。

## 右気管支初回分枝部 [R\_Bronchus\_Bifur] (bifurcation of rt. main bronchus)

	AX	COR
AX		
SAG		

矢状断でスクロールして、右気管支が上葉枝と中間幹に分かれるところの股間の鞍点をとる。

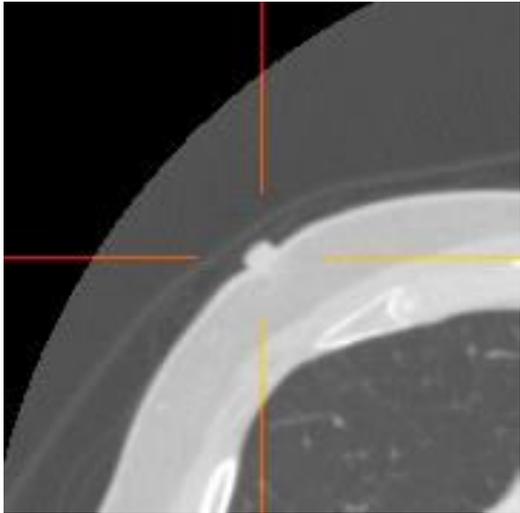
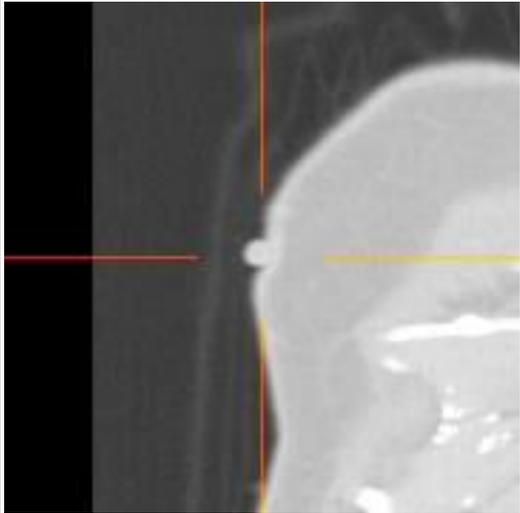
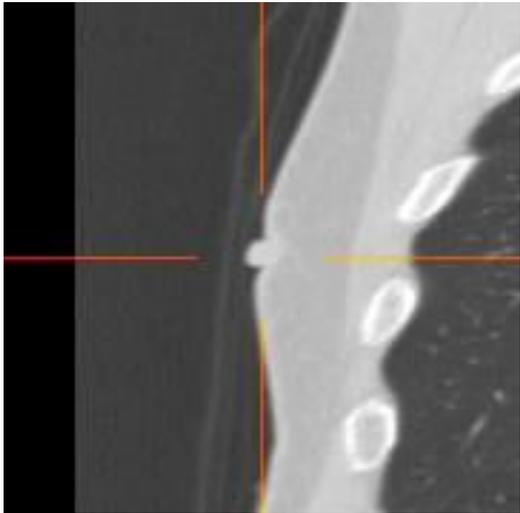
## 左気管支初回分枝部 [L\_Bronchus\_Bifur] (bifurcation of lt. main bronchus)

	AX	COR
AX		
SAG		

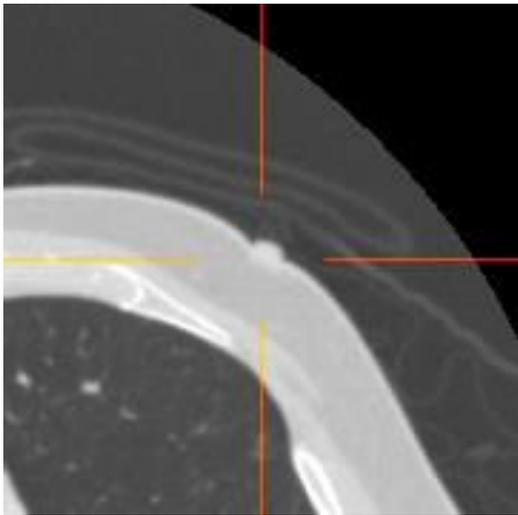
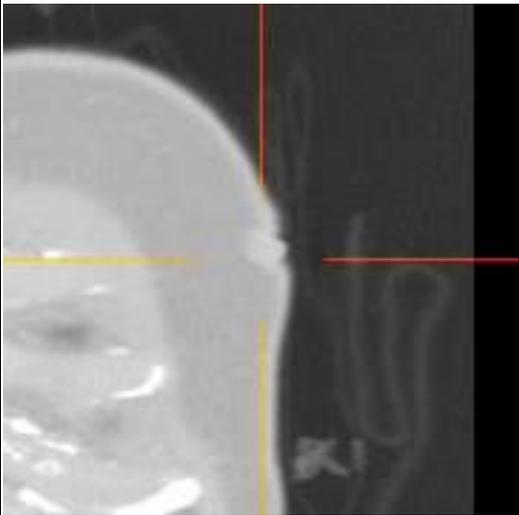
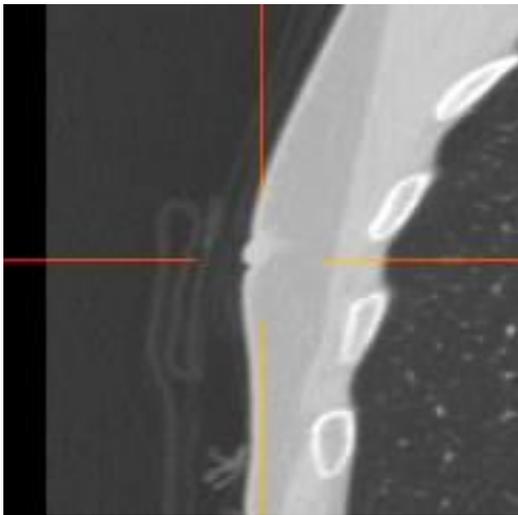
同じく矢状断でみて、左上・下葉気管支分枝部の股間の鞍点。

## Appendix B Anatomical landmark list

## 右乳頭 [R\_Nipple]

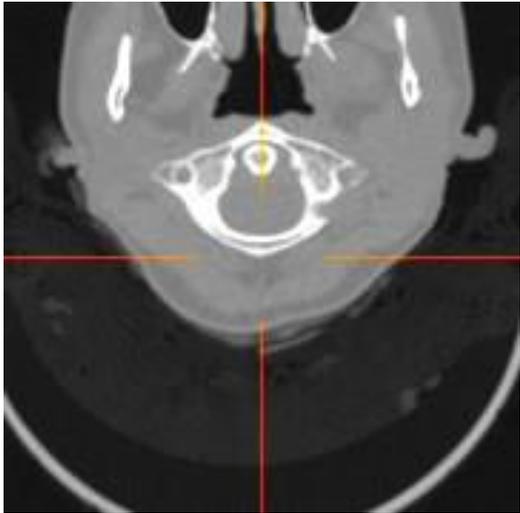
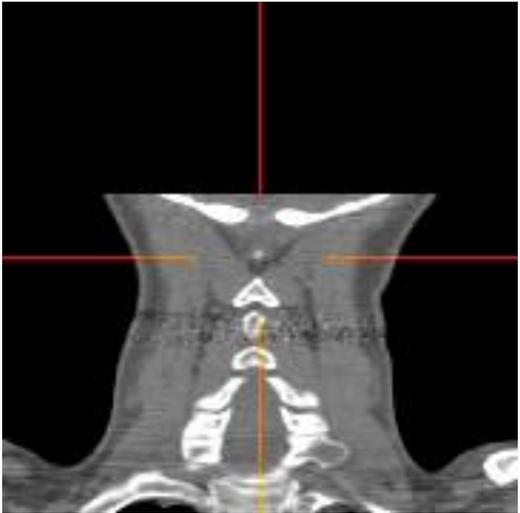
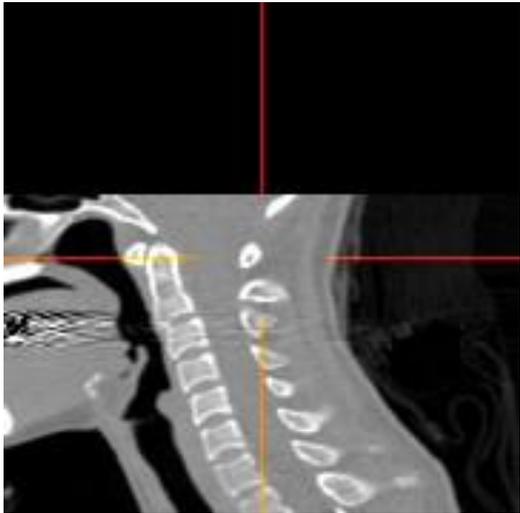
	AX	COR
AX		
SAG		

## 左乳頭 [L\_Nipple]

	AX	COR
AX		
SAG		

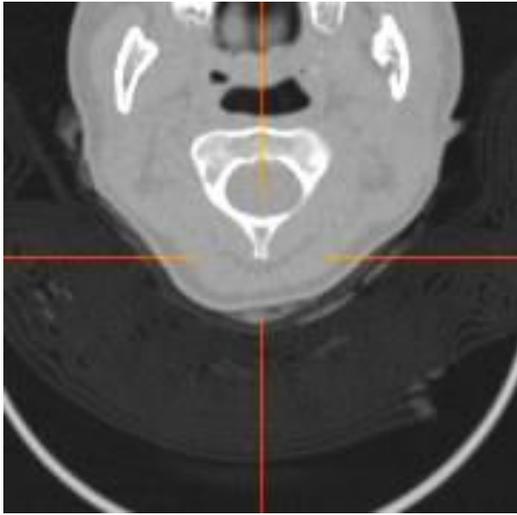
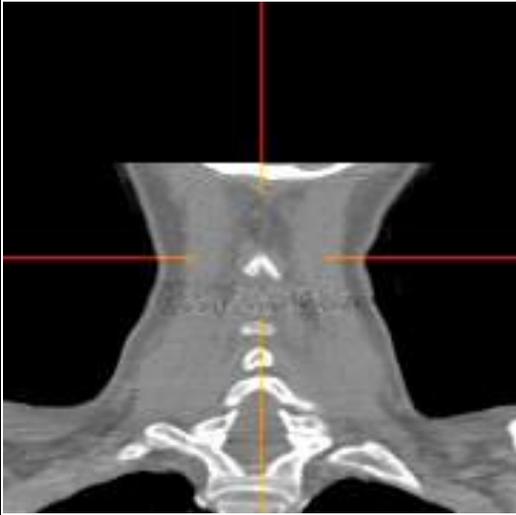
## Appendix B Anatomical landmark list

## 第 1 頤椎棘突起先端 [C1\_SpinousP] (C1 spinous process)

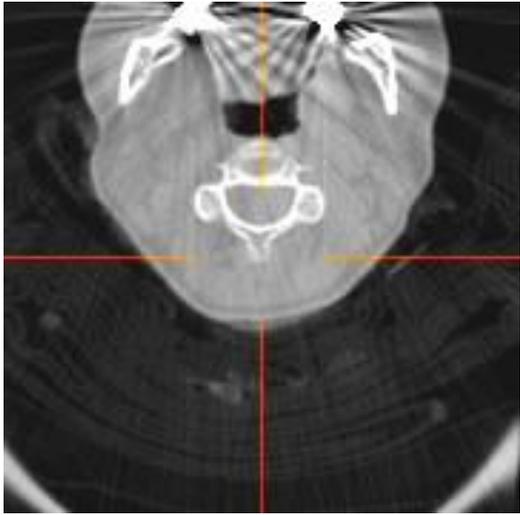
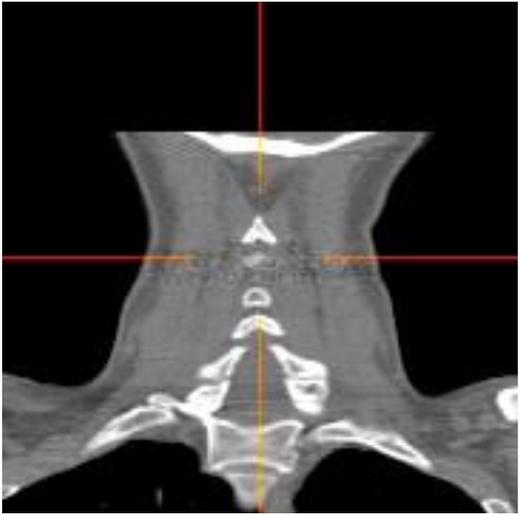
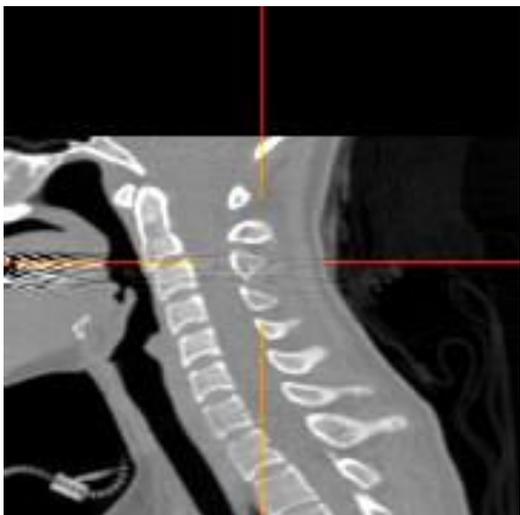
	AX	COR
AX		
SAG		

環椎では棘突起はほとんどないことが多いが、その場合は椎弓の正中背側端をとること。

## 第 2 頸椎棘突起先端 [C2\_SpinousP]

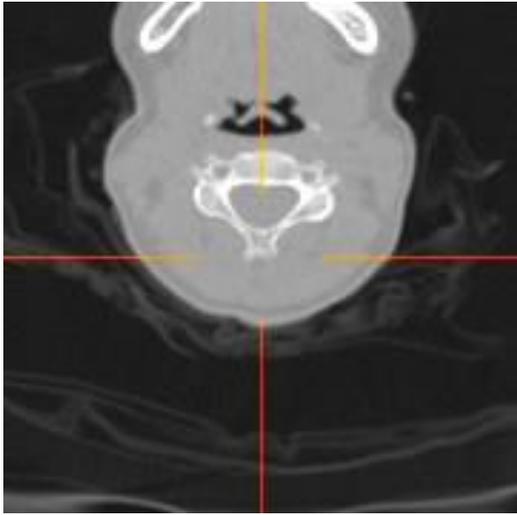
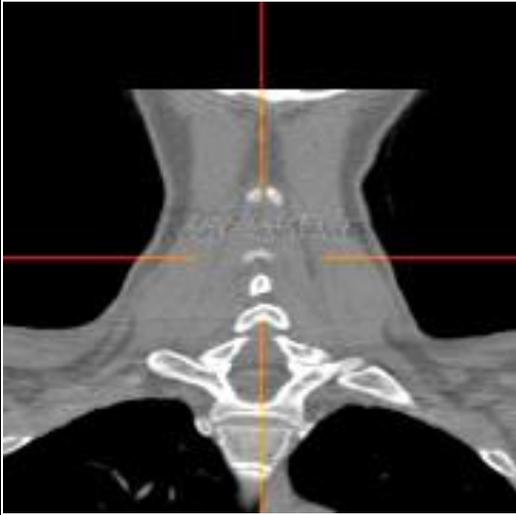
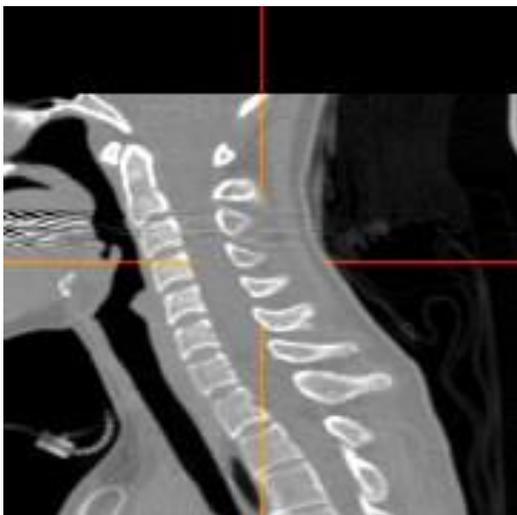
	AX	COR
AX		
SAG		

## 第3 頚椎棘突起先端 [C3\_SpinousP]

	AX	COR
AX		
SAG		

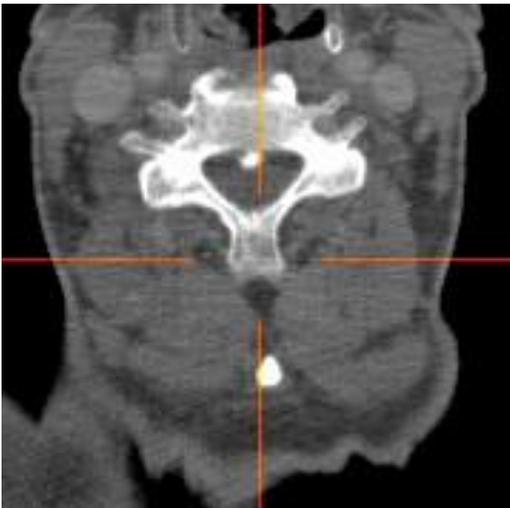
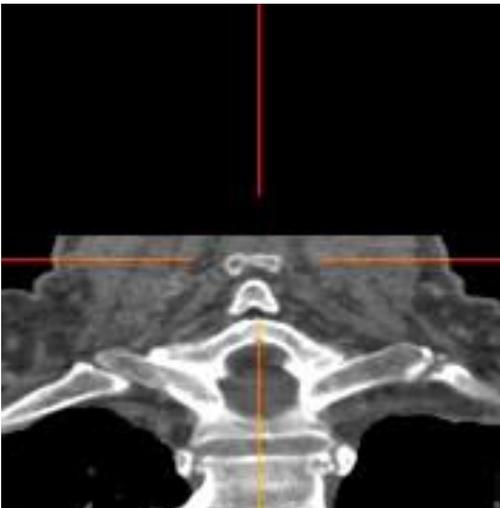
棘突起が二股に分かれているときは、その股の股間の鞍点をとる。さもなくば突起の先端をとる。

第 4 頸椎棘突起先端 [C4\_SpinousP]

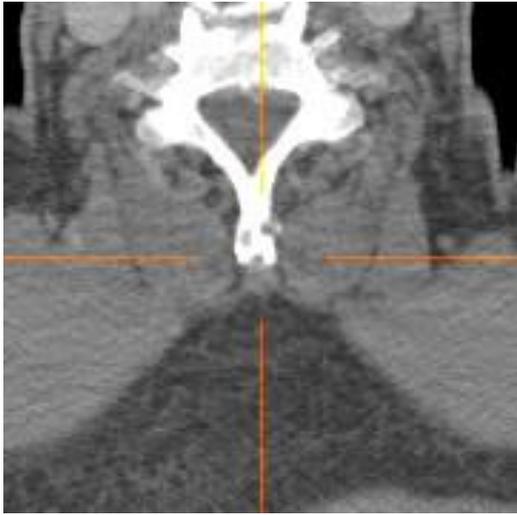
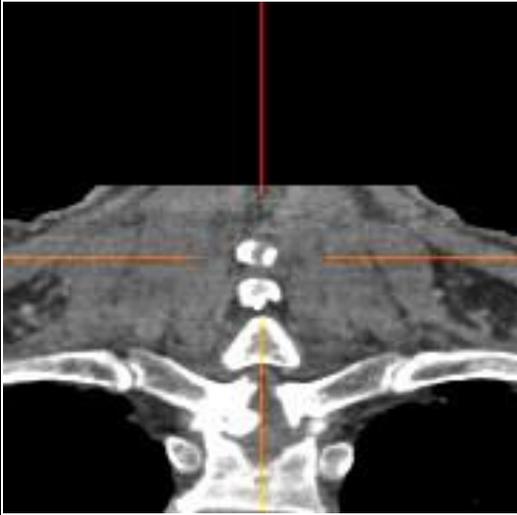
	AX	COR
AX		
SAG		

Appendix B Anatomical landmark list

第 5 頸椎棘突起先端 [C5\_SpinousP]

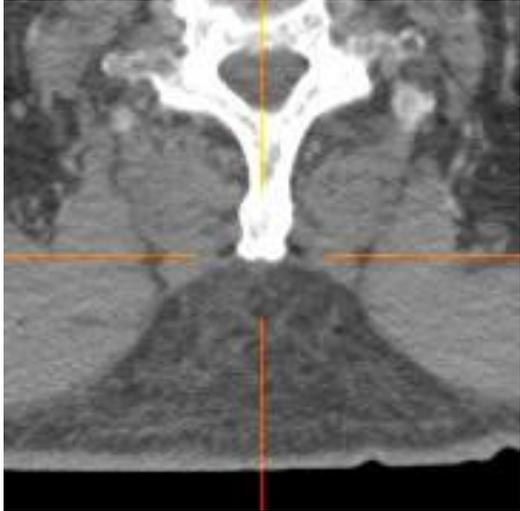
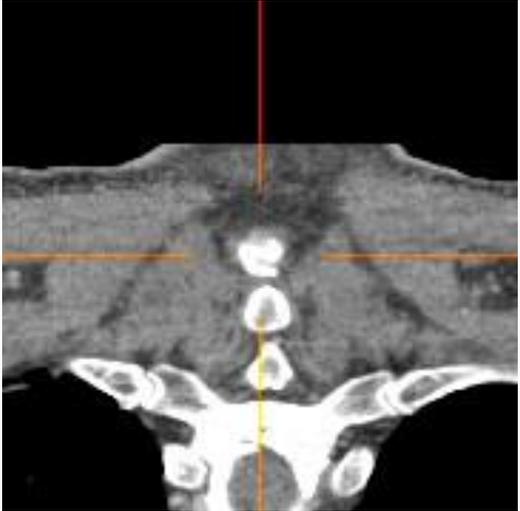
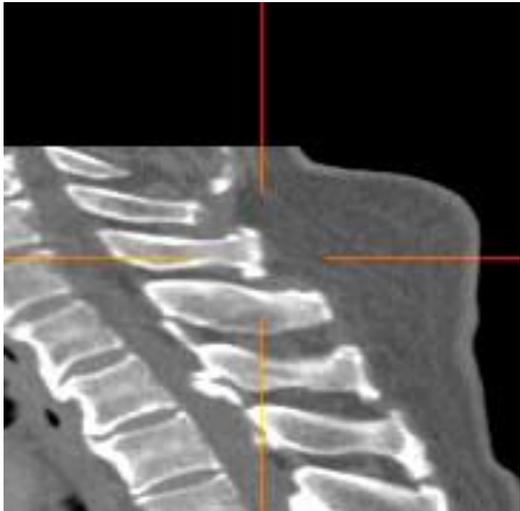
	AX	COR
AX		
SAG		

第 6 頸椎棘突起先端 [C6\_SpinousP]

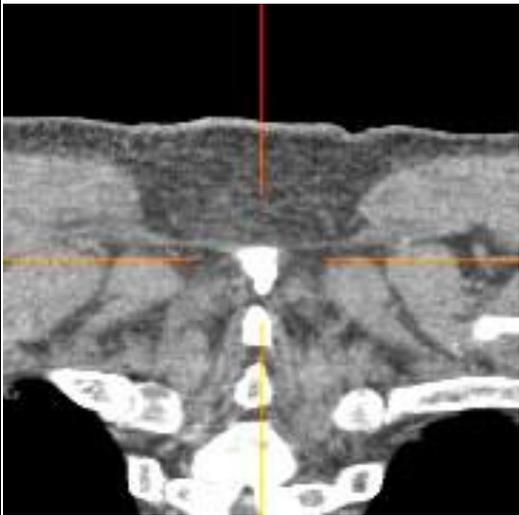
	AX	COR
AX		
SAG		

## Appendix B Anatomical landmark list

## 第 7 頸椎棘突起先端 [C7\_SpinousP]

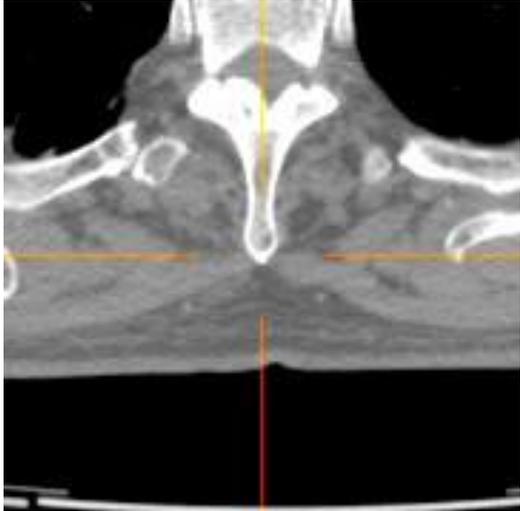
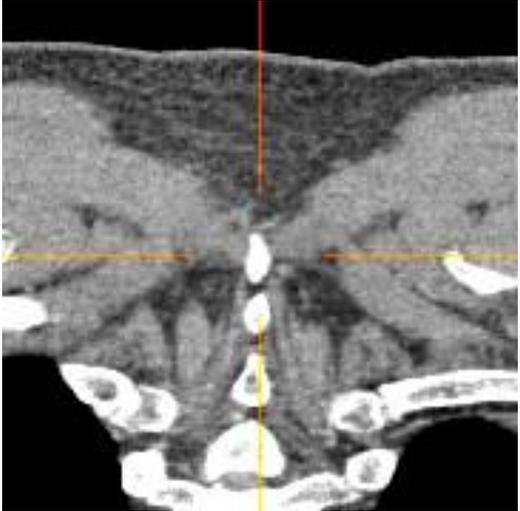
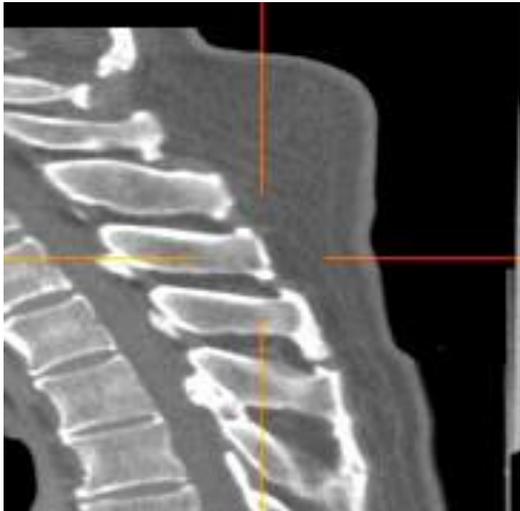
	AX	COR
AX		
SAG		

## 第 1 胸椎棘突起先端 [Th1\_SpinousP]

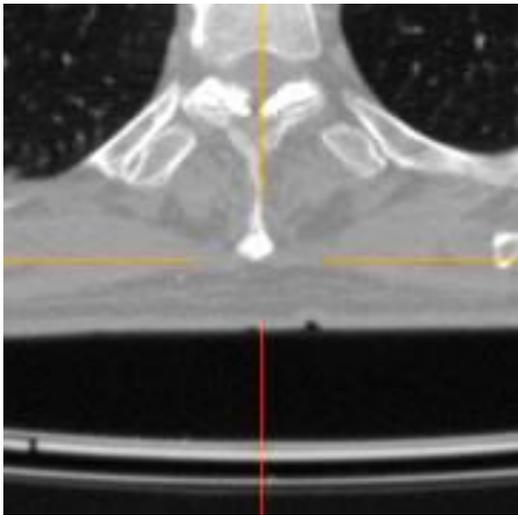
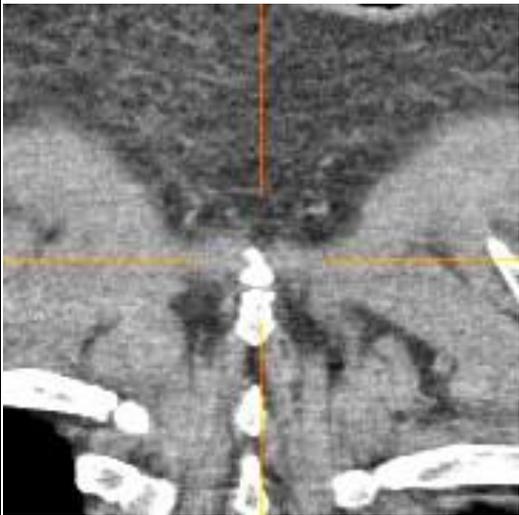
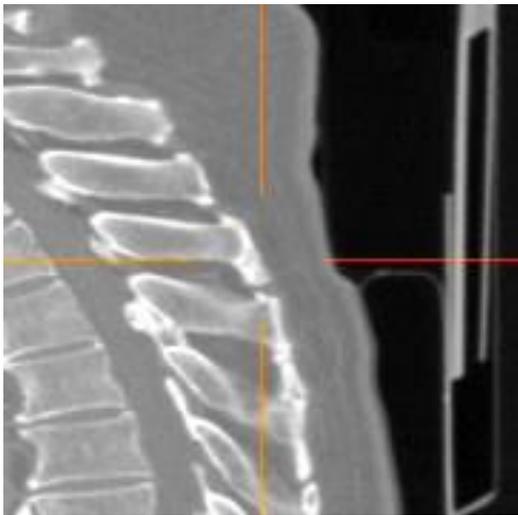
	AX	COR
AX		
SAG		

## Appendix B Anatomical landmark list

## 第 2 胸椎棘突起先端 [Th2\_SpinousP]

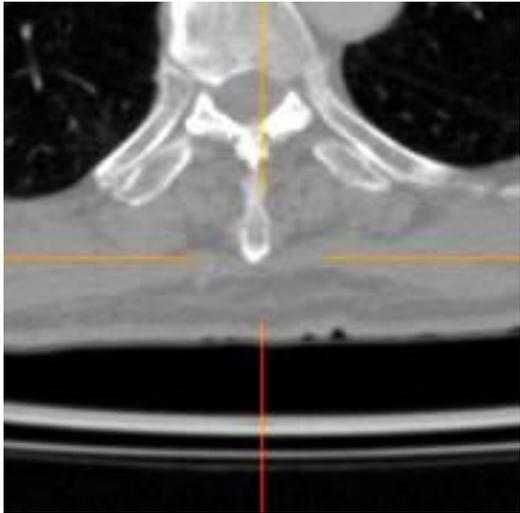
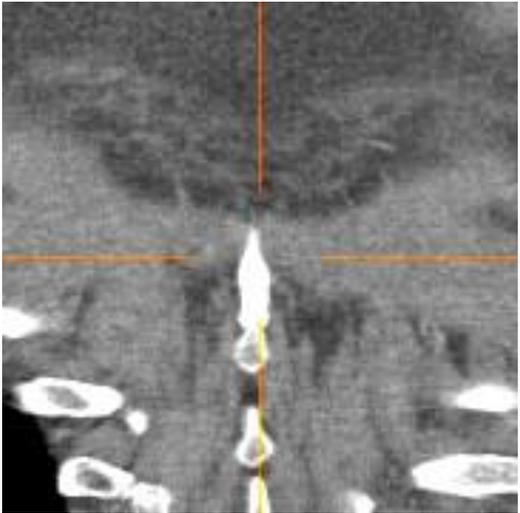
	AX	COR
AX		
SAG		

## 第 3 胸椎棘突起先端 [Th3\_SpinousP]

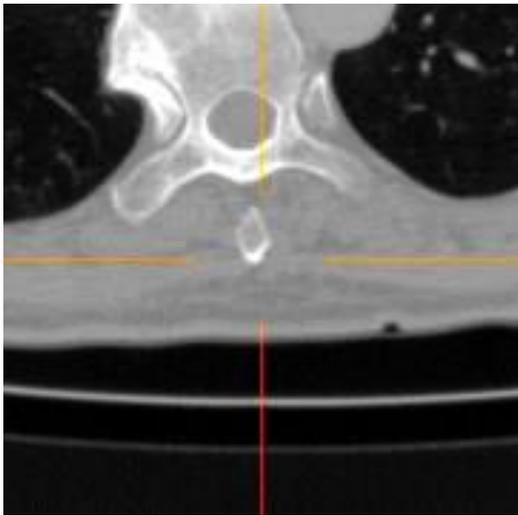
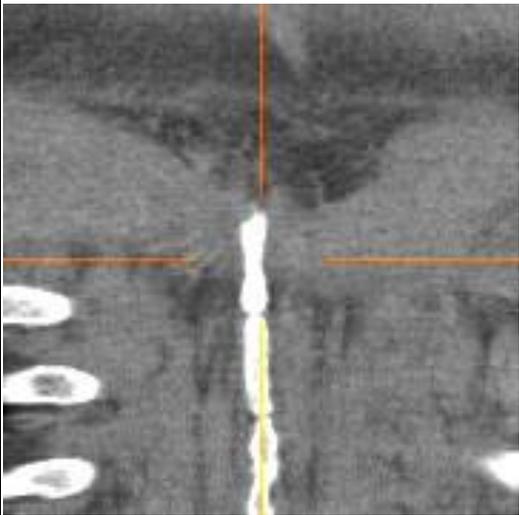
	AX	COR
AX		
SAG		

## Appendix B Anatomical landmark list

## 第 4 胸椎棘突起先端 [Th4\_SpinousP]

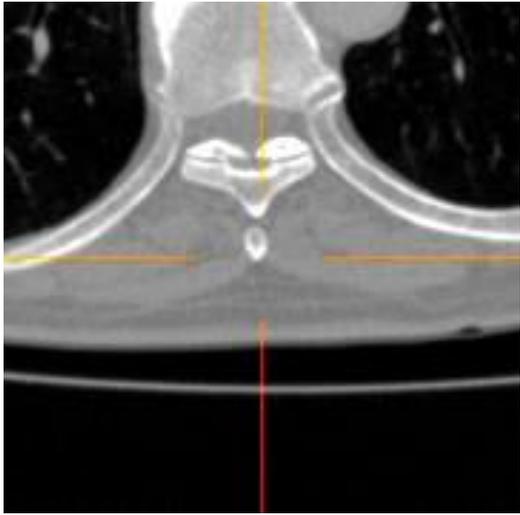
	AX	COR
AX		
SAG		

第 5 胸椎棘突起先端 [Th5\_SpinousP]

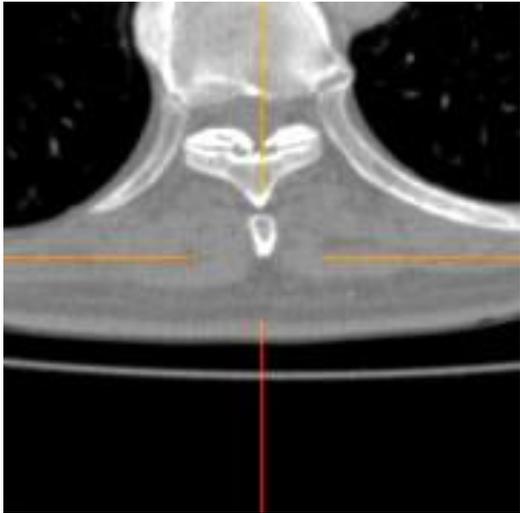
	AX	COR
AX		
SAG		

## Appendix B Anatomical landmark list

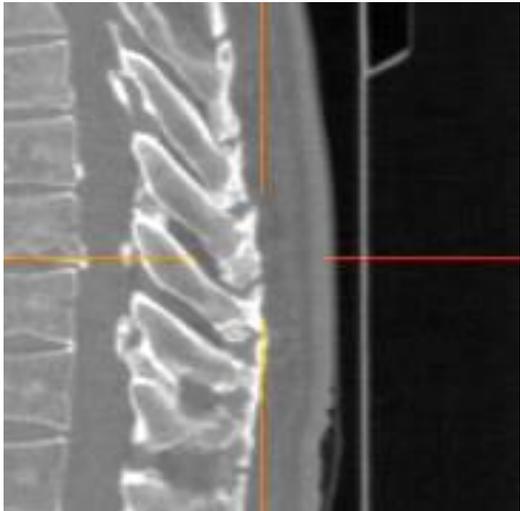
## 第 6 胸椎棘突起先端 [Th6\_SpinousP]

	AX	COR
AX		
SAG		

## 第 7 胸椎棘突起先端 [Th7\_SpinousP]

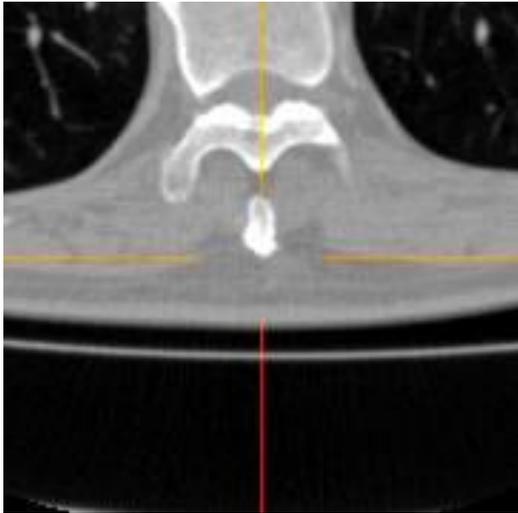
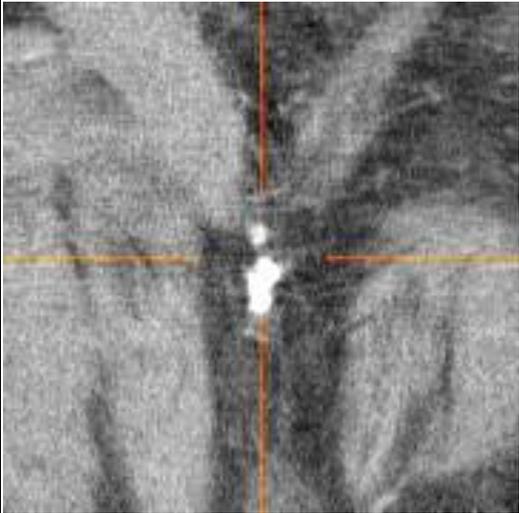
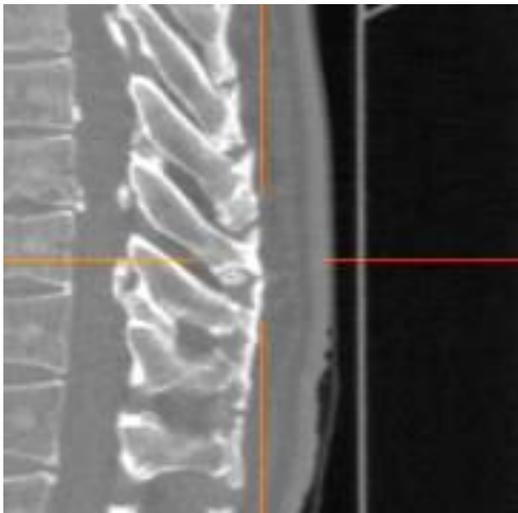
	AX	COR
AX		
SAG		

## 第 8 胸椎棘突起先端 [Th8\_SpinousP]

	AX	COR
AX		
SAG		

Th-6\_SpinousP

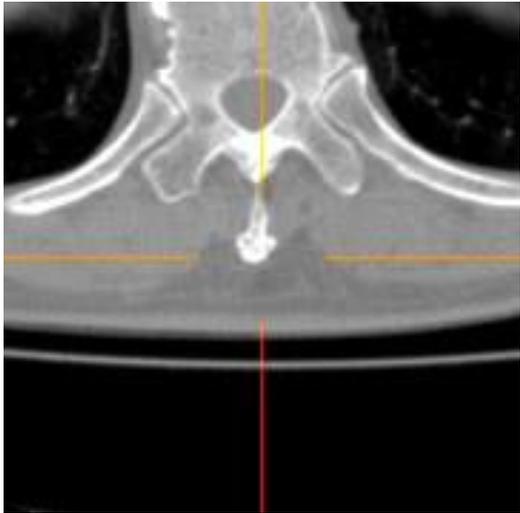
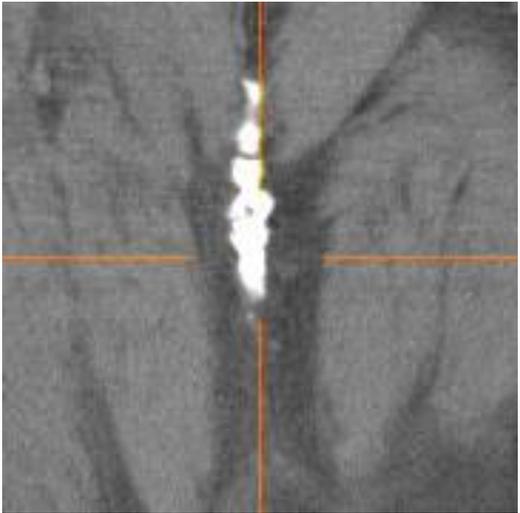
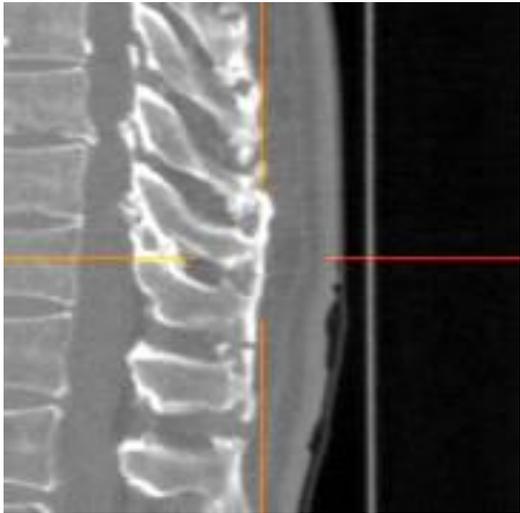
第 9 胸椎棘突起先端 [Th9\_SpinousP]

	AX	COR
AX		
SAG		

Th-5\_SpinousP

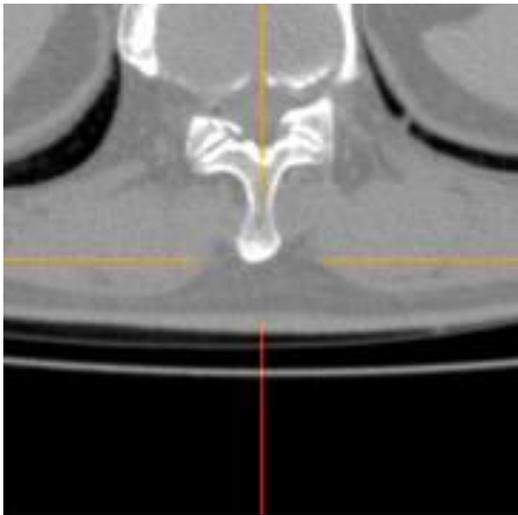
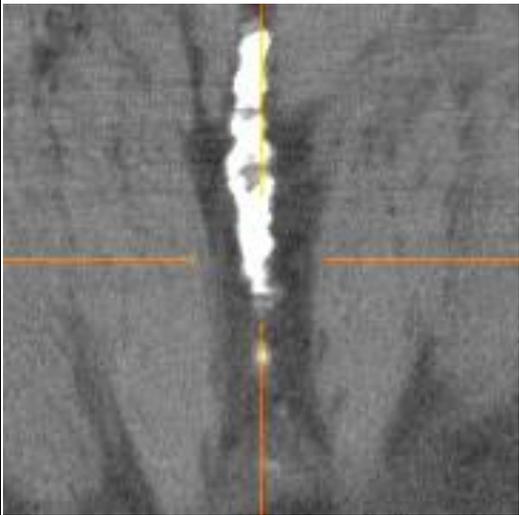
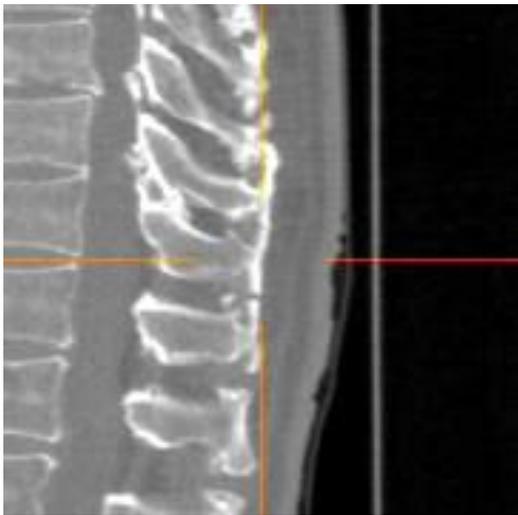
## Appendix B Anatomical landmark list

## 第 10 胸椎棘突起先端 [Th10\_SpinousP]

	AX	COR
AX		
SAG		

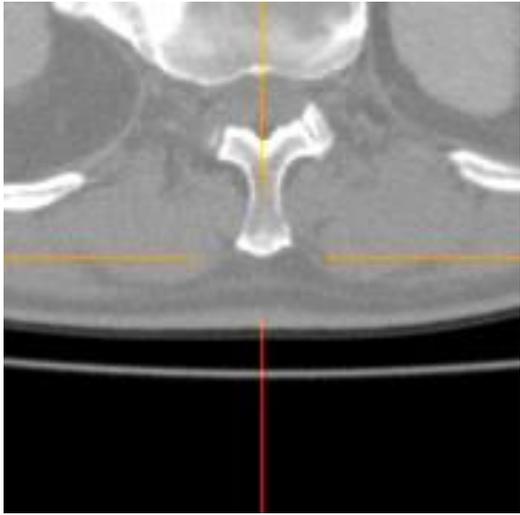
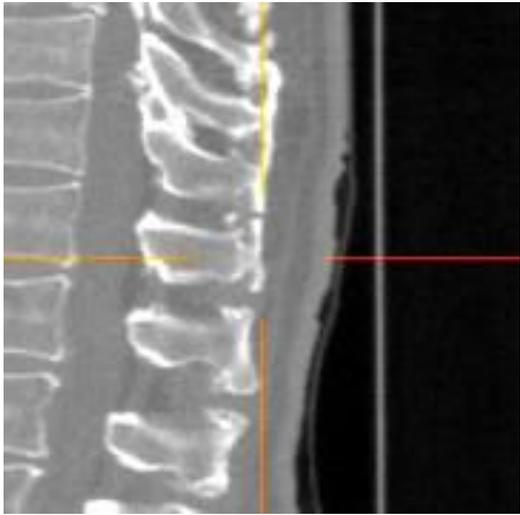
Th-4\_SpinousP

第 11 胸椎棘突起先端 [Th11\_SpinousP]

	AX	COR
AX		
SAG		

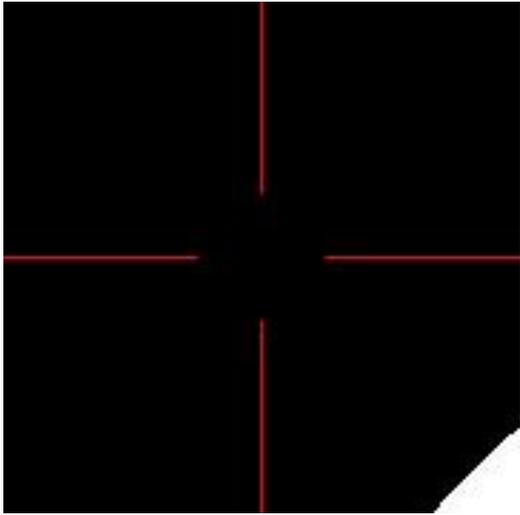
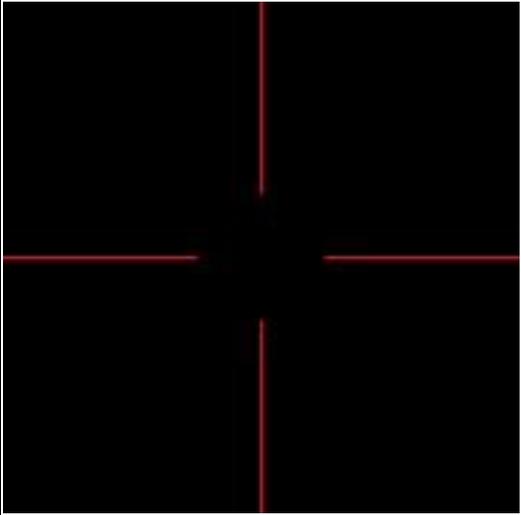
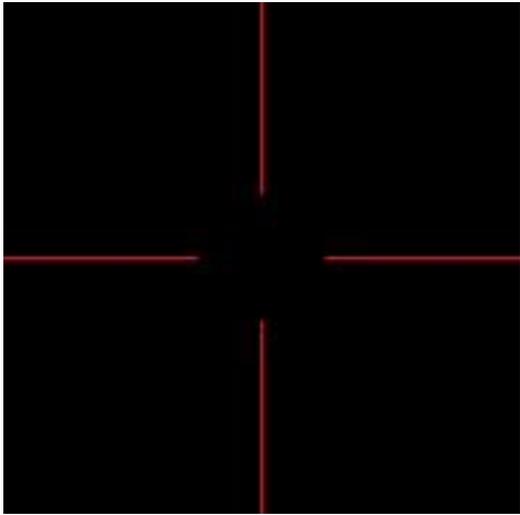
Th-3\_SpinousP

## 第 12 胸椎棘突起先端 [Th12\_SpinousP]

	AX	COR
AX		
SAG		

Th-2\_SpinousP

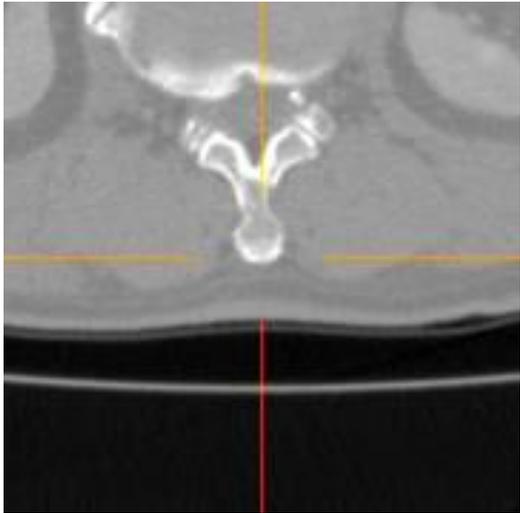
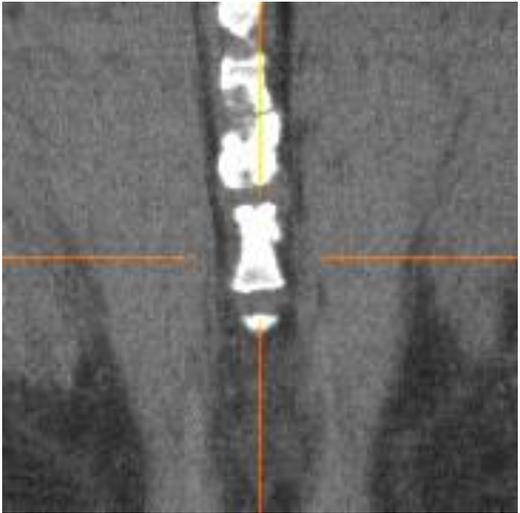
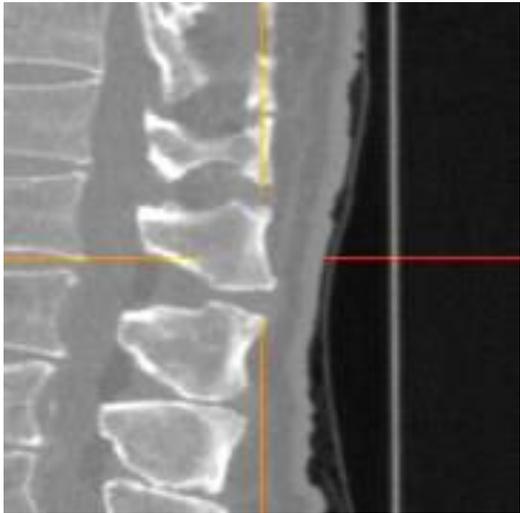
## 第 13 胸椎棘突起先端 [Th13\_SpinousP]

	AX	COR
AX		
SAG		

(破格時のみ)

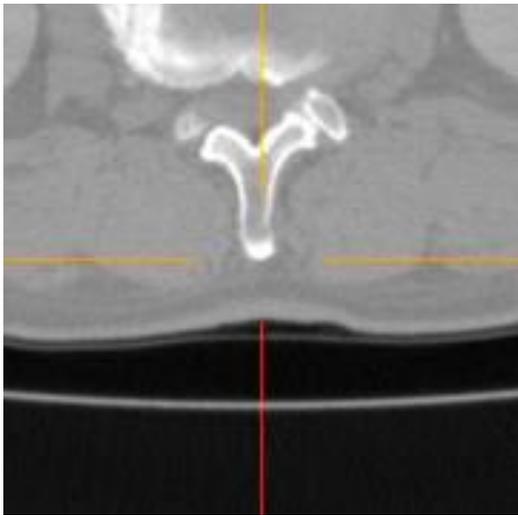
## Appendix B Anatomical landmark list

## 第 1 腰椎棘突起先端 [L1\_SpinousP]

	AX	COR
AX		
SAG		

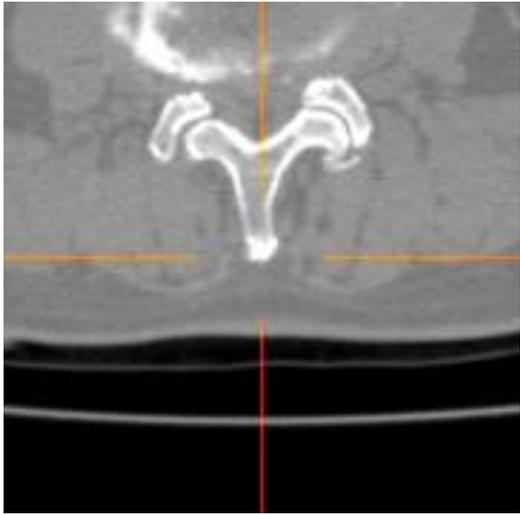
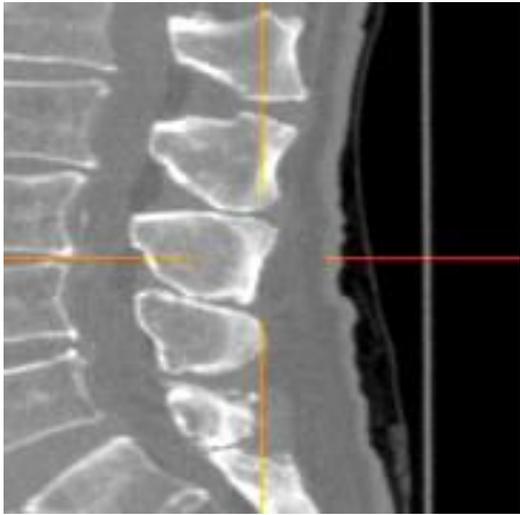
上下に長い板状の棘突起の場合は、極力中央あたりをとる。

## 第 2 腰椎棘突起先端 [L2\_SpinousP]

	AX	COR
AX		
SAG		

## Appendix B Anatomical landmark list

## 第 3 腰椎棘突起先端 [L3\_SpinousP]

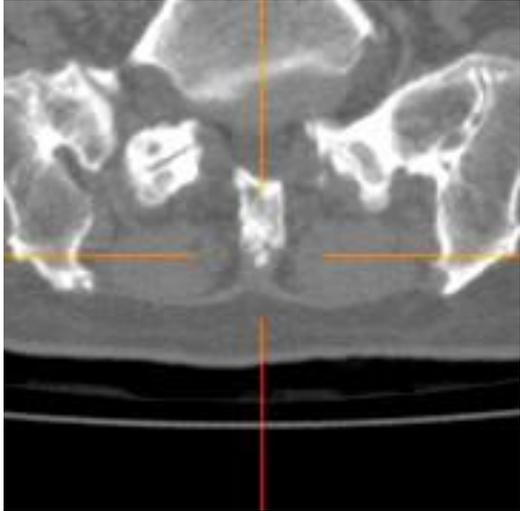
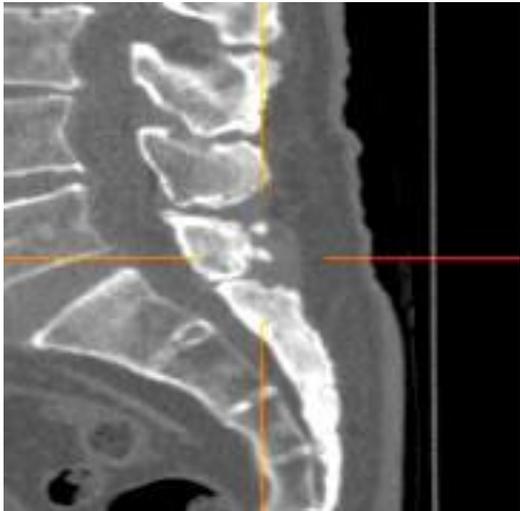
	AX	COR
AX		
SAG		

第 4 腰椎棘突起先端 [L4\_SpinousP]

	AX	COR
AX		
SAG		

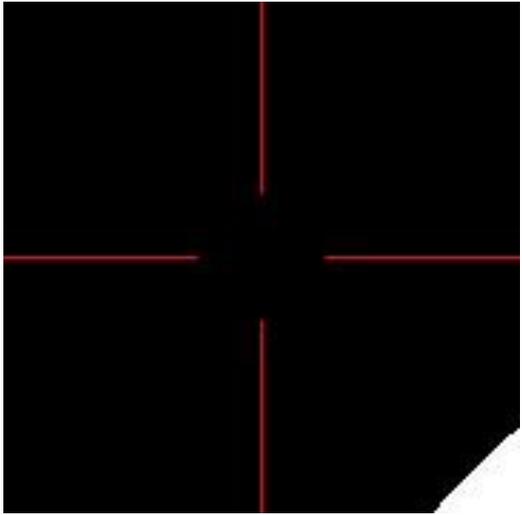
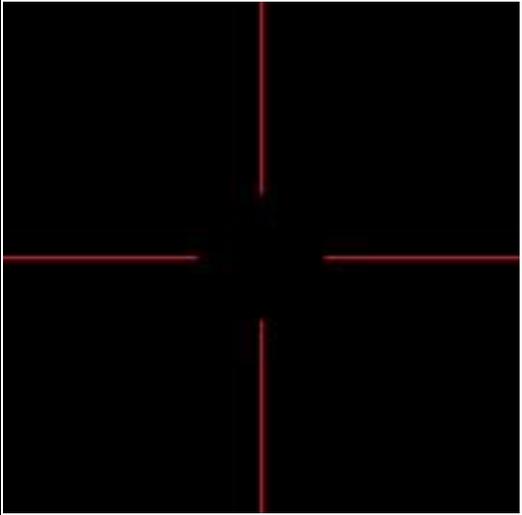
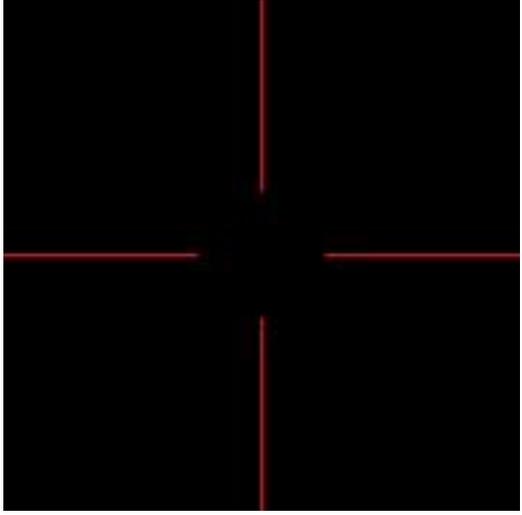
L-3\_SpinousP

第 5 腰椎棘突起先端 [L5\_SpinousP]

	AX	COR
AX		
SAG		

L-2\_SpinousP

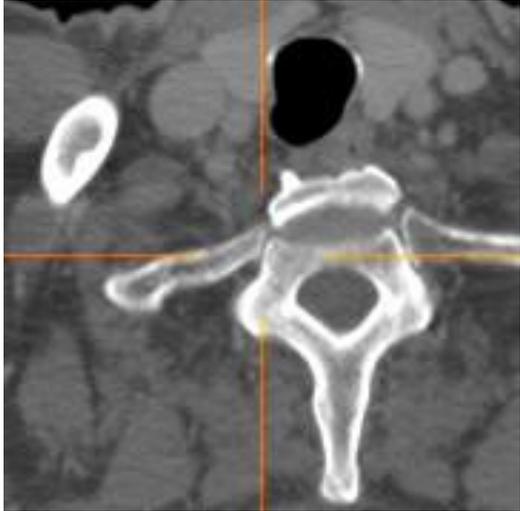
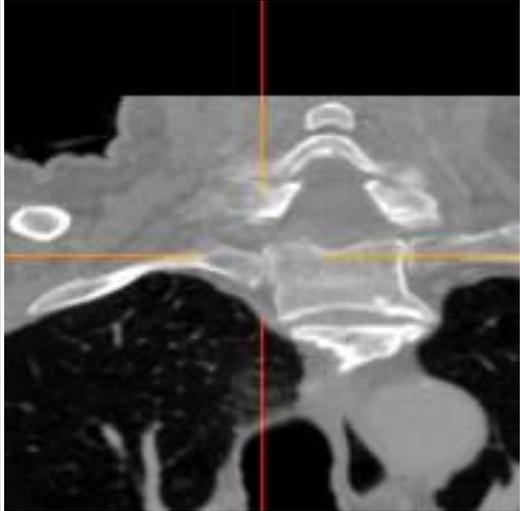
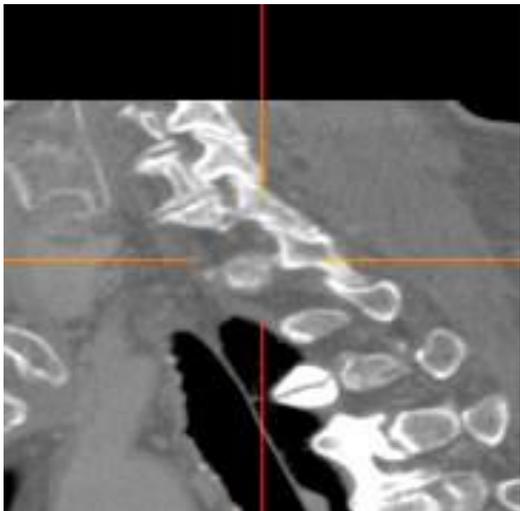
## 第 6 腰椎棘突起先端 [L6\_SpinousP]

	AX	COR
AX		
SAG		

(破格時のみ)

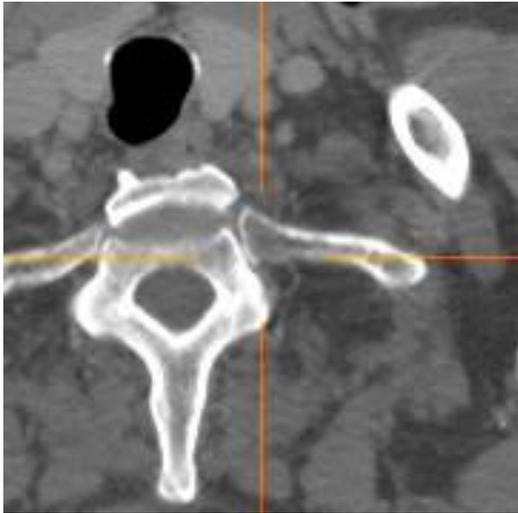
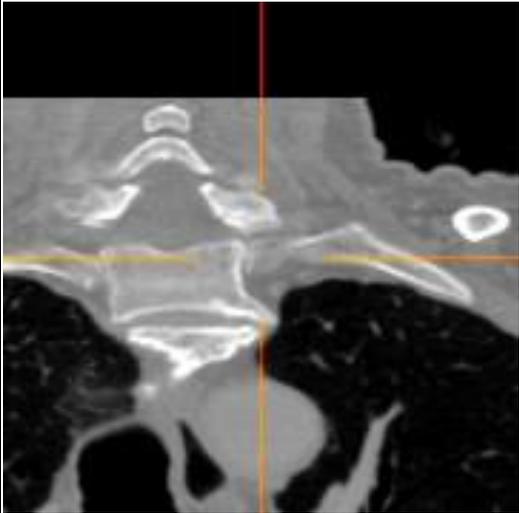
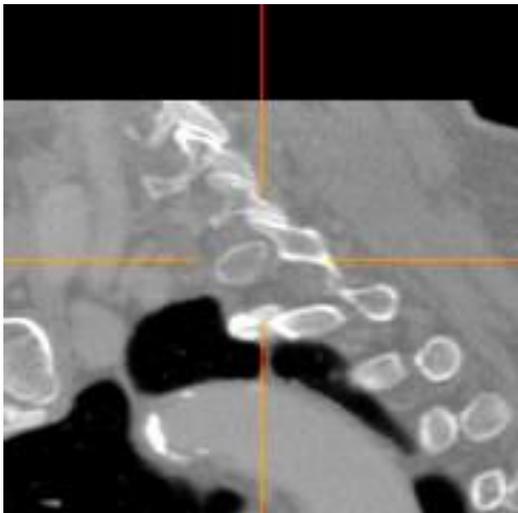
## Appendix B Anatomical landmark list

## 右第 1 肋骨近位端上縁 [R\_Rib1\_Prox] (proximal tip of rt. 1st rib)

	AX	COR
AX		
SAG		

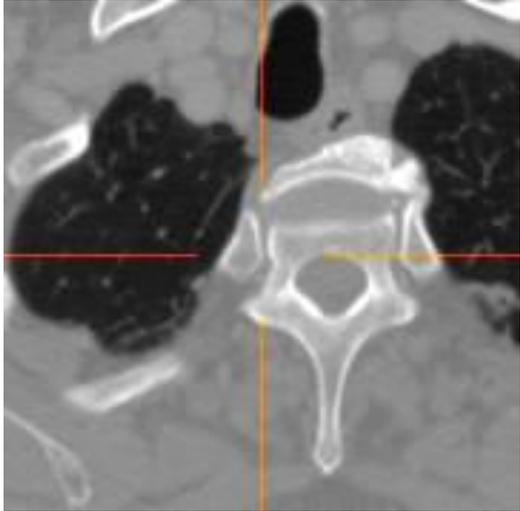
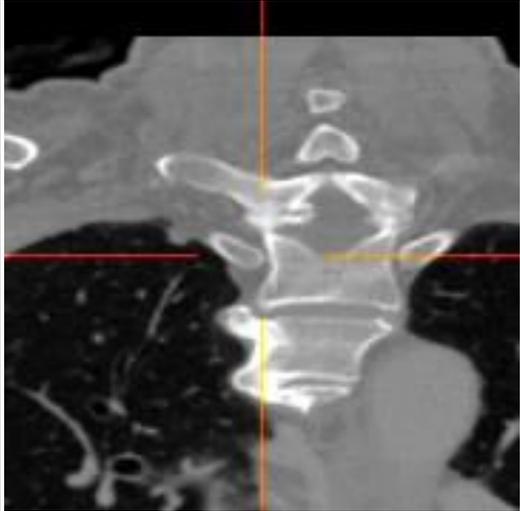
肋骨が椎骨横突起の根本前縁と接するところで、肋骨-椎骨間の関節があるところ。  
 だいたい椎弓根と同じ高さでとる。

左第 1 肋骨近位端上縁 [L\_Rib1\_Prox] (proximal tip of lt. 1st rib)

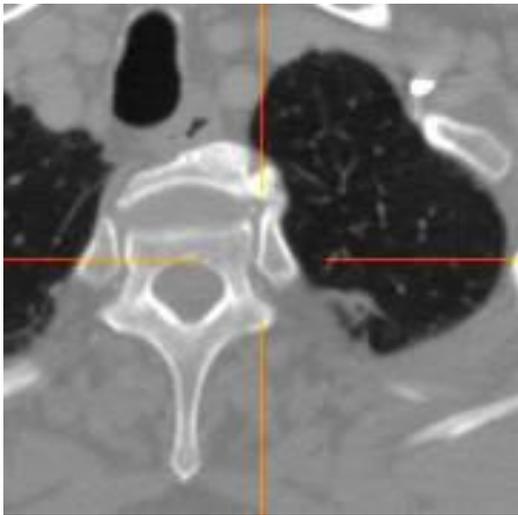
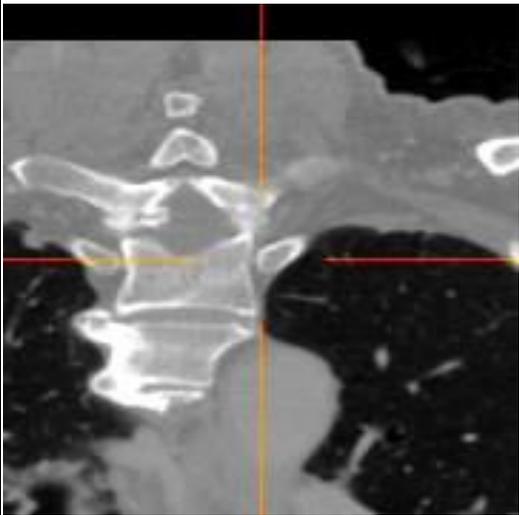
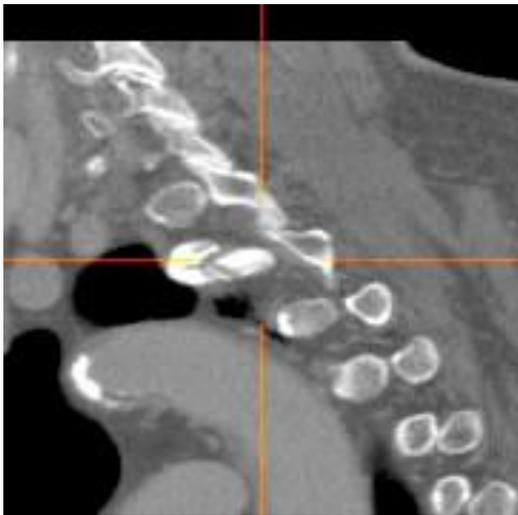
	AX	COR
AX		
SAG		

## Appendix B Anatomical landmark list

## 右第 2 肋骨近位端上縁 [R\_Rib2\_Prox]

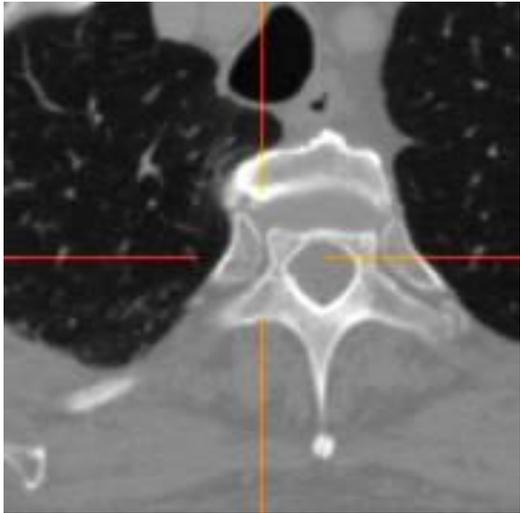
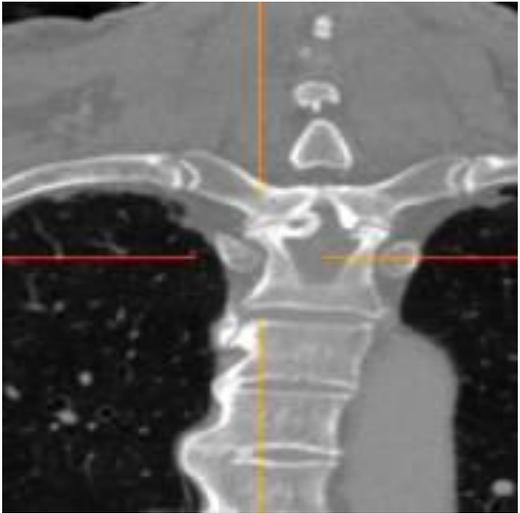
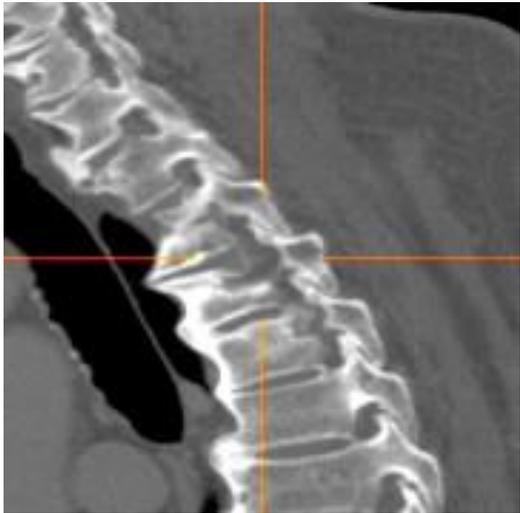
	AX	COR
AX		
SAG		

## 左第2肋骨近位端上縁 [L\_Rib2\_Prox]

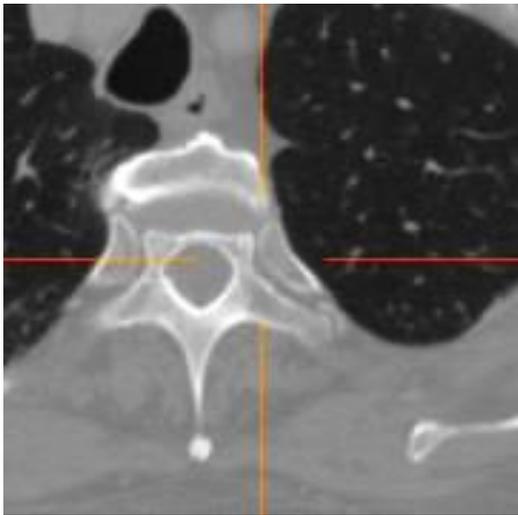
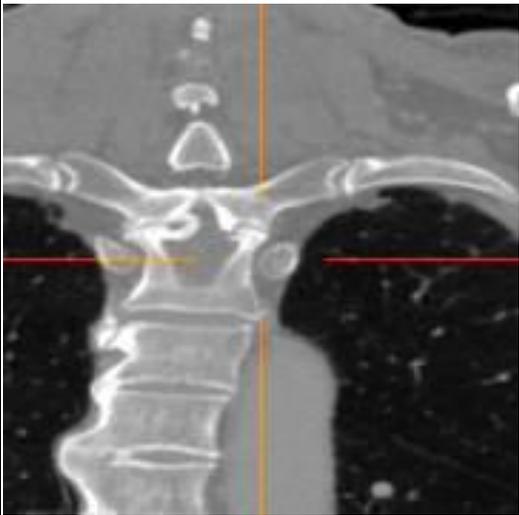
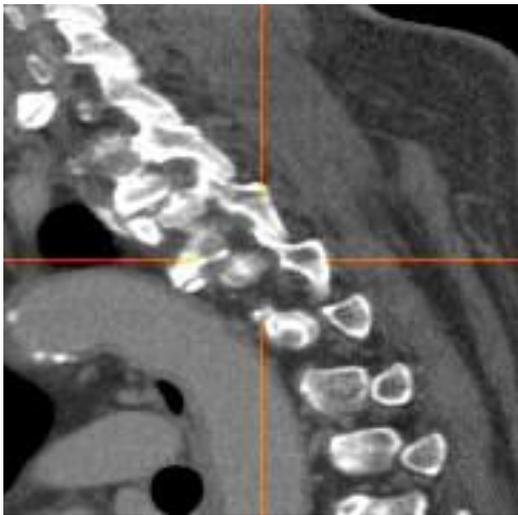
	AX	COR
AX		
SAG		

## Appendix B Anatomical landmark list

## 右第 3 肋骨近位端上縁 [R\_Rib3\_Prox]

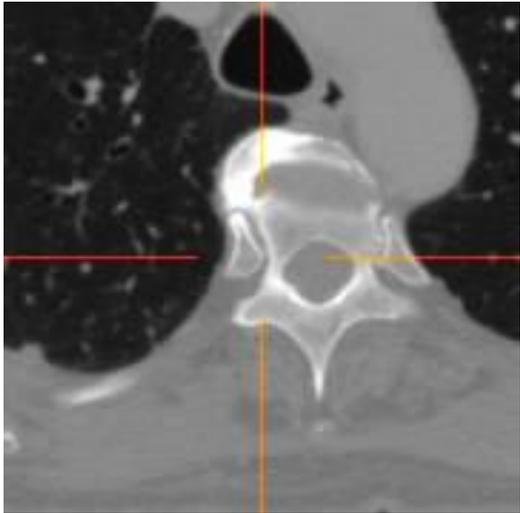
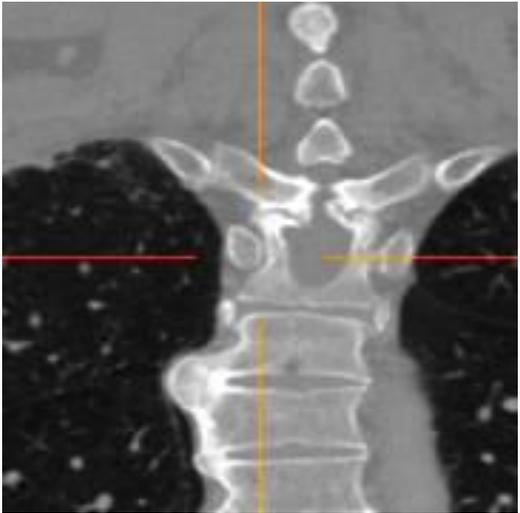
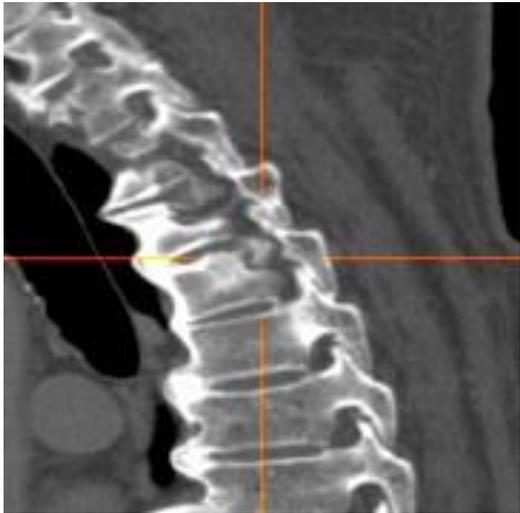
	AX	COR
AX		
SAG		

左第 3 肋骨近位端上縁 [L\_Rib3\_Prox]

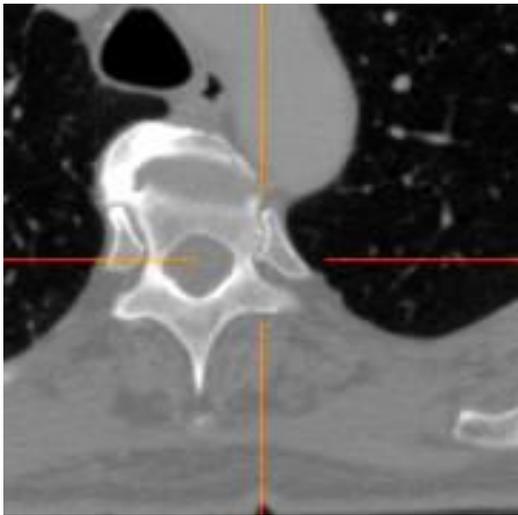
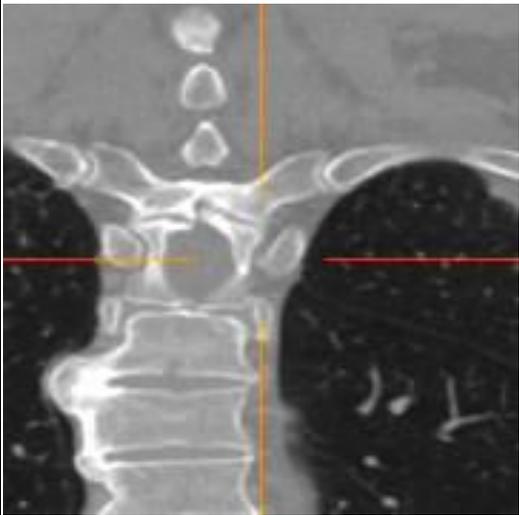
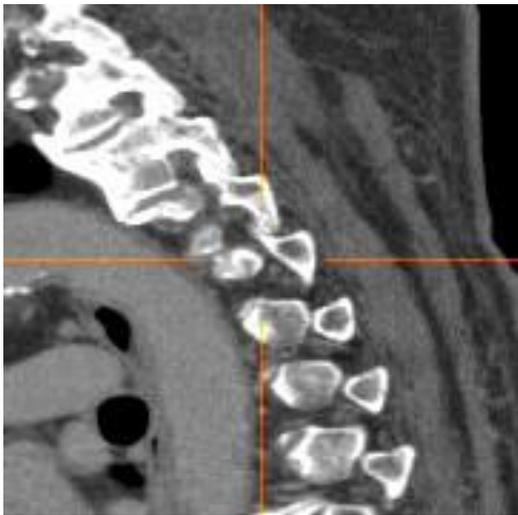
	AX	COR
AX		
SAG		

## Appendix B Anatomical landmark list

## 右第 4 肋骨近位端上縁 [R\_Rib4\_Prox]

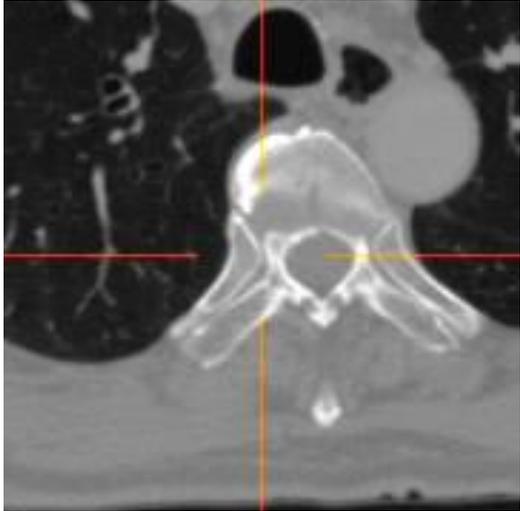
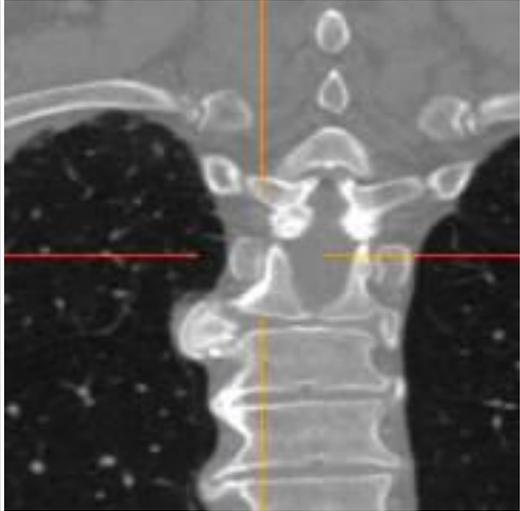
	AX	COR
AX		
SAG		

左第 4 肋骨近位端上縁 [L\_Rib4\_Prox]

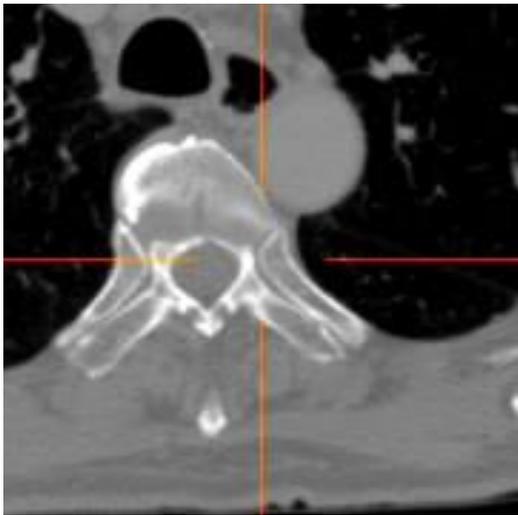
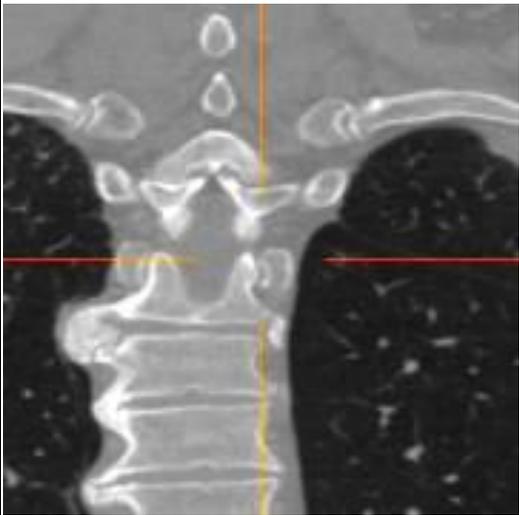
	AX	COR
AX		
SAG		

## Appendix B Anatomical landmark list

## 右第 5 肋骨近位端上縁 [R\_Rib5\_Prox]

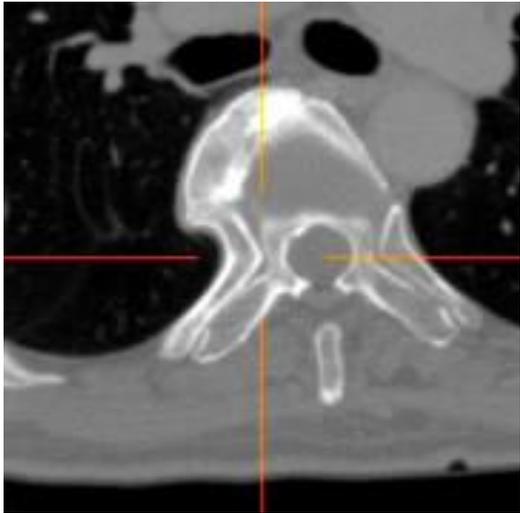
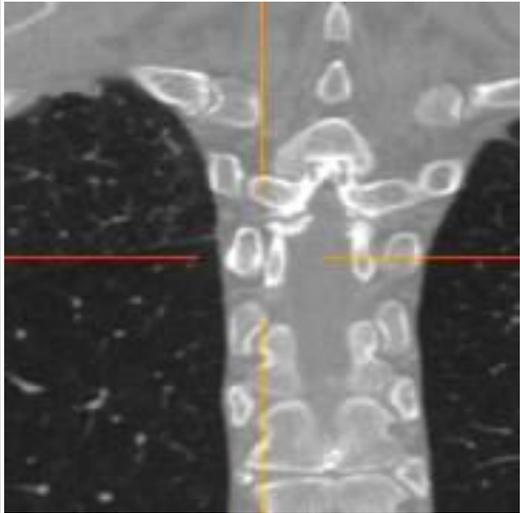
	AX	COR
AX		
SAG		

左第 5 肋骨近位端上縁 [L\_Rib5\_Prox]

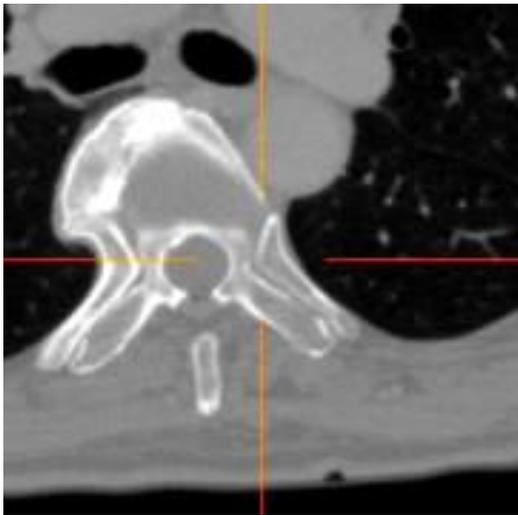
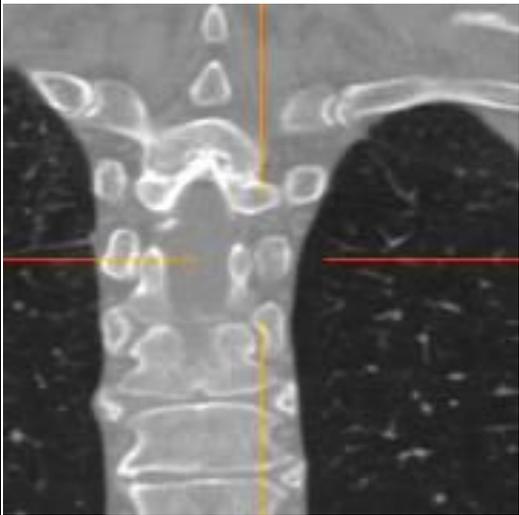
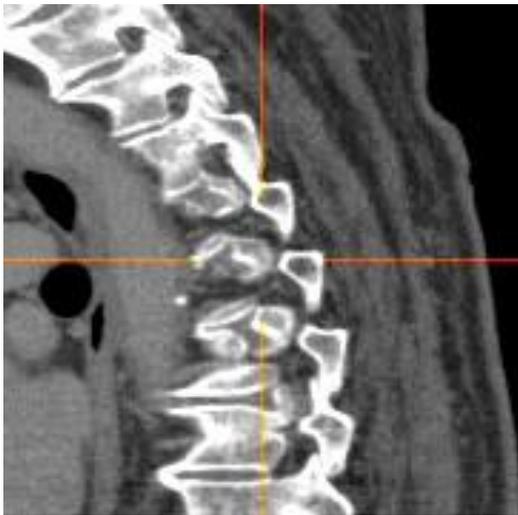
	AX	COR
AX		
SAG		

## Appendix B Anatomical landmark list

## 右第 6 肋骨近位端上縁 [R\_Rib6\_Prox]

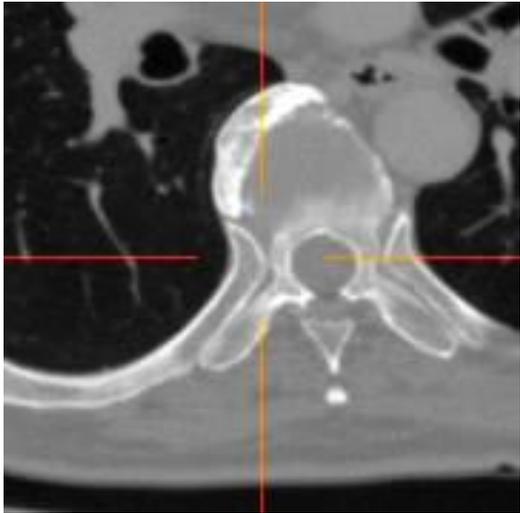
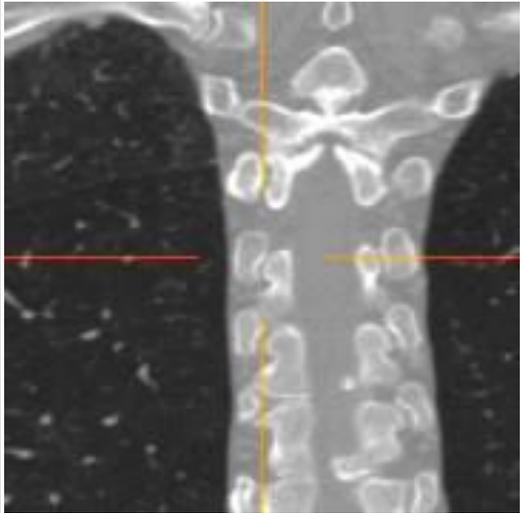
	AX	COR
AX		
SAG		

左第 6 肋骨近位端上縁 [L\_Rib6\_Prox]

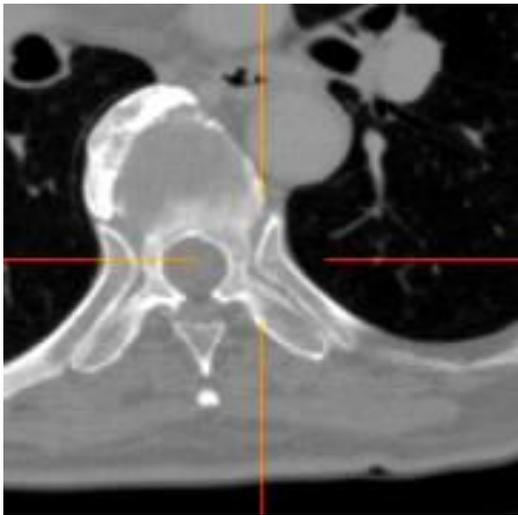
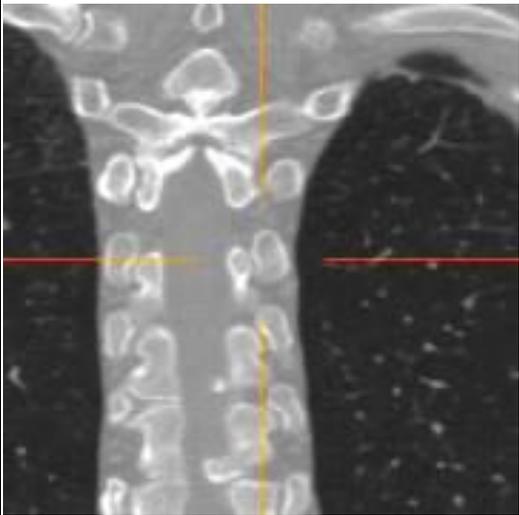
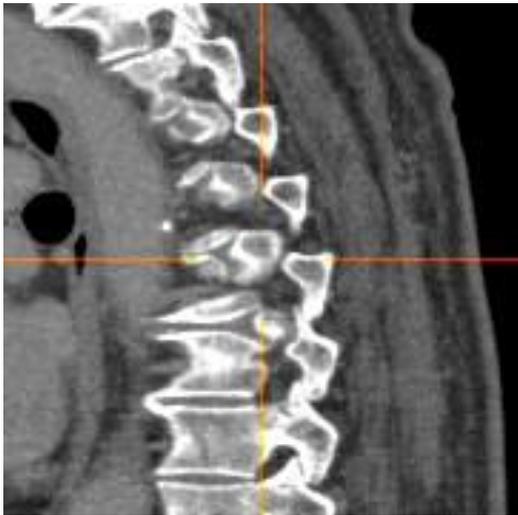
	AX	COR
AX		
SAG		

## Appendix B Anatomical landmark list

## 右第 7 肋骨近位端上縁 [R\_Rib7\_Prox]

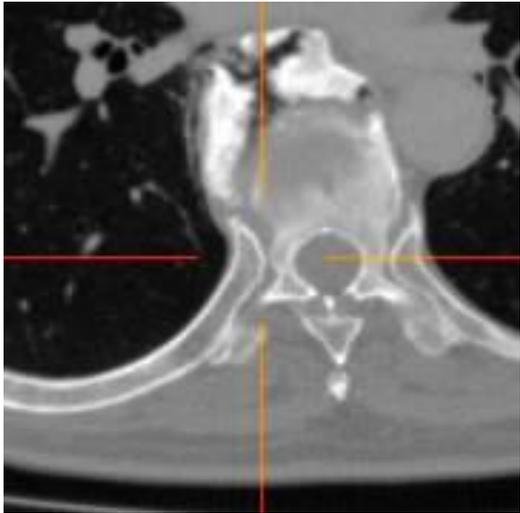
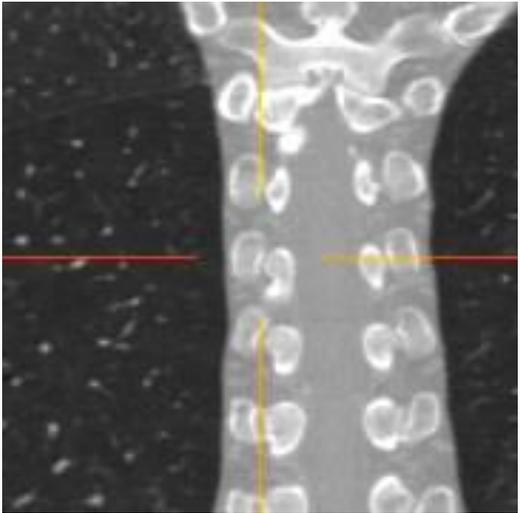
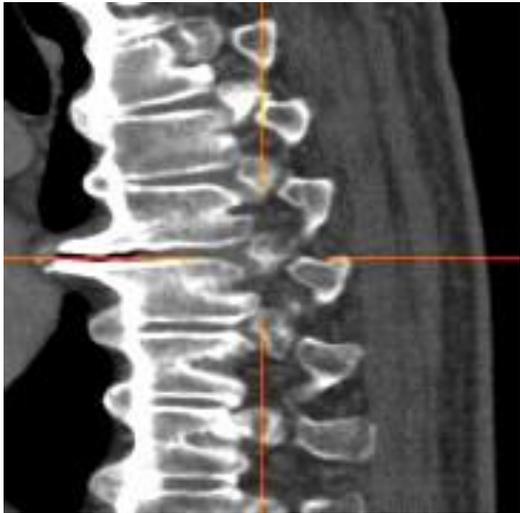
	AX	COR
AX		
SAG		

左第 7 肋骨近位端上縁 [L\_Rib7\_Prox]

	AX	COR
AX		
SAG		

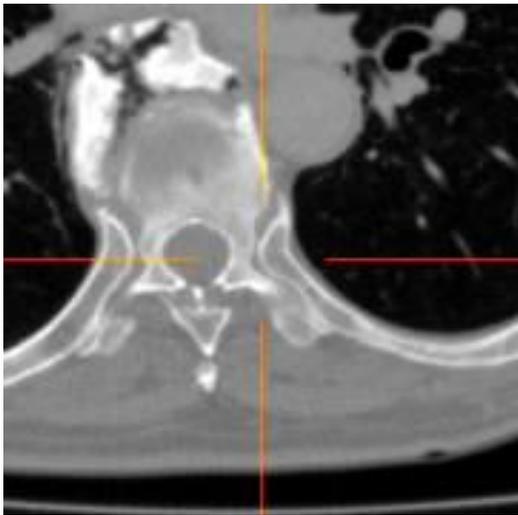
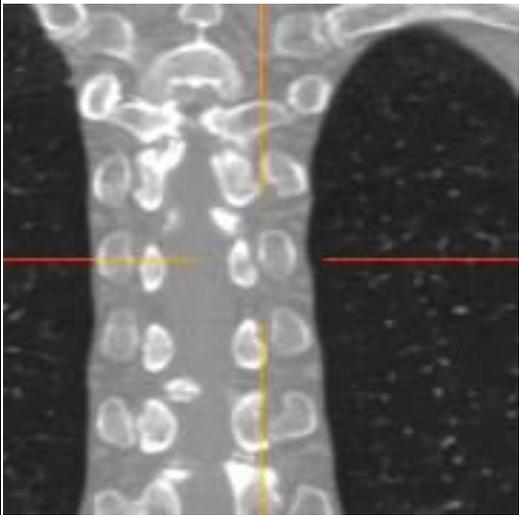
## Appendix B Anatomical landmark list

## 右第 8 肋骨近位端上縁 [R\_Rib8\_Prox]

	AX	COR
AX		
SAG		

R\_Rib-6\_Prox

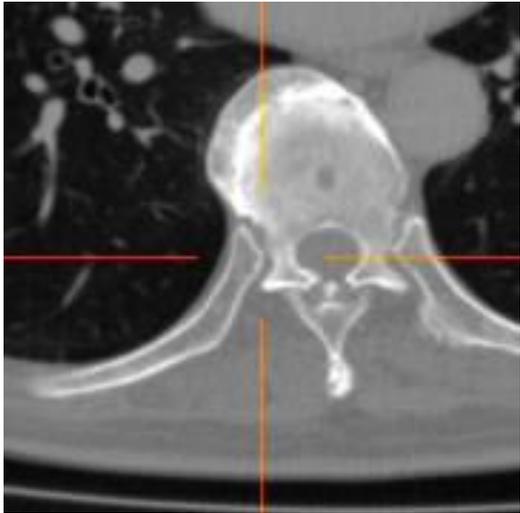
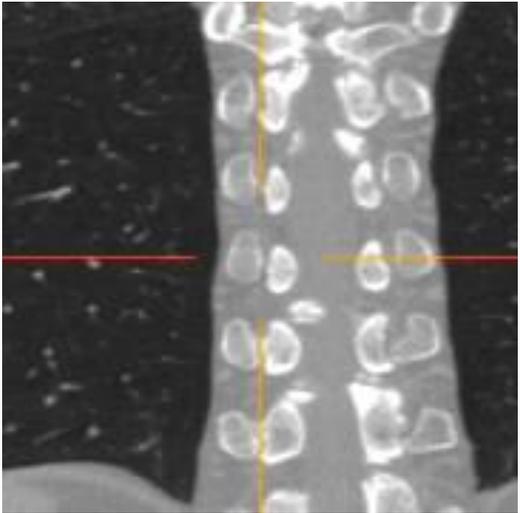
## 左第 8 肋骨近位端上縁 [L\_Rib8\_Prox]

	AX	COR
AX		
SAG		

L\_Rib-6\_Prox

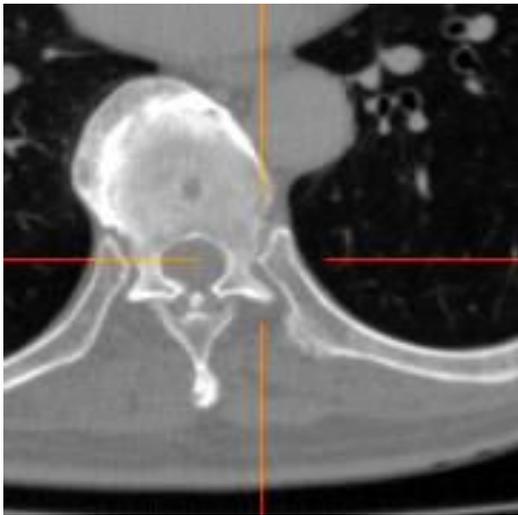
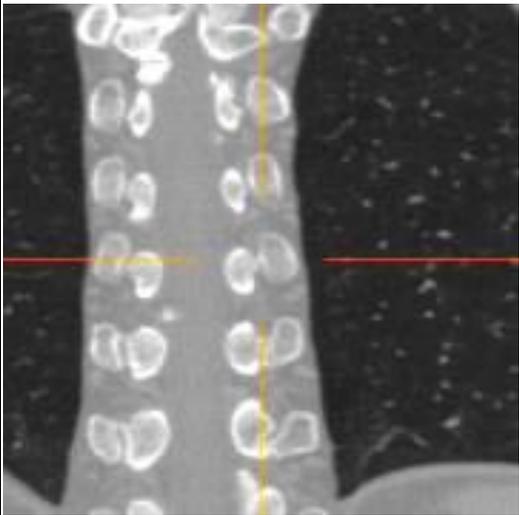
## Appendix B Anatomical landmark list

## 右第 9 肋骨近位端上縁 [R\_Rib9\_Prox]

	AX	COR
AX		
SAG		

R\_Rib-5\_Prox

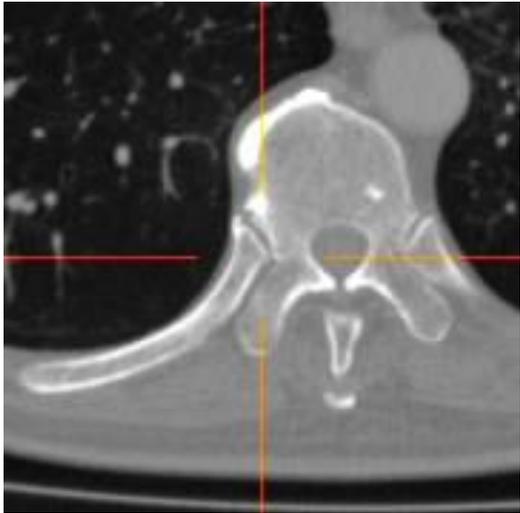
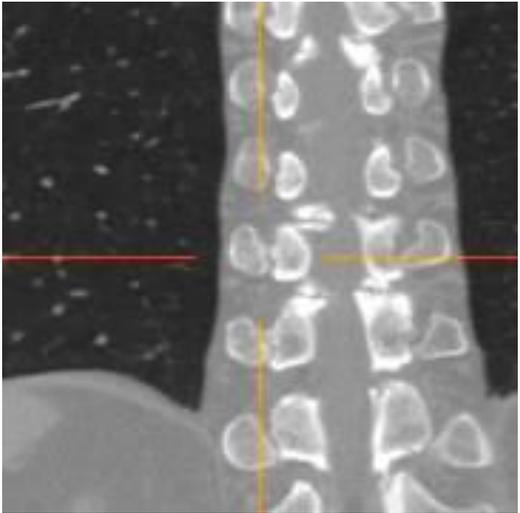
左第 9 肋骨近位端上縁 [L\_Rib9\_Prox]

	AX	COR
AX		
SAG		

L\_Rib-5\_Prox

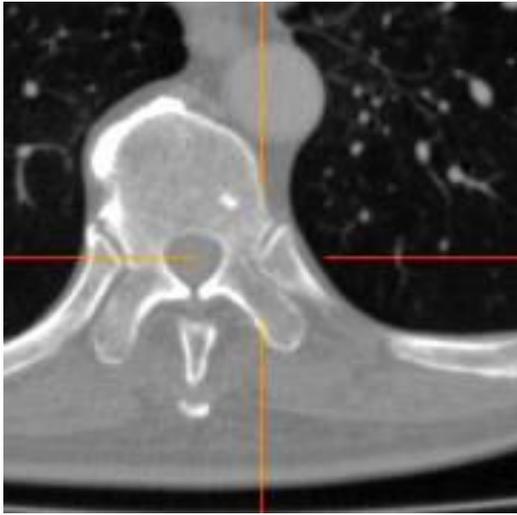
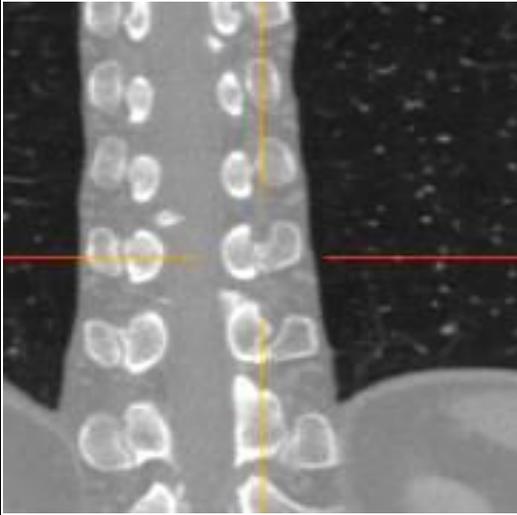
## Appendix B Anatomical landmark list

## 右第 10 肋骨近位端上縁 [R\_Rib10\_Prox]

	AX	COR
AX		
SAG		

R\_Rib-4\_Prox

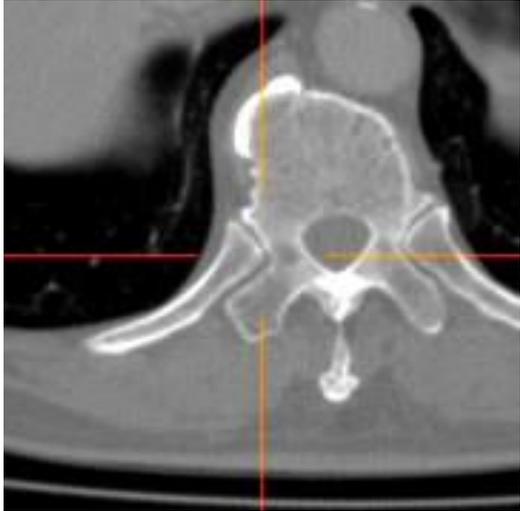
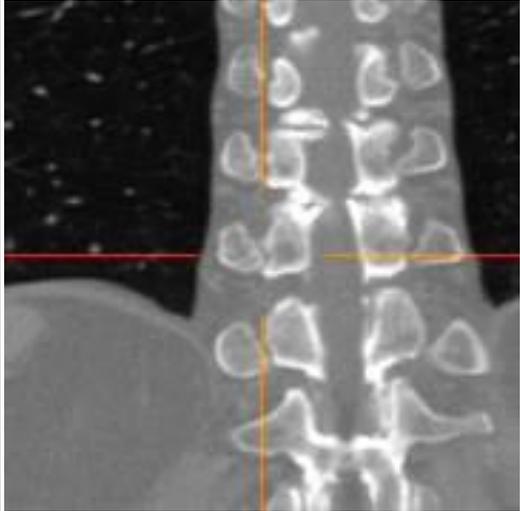
左第 10 肋骨近位端上縁 [L\_Rib10\_Prox]

	AX	COR
AX		
SAG		

L\_Rib-4\_Prox

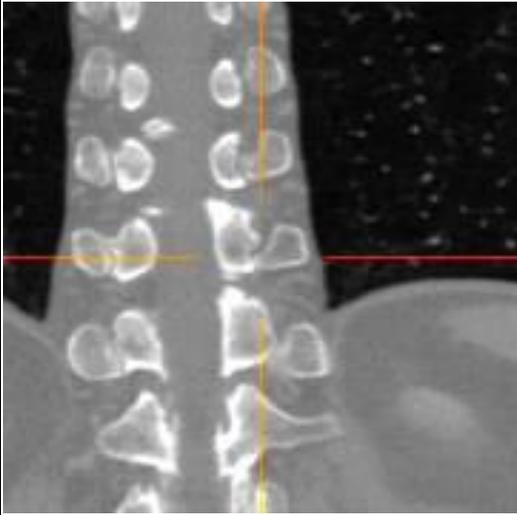
## Appendix B Anatomical landmark list

## 右第 11 肋骨近位端上縁 [R\_Rib11\_Prox]

	AX	COR
AX		
SAG		

R\_Rib-3\_Prox

## 左第 11 肋骨近位端上縁 [L\_Rib11\_Prox]

	AX	COR
AX		
SAG		

L\_Rib-3\_Prox

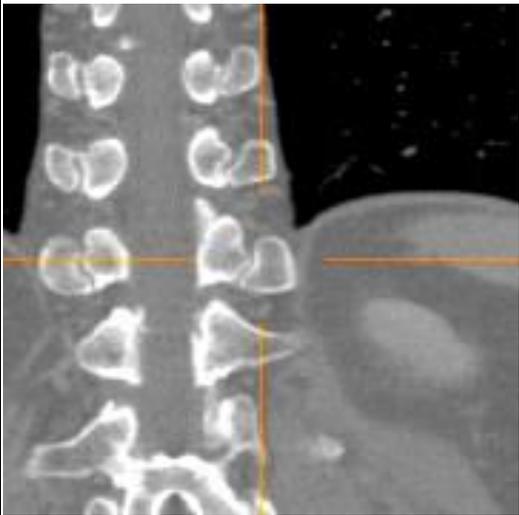
## Appendix B Anatomical landmark list

## 右第 12 肋骨近位端上縁 [R\_Rib12\_Prox]

	AX	COR
AX		
SAG		

第 12 肋骨は横突起との関係がやや他と異なる。付図を参照。

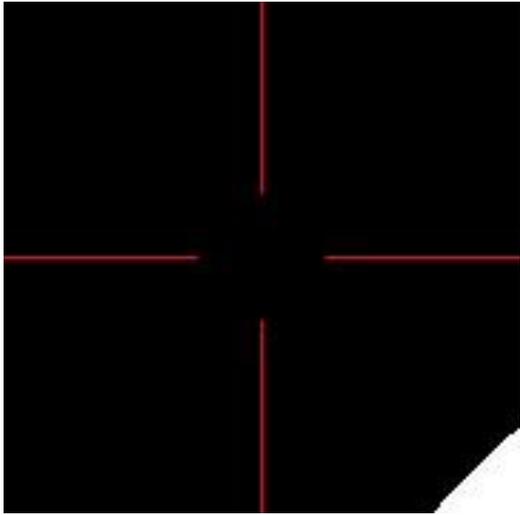
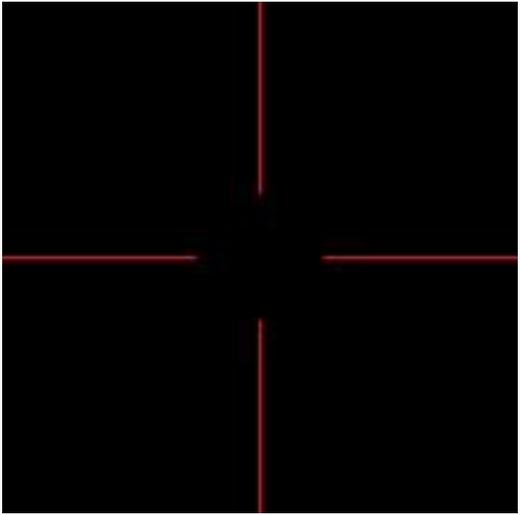
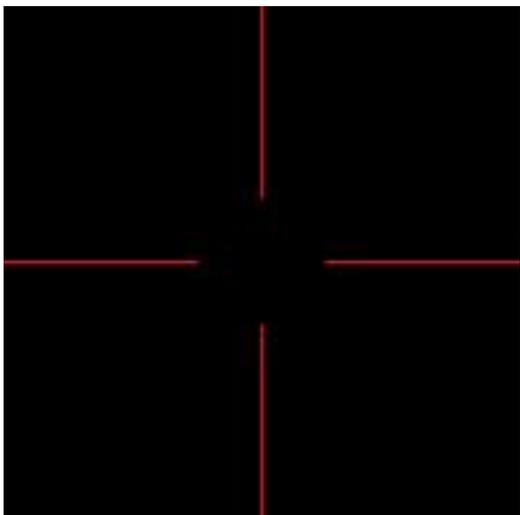
左第 12 肋骨近位端上縁 [L\_Rib12\_Prox]

	AX	COR
AX		
SAG		

L\_Rib-2\_Prox

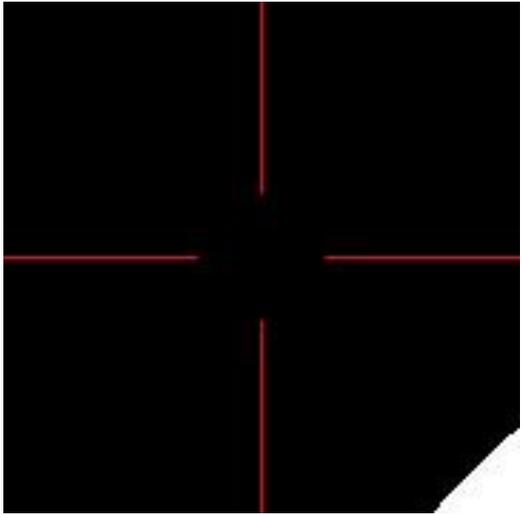
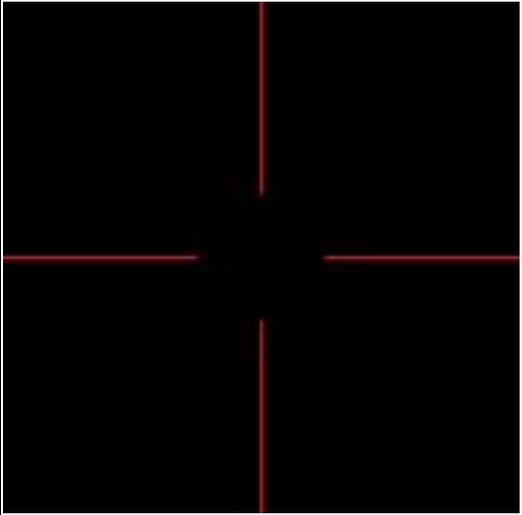
## Appendix B Anatomical landmark list

## 右第 13 肋骨近位端上縁 [R\_Rib13\_Prox]

	AX	COR
AX	 An axial view of the proximal end of the 13th rib. A red crosshair is centered on the rib head. The image is mostly black with a white triangular artifact in the bottom right corner.	 A coronal view of the proximal end of the 13th rib. A red crosshair is centered on the rib head. The image is mostly black.
SAG	 A sagittal view of the proximal end of the 13th rib. A red crosshair is centered on the rib head. The image is mostly black.	

(破格時のみ)

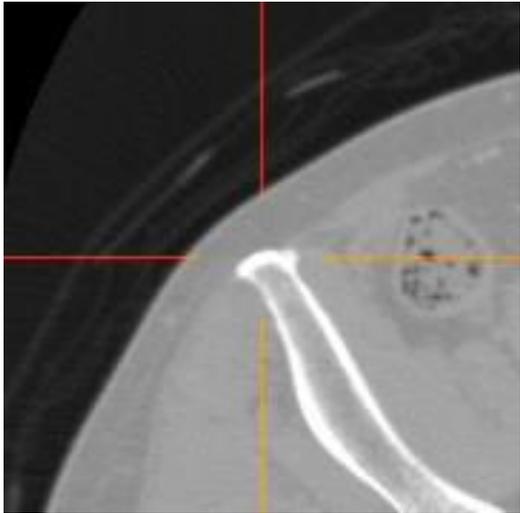
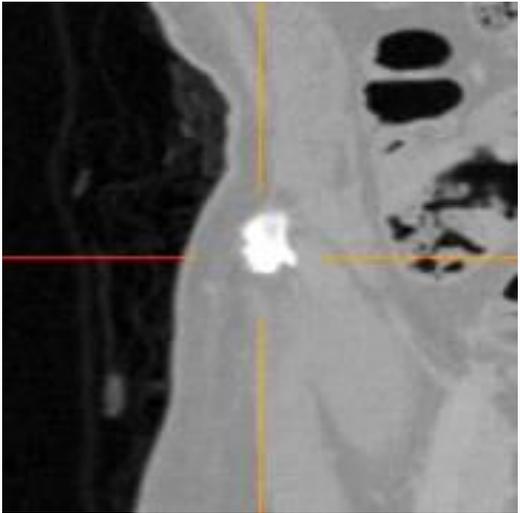
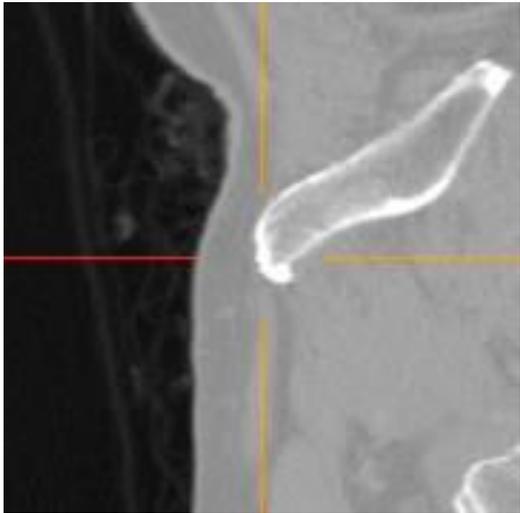
## 左第 13 肋骨近位端上縁 [L\_Rib13\_Prox]

	AX	COR
AX	 An axial CT scan image of the proximal end of the 13th rib. A red crosshair is overlaid on the image, with the vertical line passing through the center of the rib's cross-section and the horizontal lines intersecting it.	 A coronal CT scan image of the proximal end of the 13th rib. A red crosshair is overlaid on the image, with the vertical line passing through the center of the rib's length and the horizontal lines intersecting it.
SAG	 A sagittal CT scan image of the proximal end of the 13th rib. A red crosshair is overlaid on the image, with the vertical line passing through the center of the rib's length and the horizontal lines intersecting it.	

(破格時のみ)

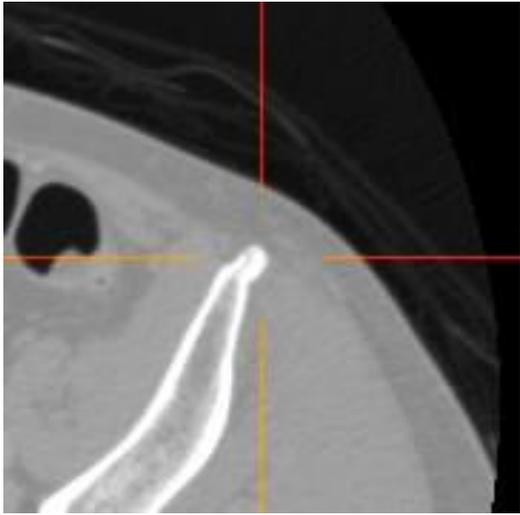
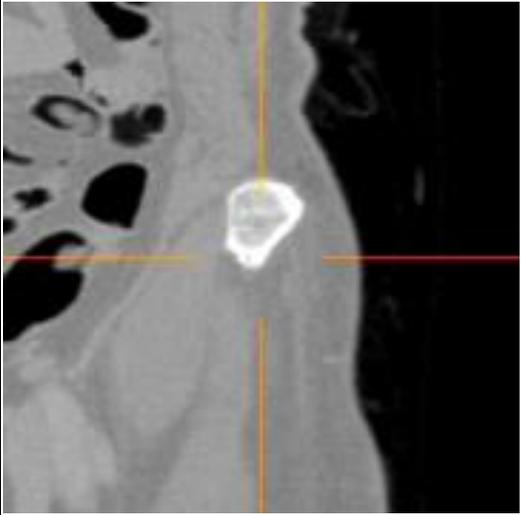
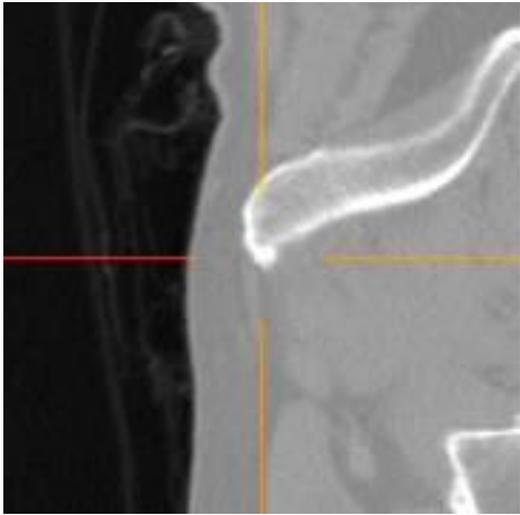
## Appendix B Anatomical landmark list

## 右上前腸骨棘 [R\_AntSupIliacSp] (rt. anterior superior iliac spine)

	AX	COR
AX		
SAG		

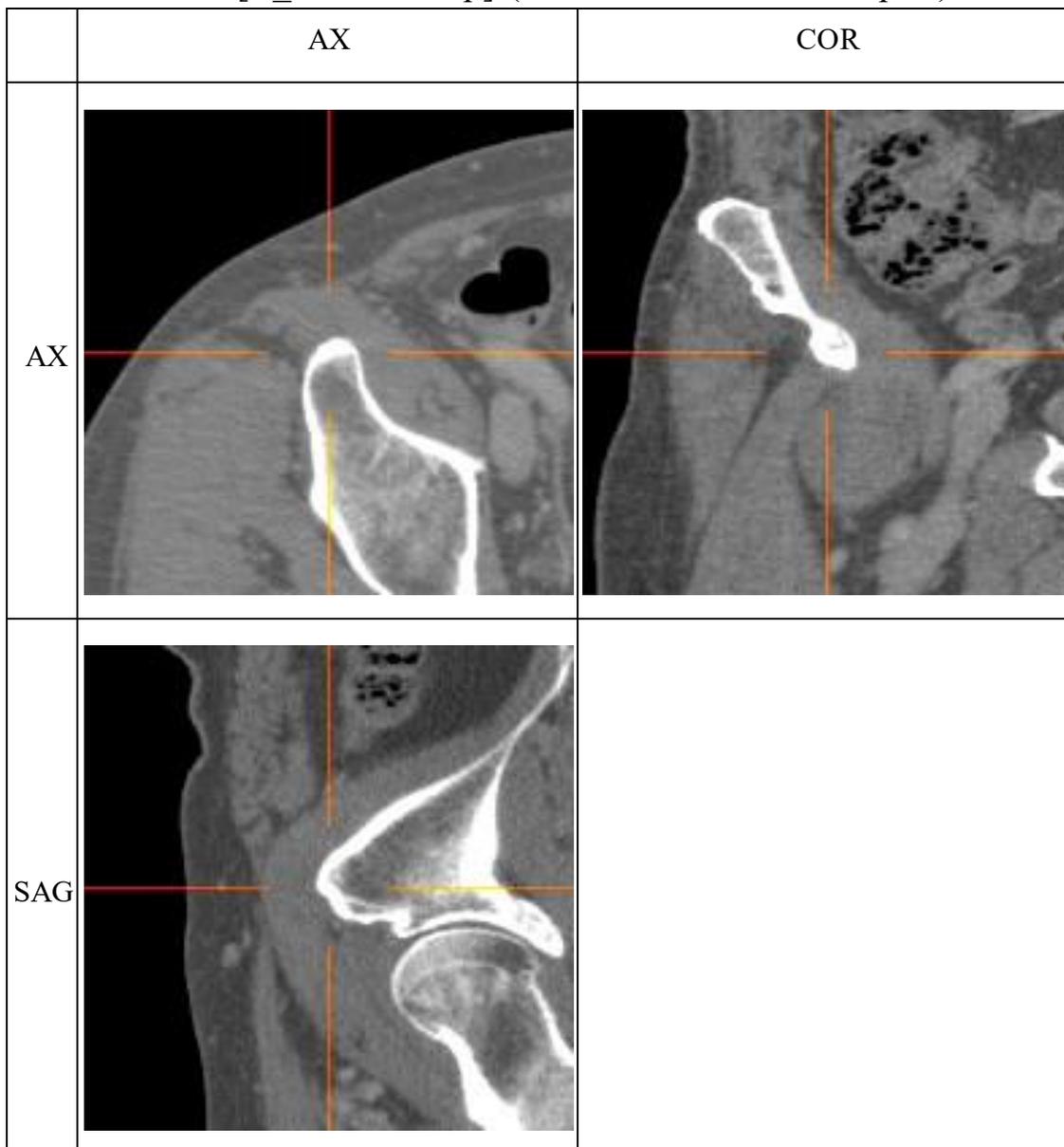
筋の起始部であることも手掛かりに点を決めること。

## 左上前腸骨棘 [L\_AntSupIliacSp] (lt. anterior superior iliac spine)

	AX	COR
AX		
SAG		

## Appendix B Anatomical landmark list

右下前腸骨棘 [R\_AntInfIliacSp] (rt. anterior inferior iliac spine)



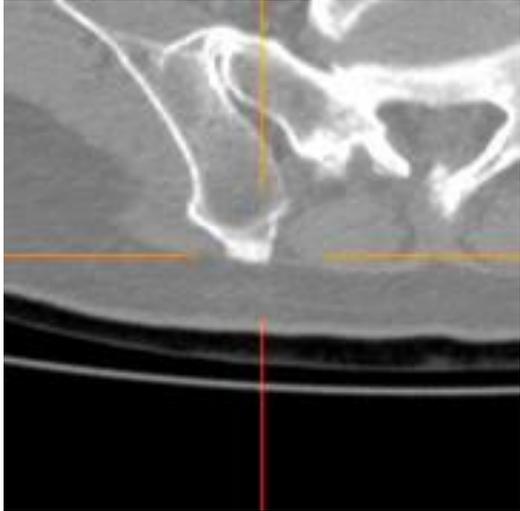
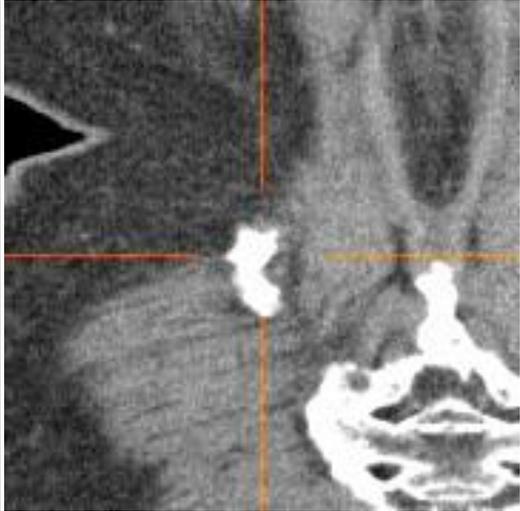
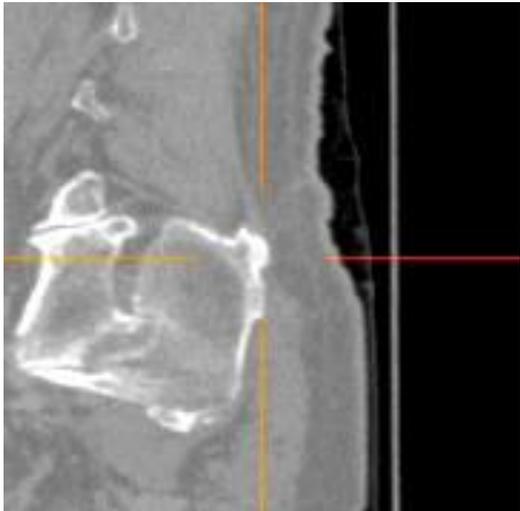
突起としては微妙なことが多いが、臼蓋直上の小さな構造として同定できる。

左下前腸骨棘 [L\_AntInflIiacSp] (lt. anterior inferior iliac spine)

	AX	COR
AX		
SAG		

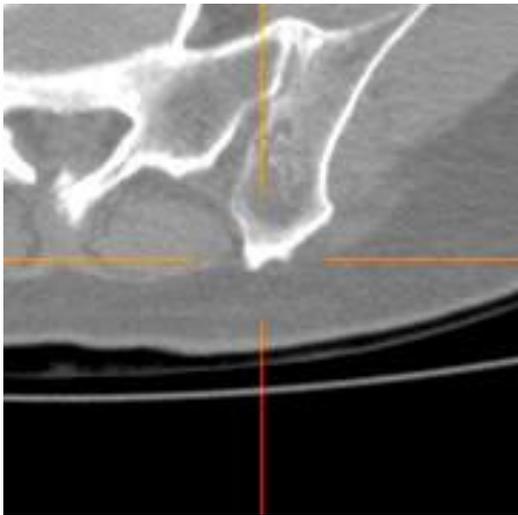
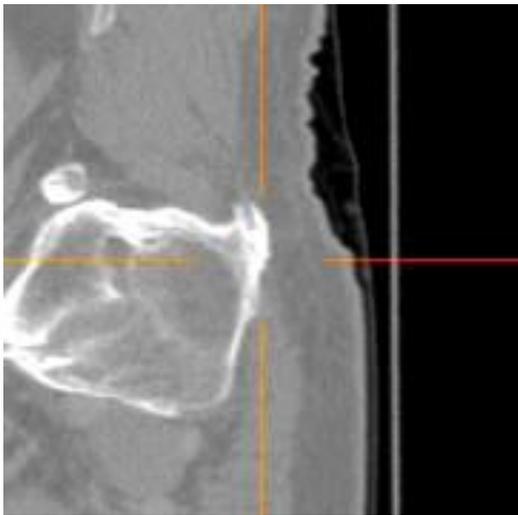
## Appendix B Anatomical landmark list

## 右上後腸骨棘 [R\_PostSupIliacSp] (rt. posterior superior iliac spine)

	AX	COR
AX		
SAG		

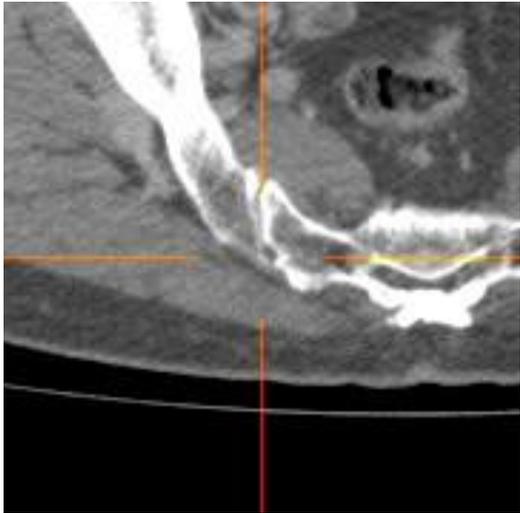
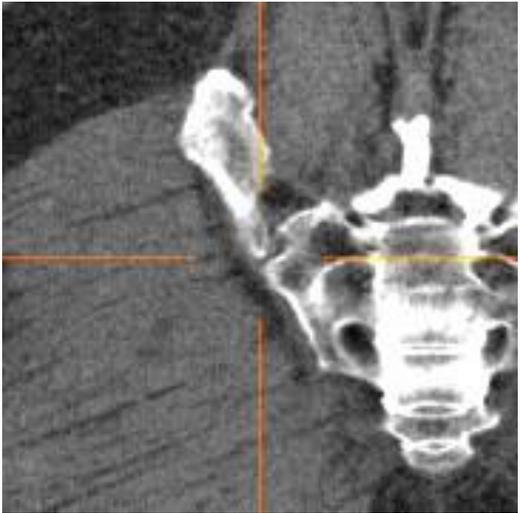
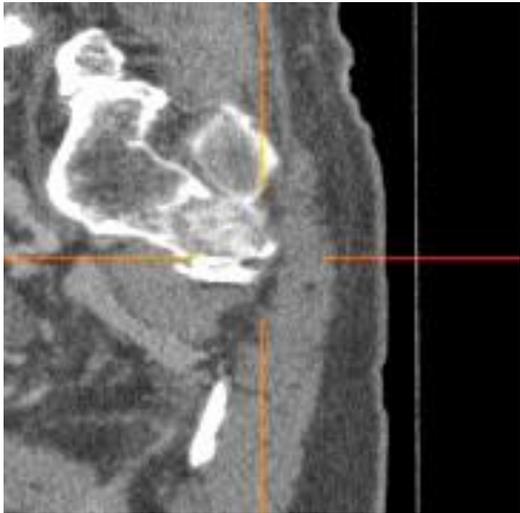
一点として同定するのはかなり難しい。付図を参考に。

## 左上後腸骨棘 [L\_PostSupIliacSp] (lt. posterior superior iliac spine)

	AX	COR
AX		
SAG		

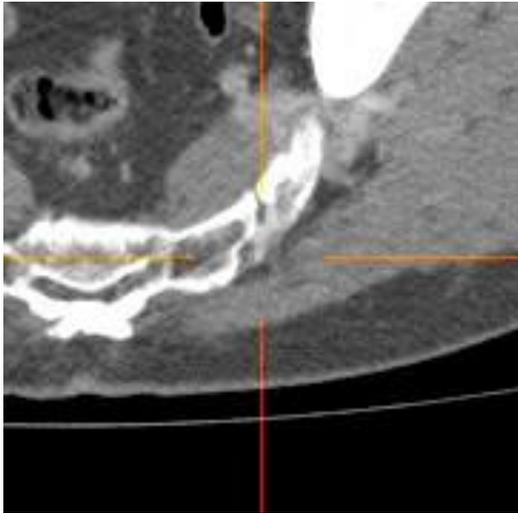
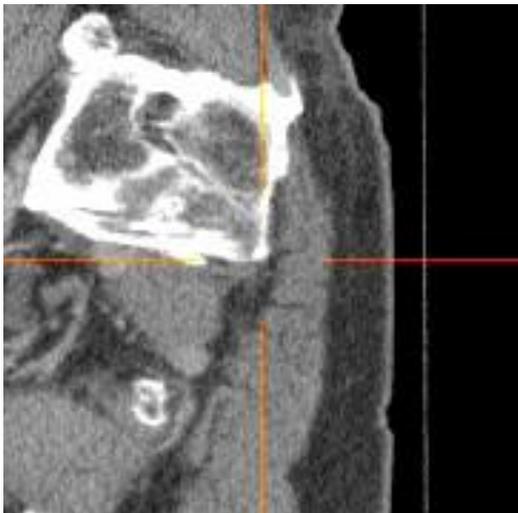
## Appendix B Anatomical landmark list

## 右下後腸骨棘 [R\_PostInfIliacSp] (rt. posterior inferior iliac spine)

	AX	COR
AX		
SAG		

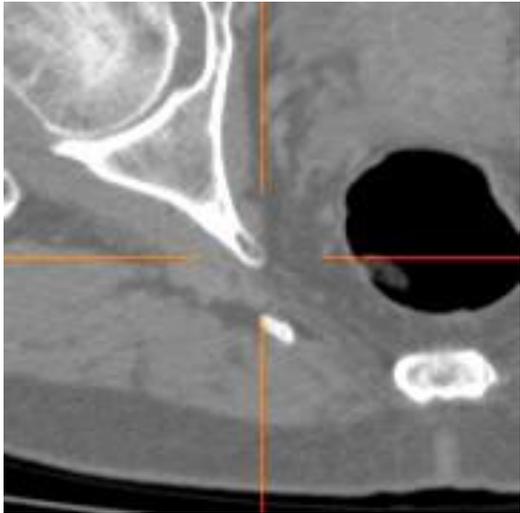
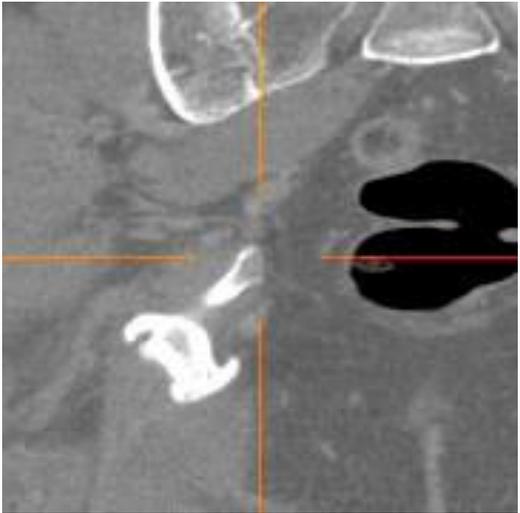
本研究では仙腸関節の下端あたりを目印に定めている。矢状断での形態を参考に。

## 左下後腸骨棘 [L\_PostInfIliacSp] (lt. posterior inferior iliac spine)

	AX	COR
AX		
SAG		

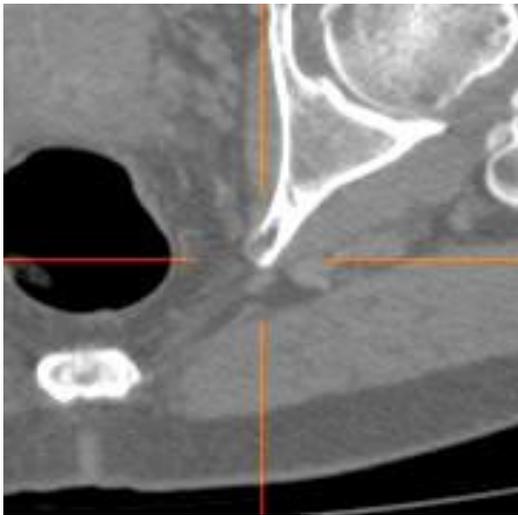
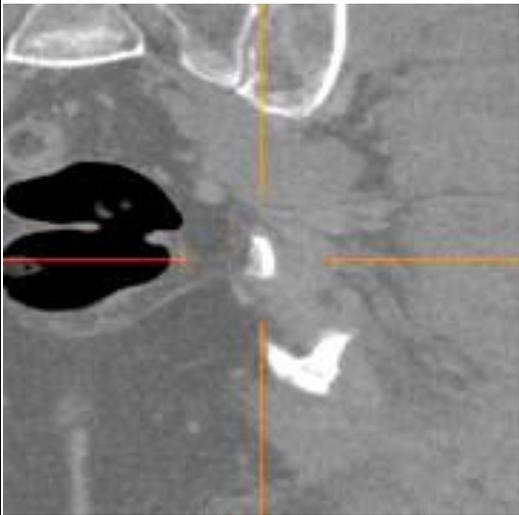
## Appendix B Anatomical landmark list

## 右坐骨棘 [R\_IschiaticSp] (rt. ischiatic spine)

	AX	COR
AX		
SAG		

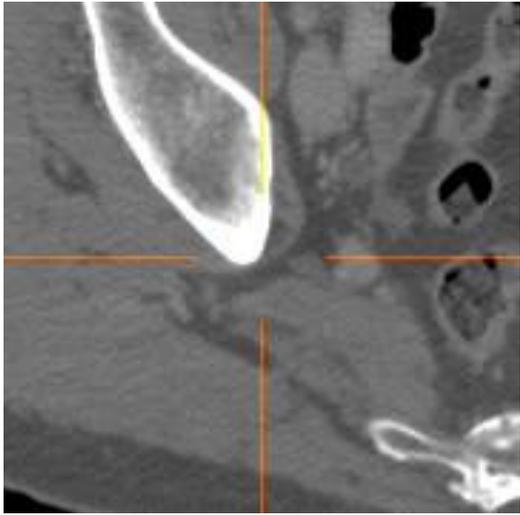
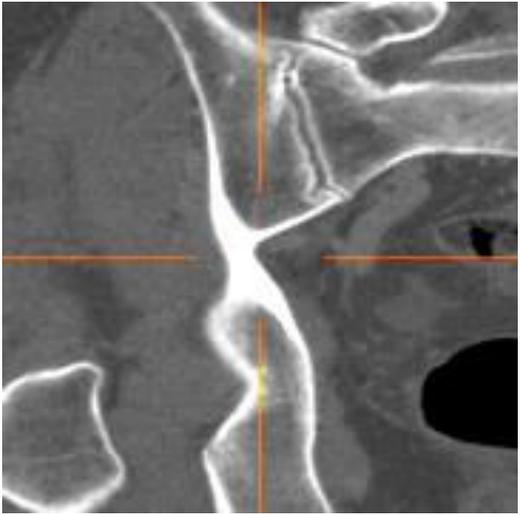
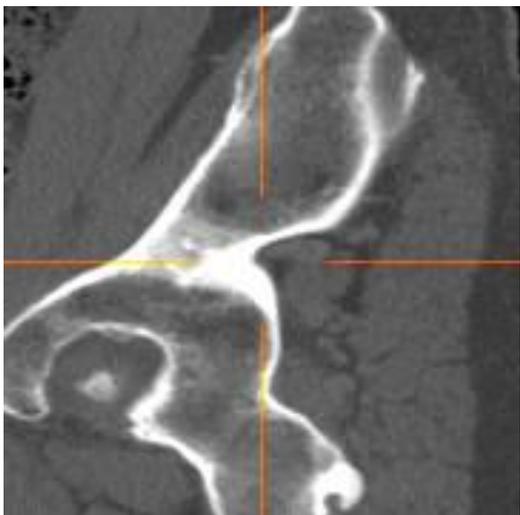
たいてい、非常に骨皮質が薄い小さな突起として同定される。その先端をとる。

## 左坐骨棘 [L\_IschiaticSp] (lt. ischiatic spine)

	AX	COR
AX		
SAG		

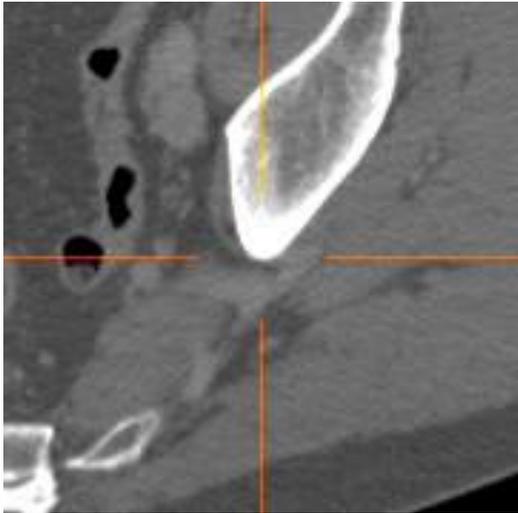
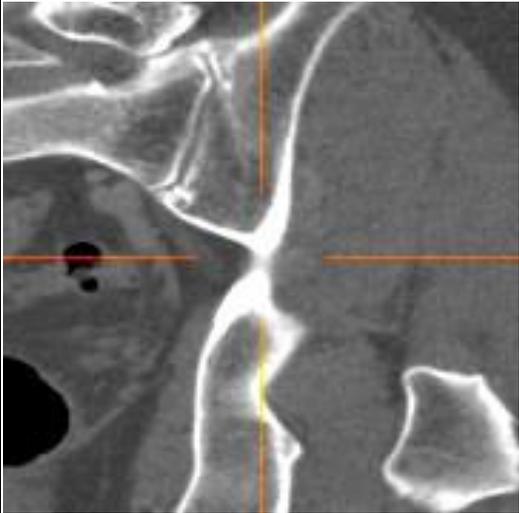
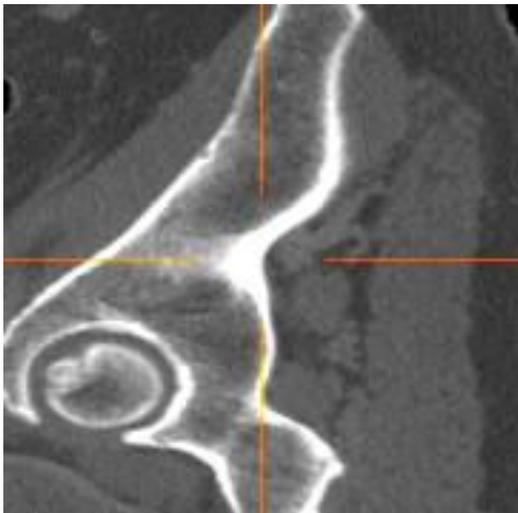
## Appendix B Anatomical landmark list

## 右大坐骨切痕 [R\_G\_SciaticNotch] (rt. greater sciatic notch)

	AX	COR
AX		
SAG		

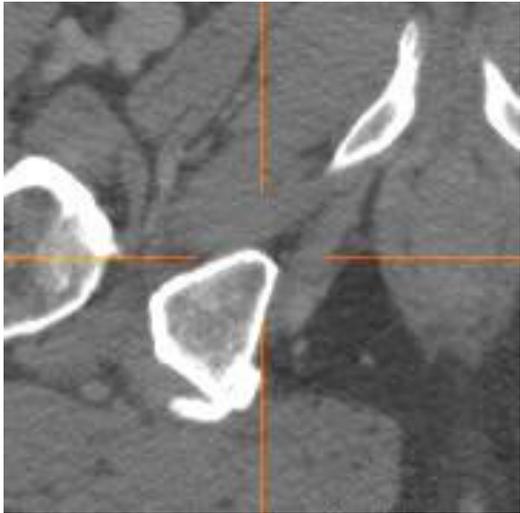
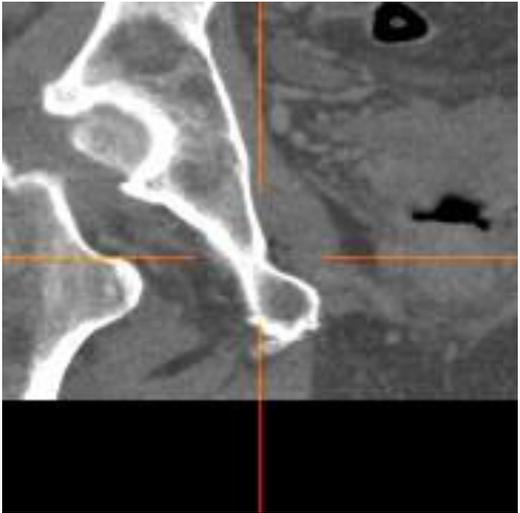
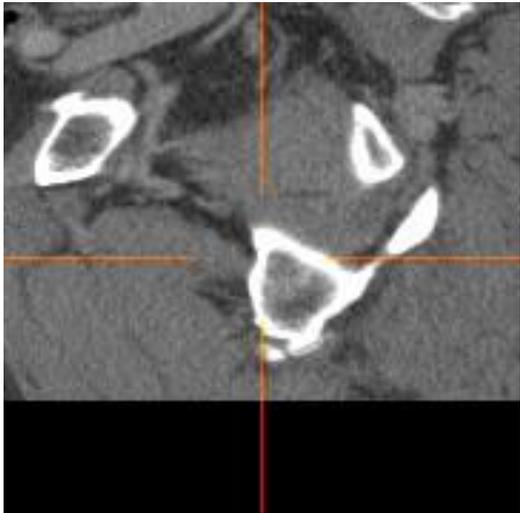
坐骨の切れ込みの最深部の骨皮質をとること。

## 左大坐骨切痕 [L\_G\_SciaticNotch] (lt. greater sciatic notch)

	AX	COR
AX		
SAG		

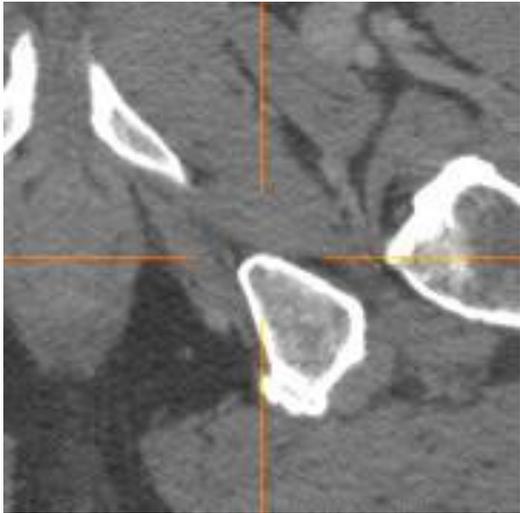
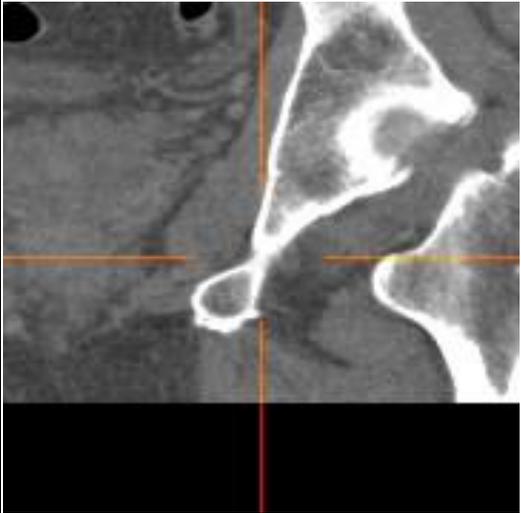
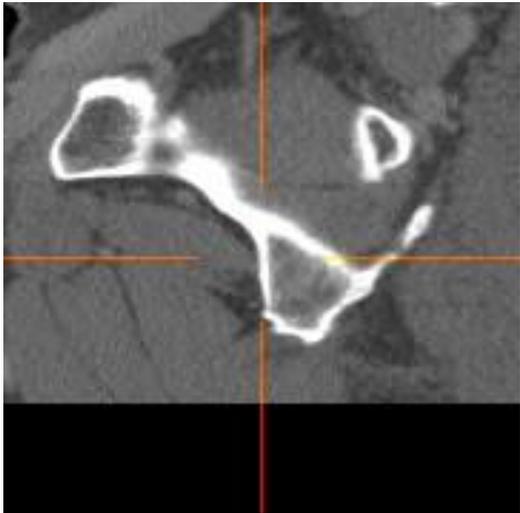
## Appendix B Anatomical landmark list

## 右閉鎖孔外側縁 [R\_Obturator\_Lat] (rt. obturator lateral margin)

	AX	COR
AX		
SAG		

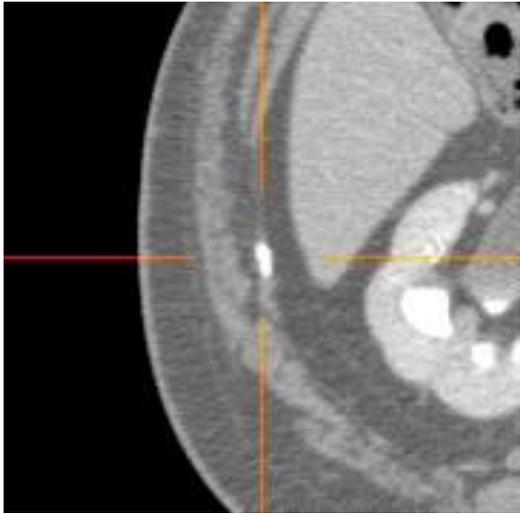
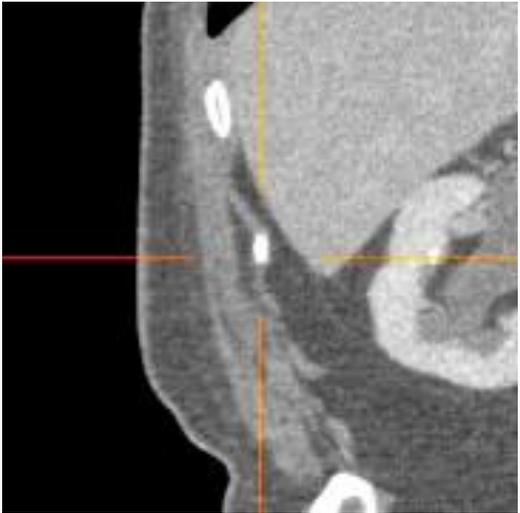
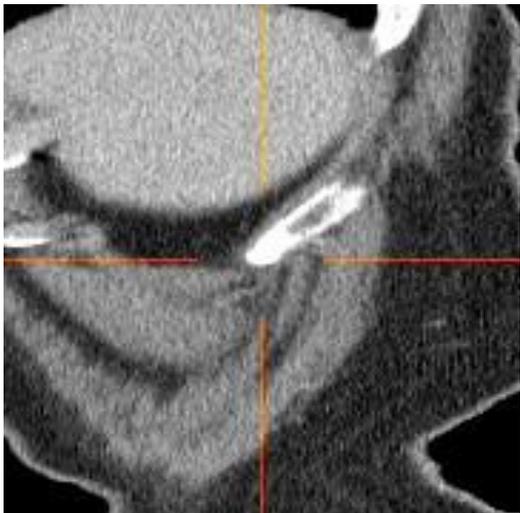
本研究では、冠状断でみて閉鎖孔の後縁、前後の坐骨枝が合するところを便宜的にとっている。

## 左閉鎖孔外側縁 [L\_Obturator\_Lat] (lt. obturator lateral margin)

	AX	COR
AX		
SAG		

## Appendix B Anatomical landmark list

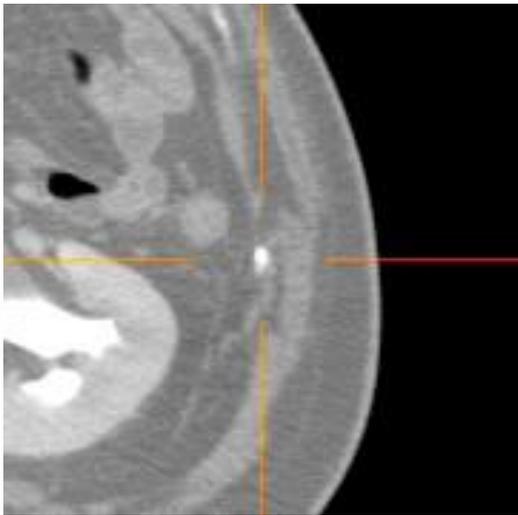
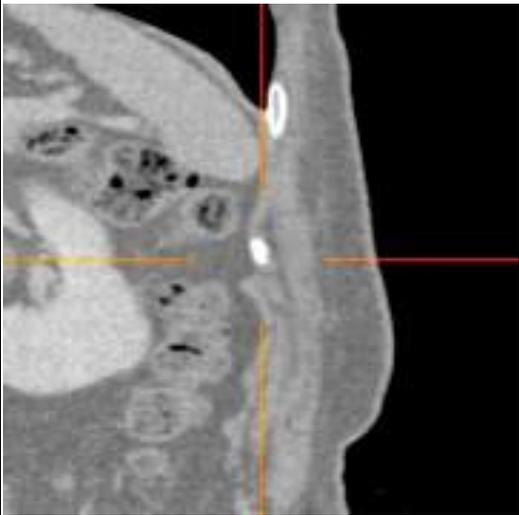
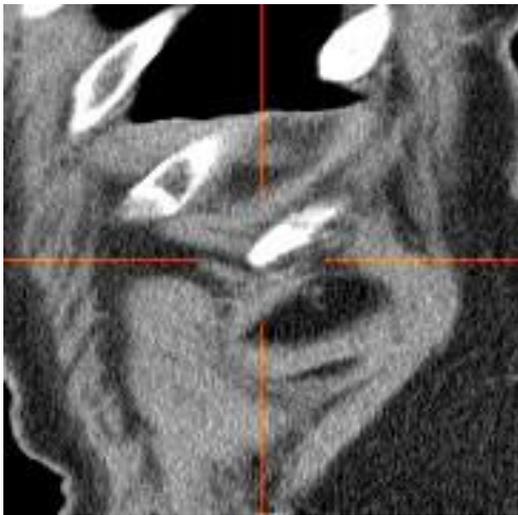
## 右第 2 最下位肋骨先端 [R\_11thRib\_Tip]

	AX	COR
AX		
SAG		

肋骨の数に破格があるときは、その破格に沿って下から 2 番目の肋骨の先端をとる。

肋骨と腰椎横突起の区別がつけがたいときは相談してください。

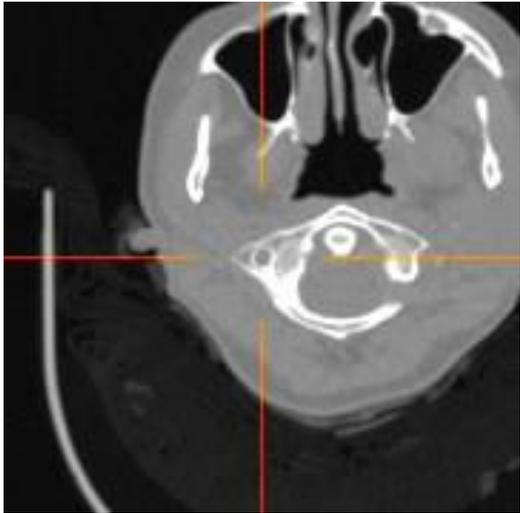
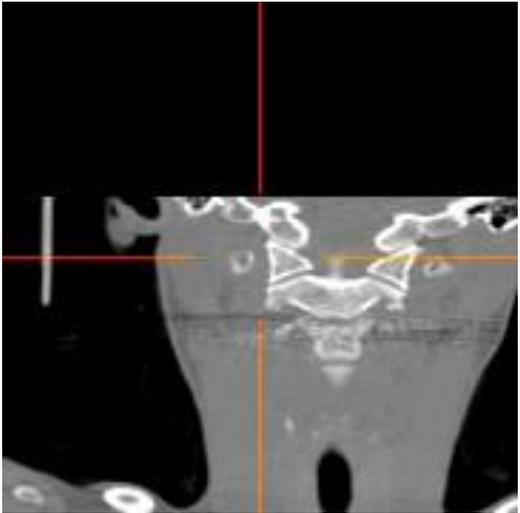
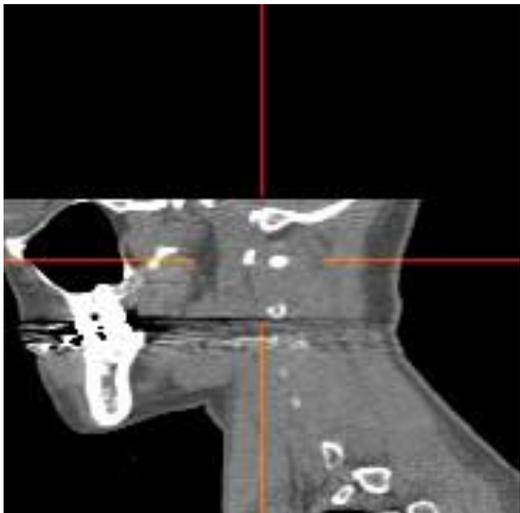
## 左第 2 最下位肋骨先端 [L\_11thRib\_Tip]

	AX	COR
AX		
SAG		

L\_Rib-2\_Tip

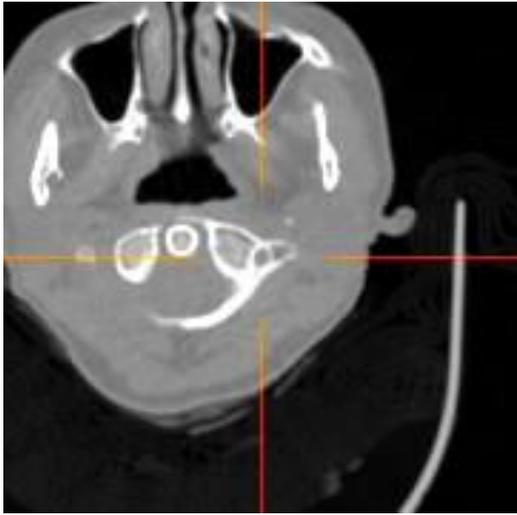
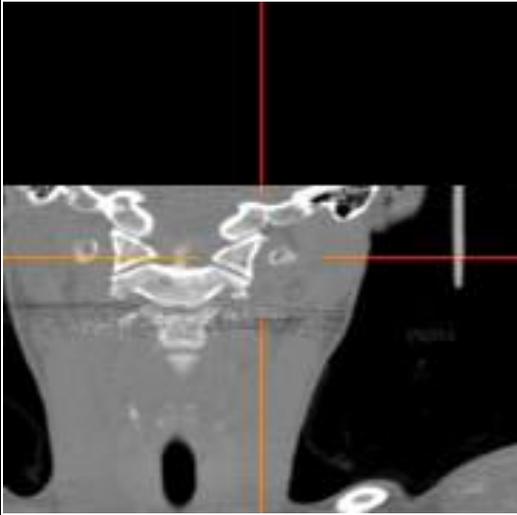
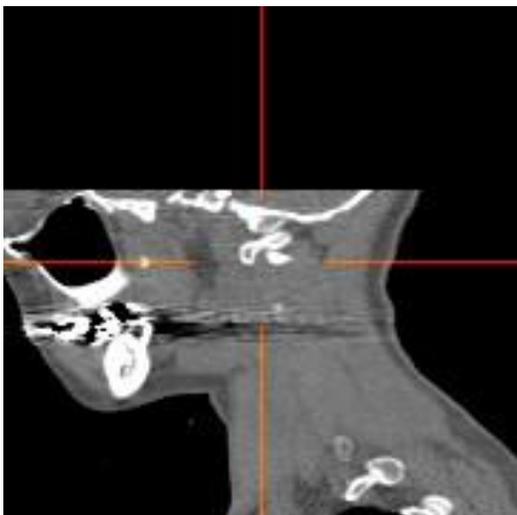
## Appendix B Anatomical landmark list

## 右第 1 頸椎横突孔 [R\_C1\_TransvForamem] (rt. C1 transverse foramen)

	AX	COR
AX		
SAG		

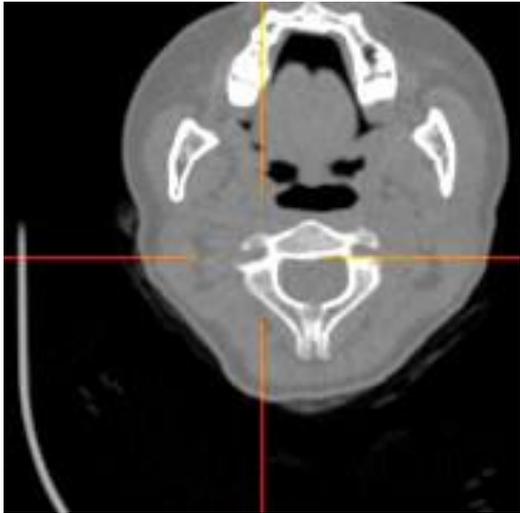
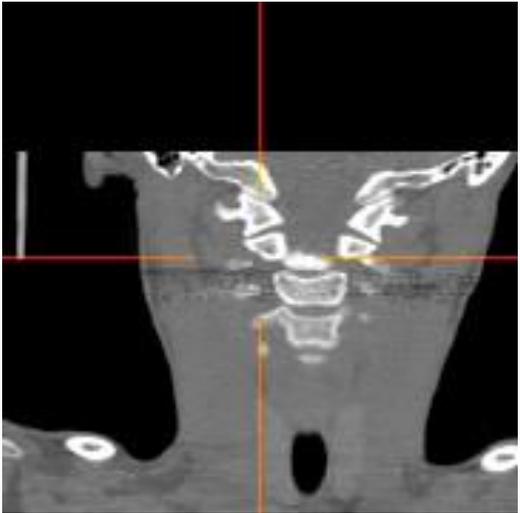
横突孔の孔の中心をとる。椎骨動脈低形成などのため孔がないときはその旨破格として記載を。

## 左第 1 頸椎横突孔 [L\_C1\_TransvForamem] (lt. C1 transverse foramen)

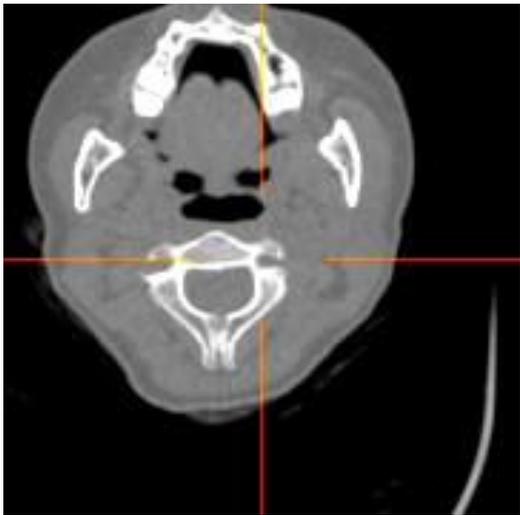
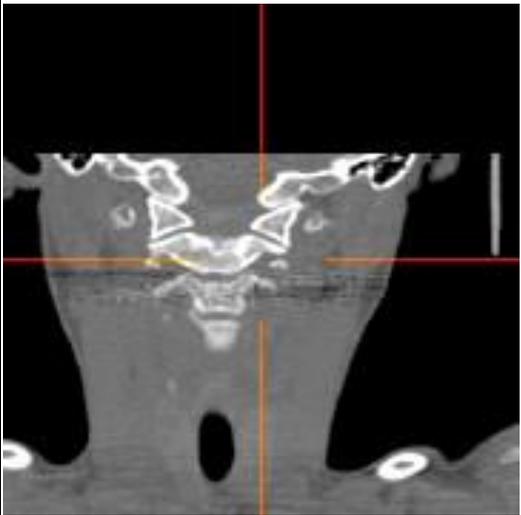
	AX	COR
AX		
SAG		

## Appendix B Anatomical landmark list

## 右第 2 頸椎横突孔 [R\_C2\_TransvForamem]

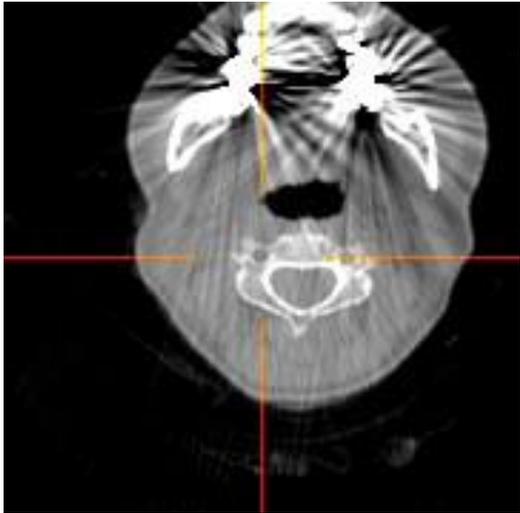
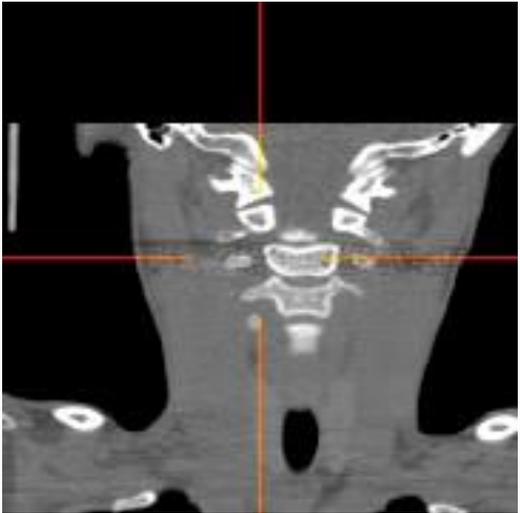
	AX	COR
AX		
SAG		

## 左第2頸椎横突孔 [L\_C2\_TransvForamem]

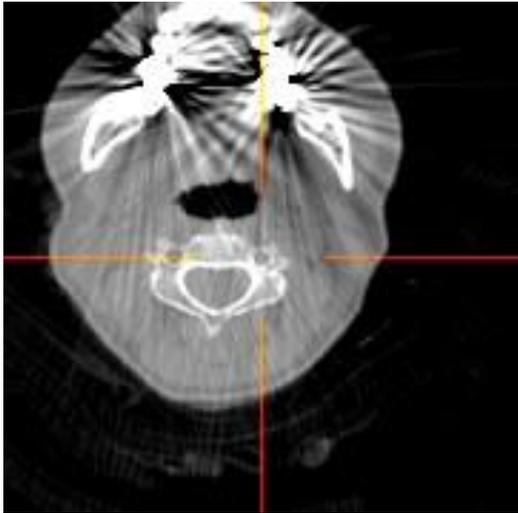
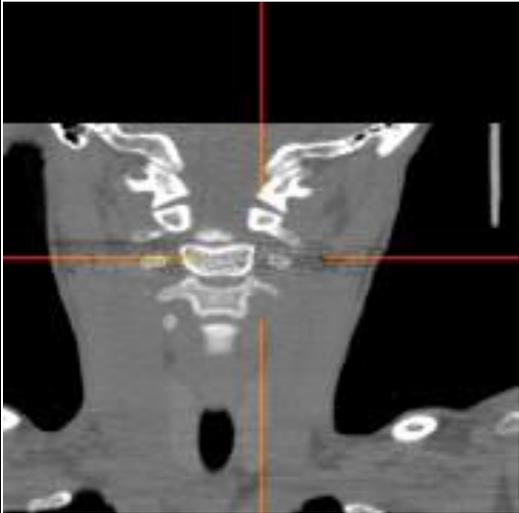
	AX	COR
AX		
SAG		

## Appendix B Anatomical landmark list

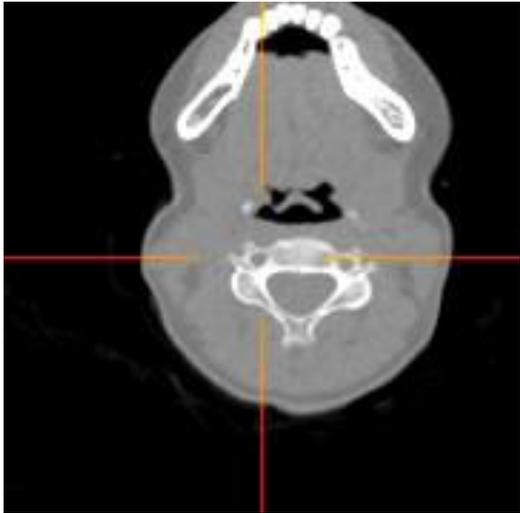
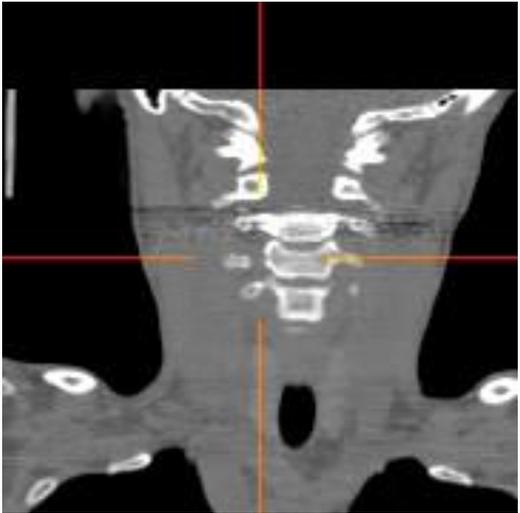
## 右第 3 頸椎横突孔 [R\_C3\_TransvForamem]

	AX	COR
AX		
SAG		

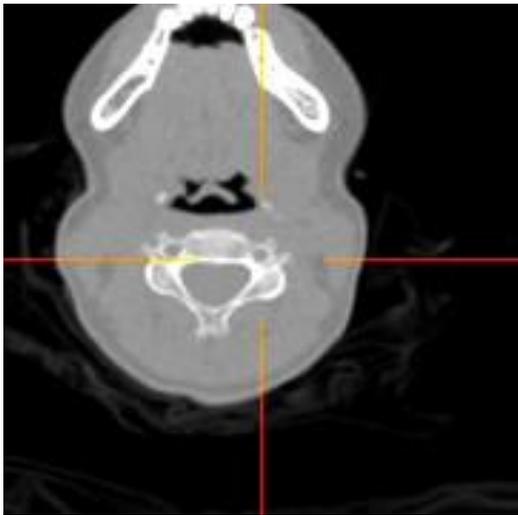
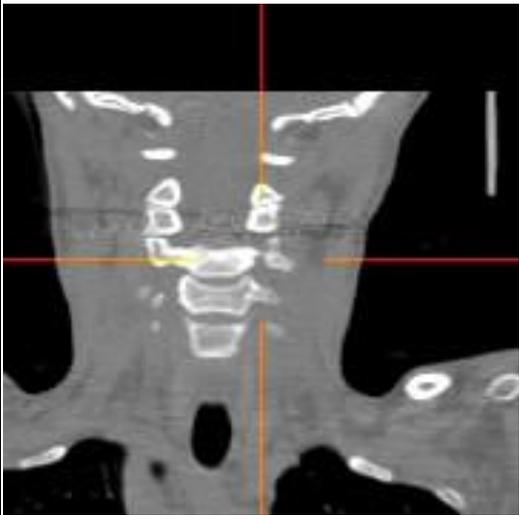
## 左第3頸椎横突孔 [L\_C3\_TransvForamem]

	AX	COR
AX		
SAG		

## 右第 4 頸椎横突孔 [R\_C4\_TransvForamem]

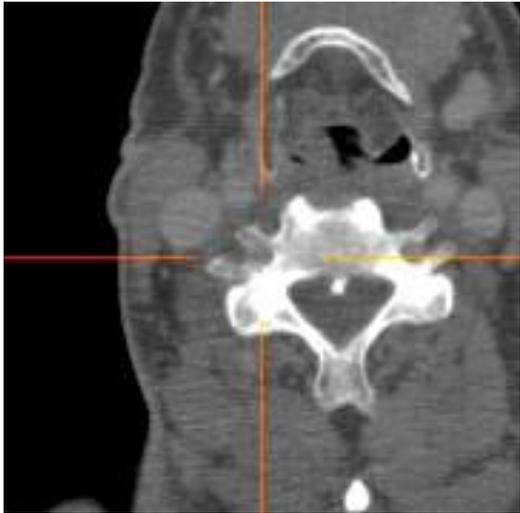
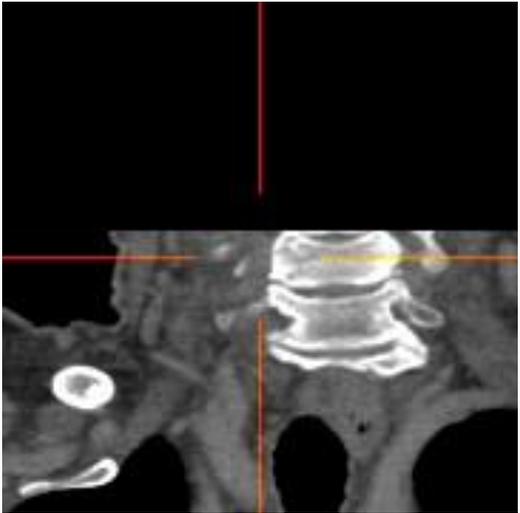
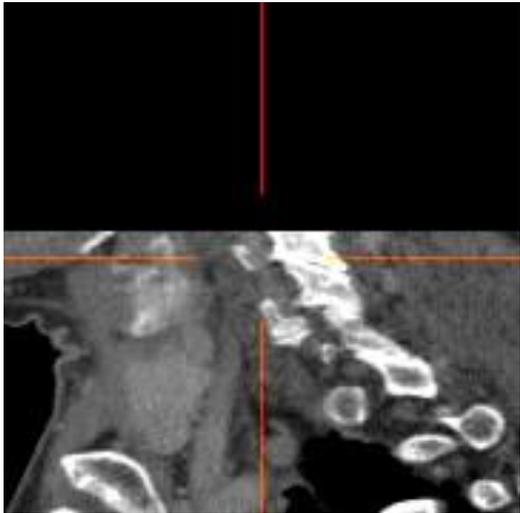
	AX	COR
AX		
SAG		

## 左第 4 頸椎横突孔 [L\_C4\_TransvForamem]

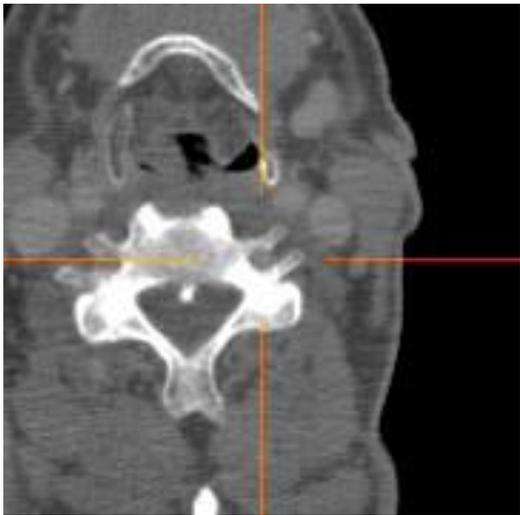
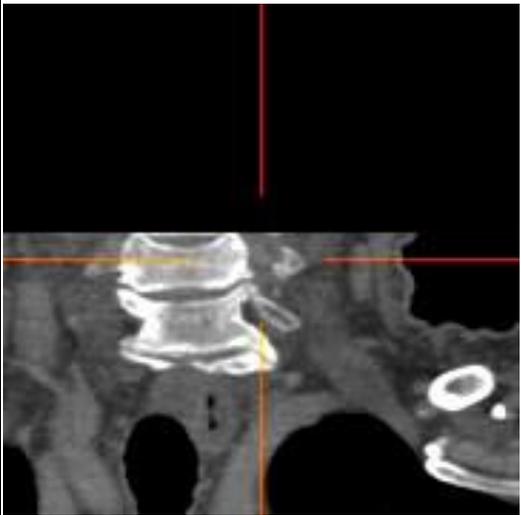
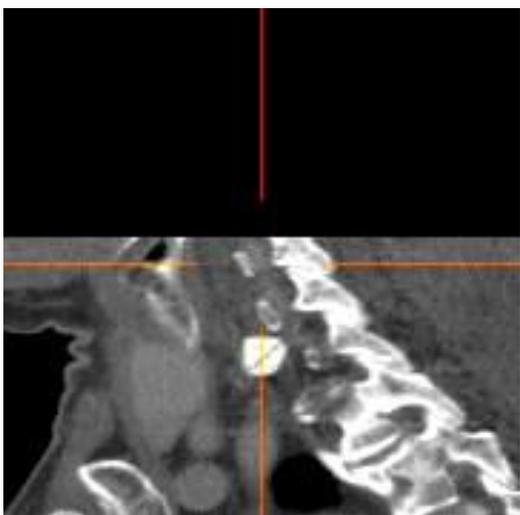
	AX	COR
AX		
SAG		

## Appendix B Anatomical landmark list

## 右第 5 頸椎横突孔 [R\_C5\_TransvForamem]

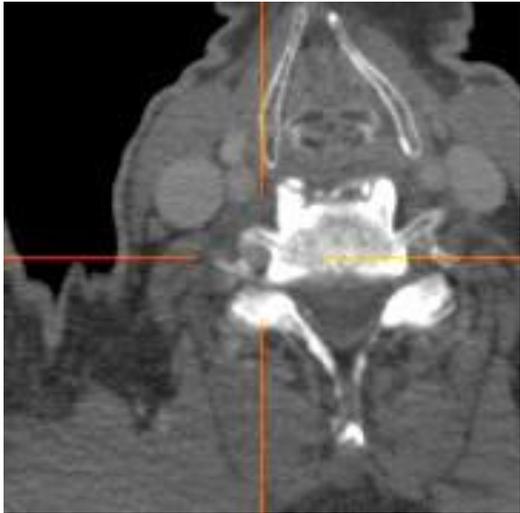
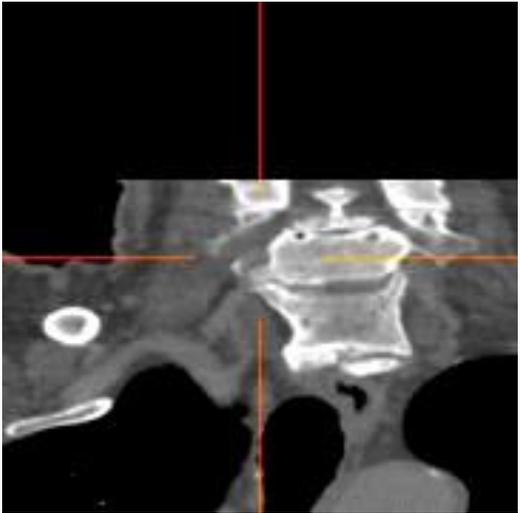
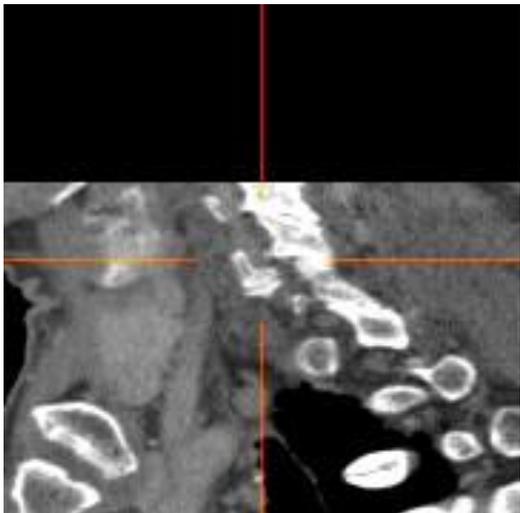
	AX	COR
AX		
SAG		

## 左第 5 頸椎横突孔 [L\_C5\_TransvForamem]

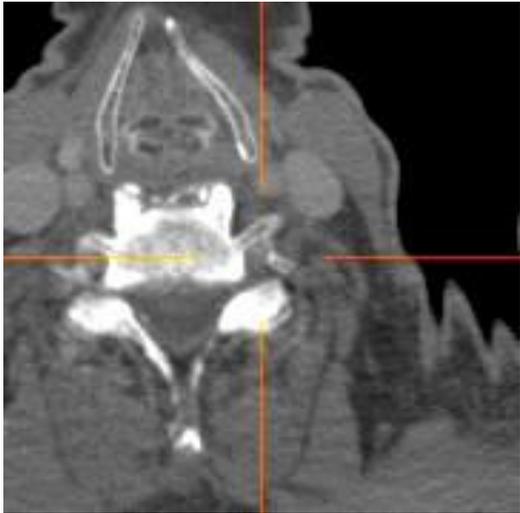
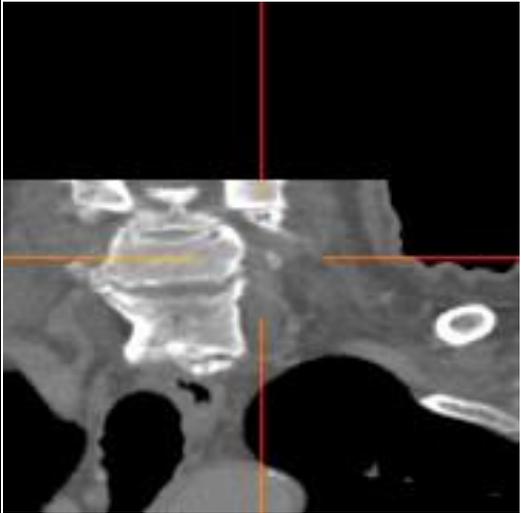
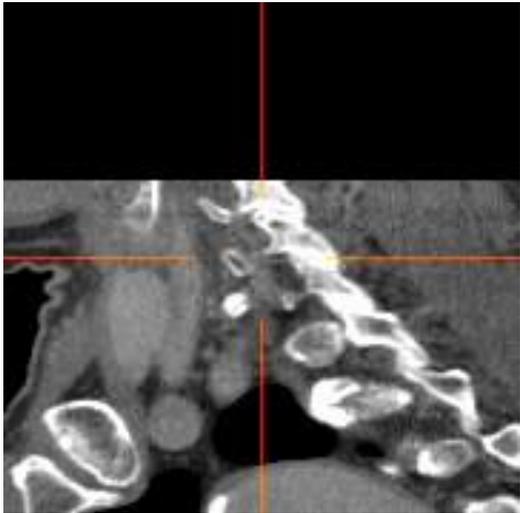
	AX	COR
AX		
SAG		

## Appendix B Anatomical landmark list

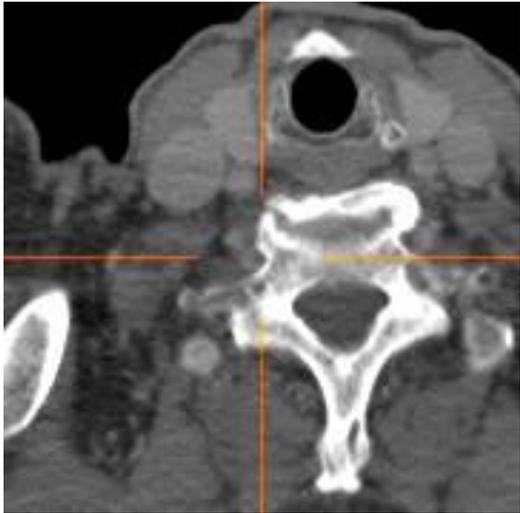
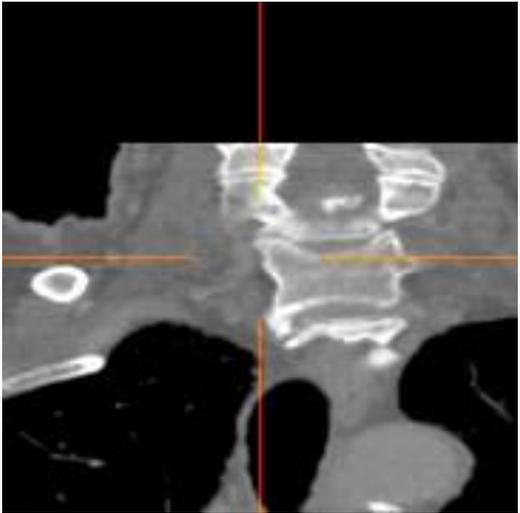
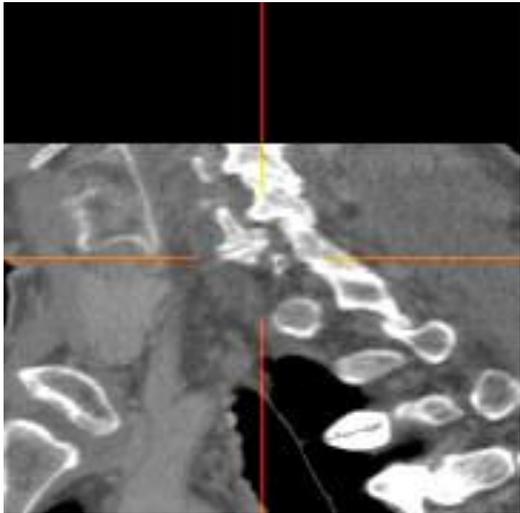
## 右第 6 頸椎横突孔 [R\_C6\_TransvForamem]

	AX	COR
AX		
SAG		

## 左第 6 頸椎横突孔 [L\_C6\_TransvForamem]

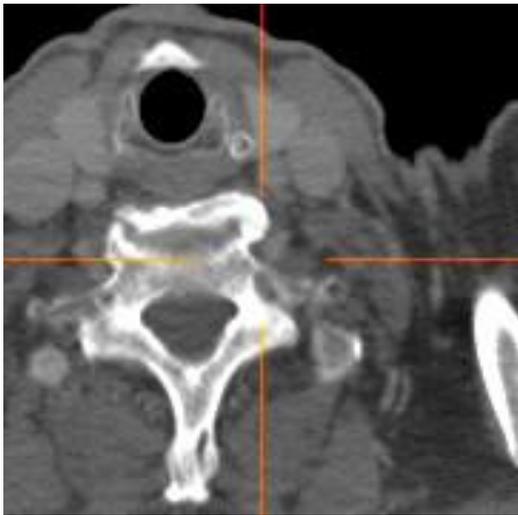
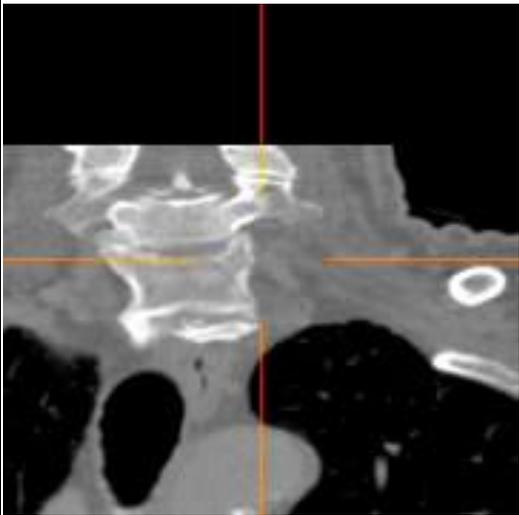
	AX	COR
AX		
SAG		

## 右第 7 頸椎横突孔 [R\_C7\_TransvForamem]

	AX	COR
AX		
SAG		

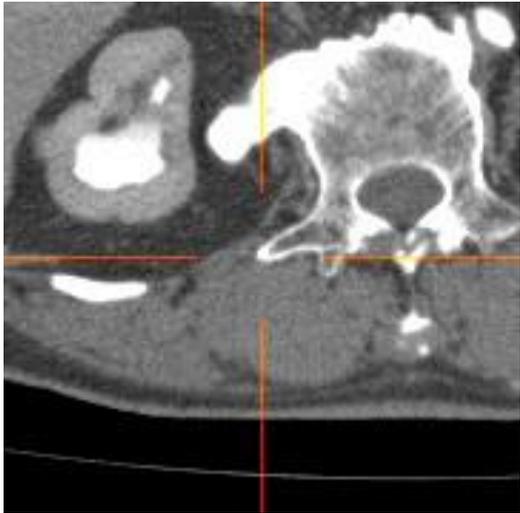
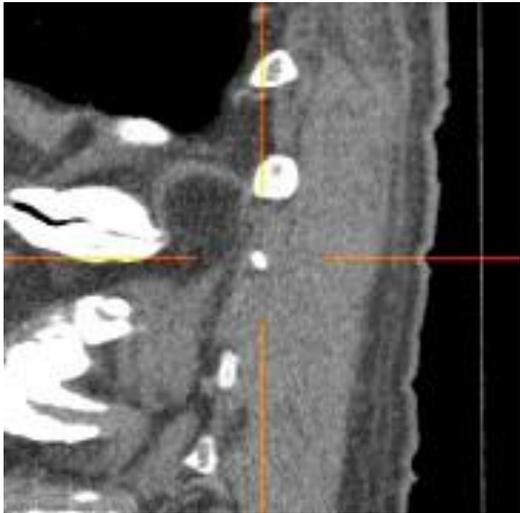
第 7 頸椎には横突孔と呼ぶべき構造がないことが多い。代わりに横突起にあたるとても細い構造が 2 本同定されるはずなので、その間の根本近くをとる。

## 左第 7 頸椎横突孔 [L\_C7\_TransvForamem]

	AX	COR
AX		
SAG		

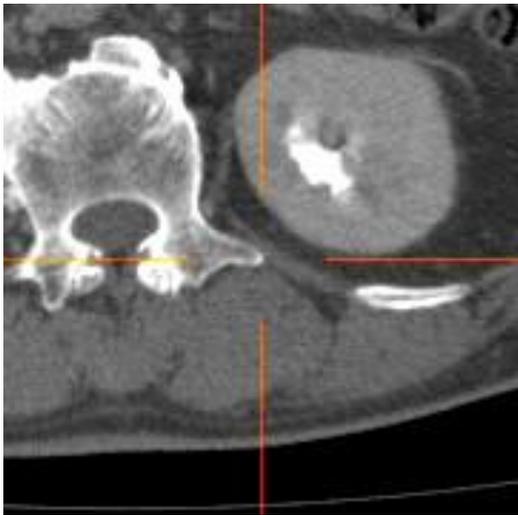
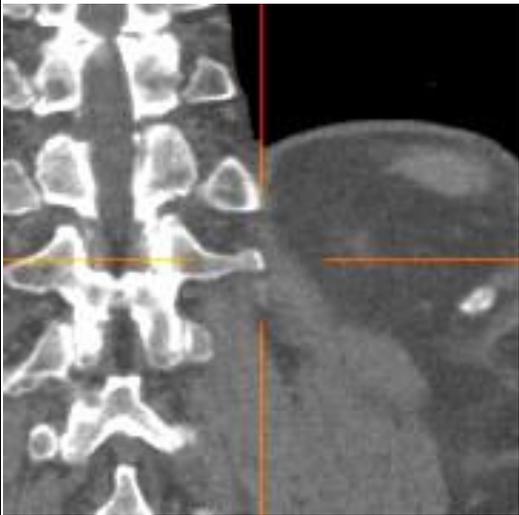
## Appendix B Anatomical landmark list

右第 1 腰椎横突起先端 [R\_L1\_TransverseP] (rt. L1 transverse process tip)

	AX	COR
AX		
SAG		

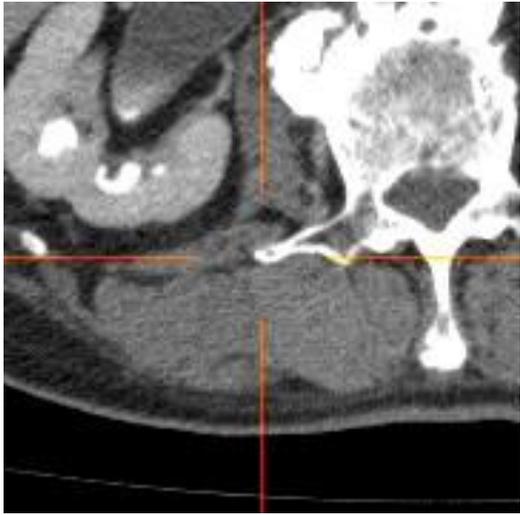
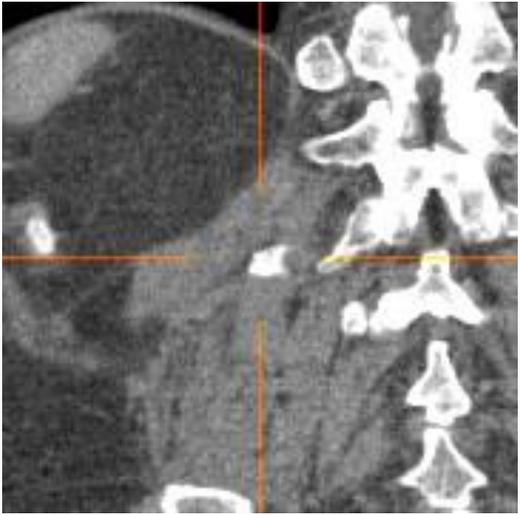
横突起の先端の骨皮質上をとる。

左第 1 腰椎横突起先端 [L\_L1\_TransverseP] (lt. L1 transverse process tip)

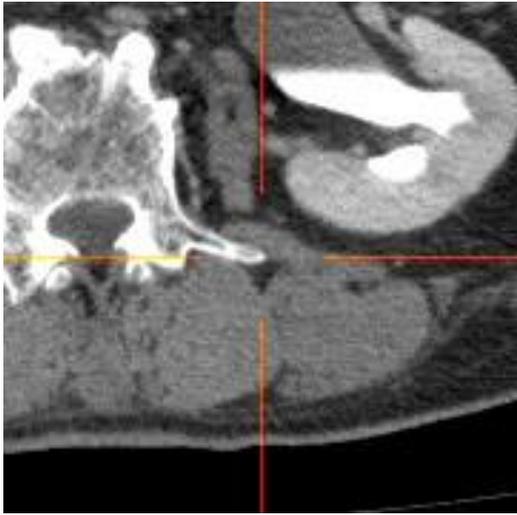
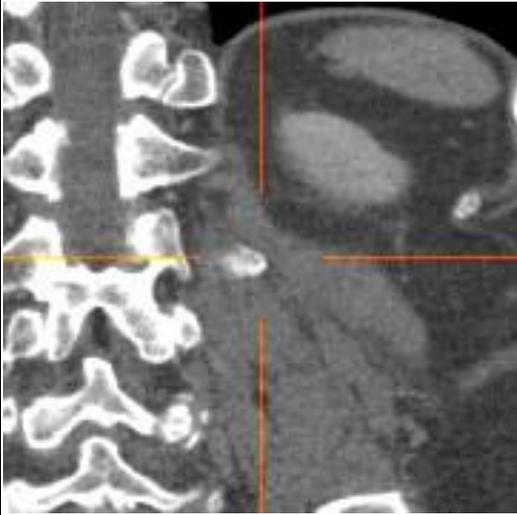
	AX	COR
AX		
SAG		

## Appendix B Anatomical landmark list

## 右第 2 腰椎横突起先端 [R\_L2\_TransverseP]

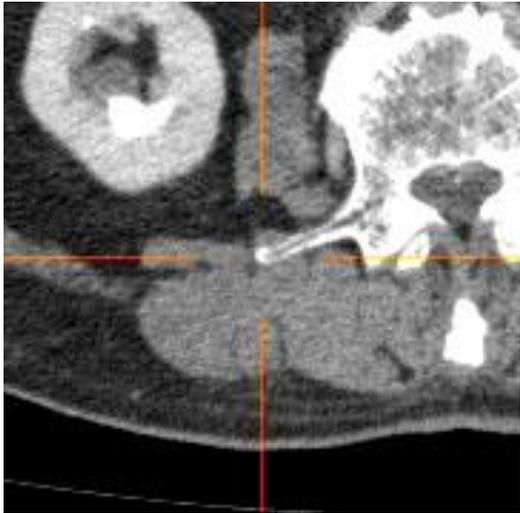
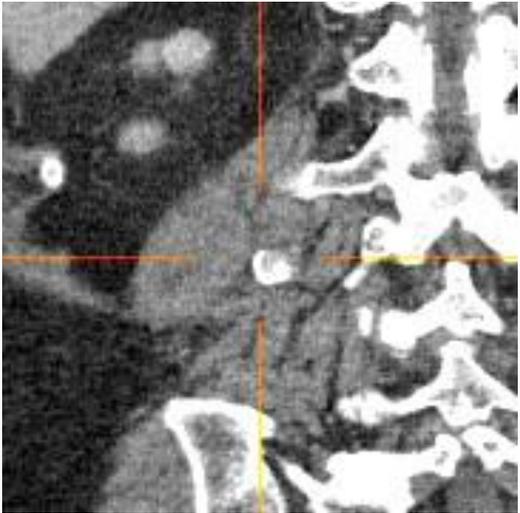
	AX	COR
AX		
SAG		

## 左第 2 腰椎横突起先端 [L\_L2\_TransverseP]

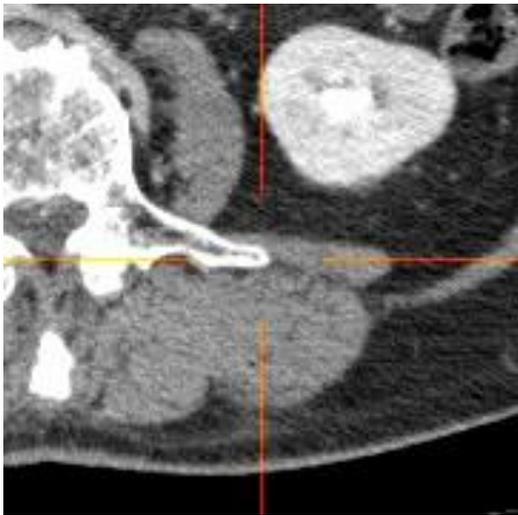
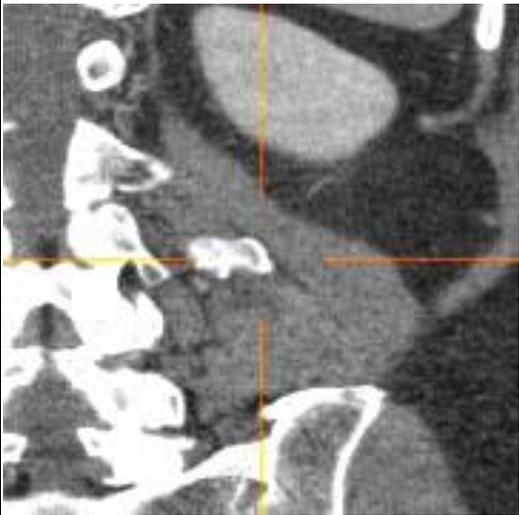
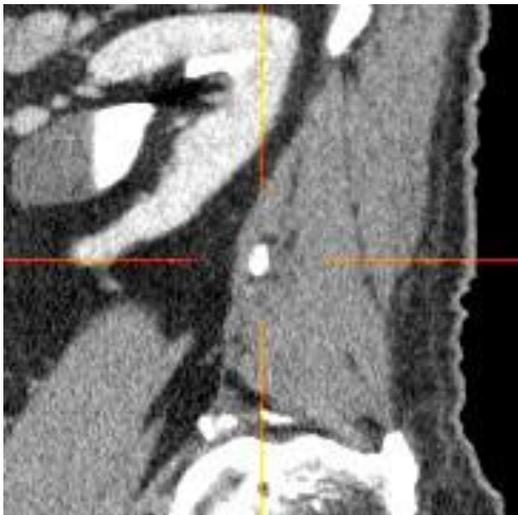
	AX	COR
AX		
SAG		

## Appendix B Anatomical landmark list

## 右第 3 腰椎横突起先端 [R\_L3\_TransverseP]

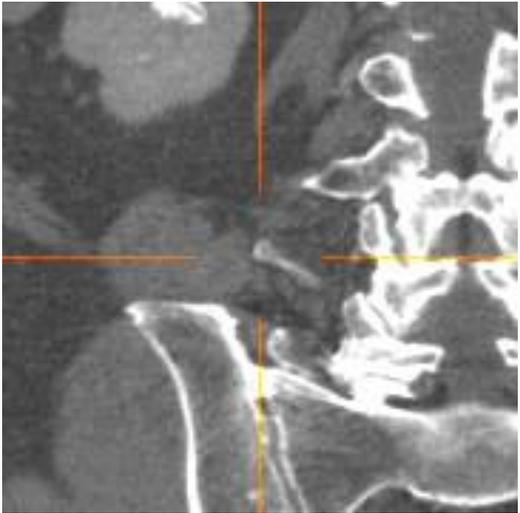
	AX	COR
AX		
SAG		

## 左第3腰椎横突起先端 [L\_L3\_TransverseP]

	AX	COR
AX		
SAG		

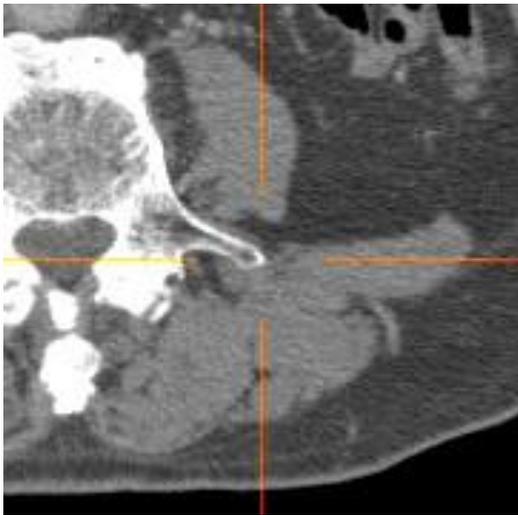
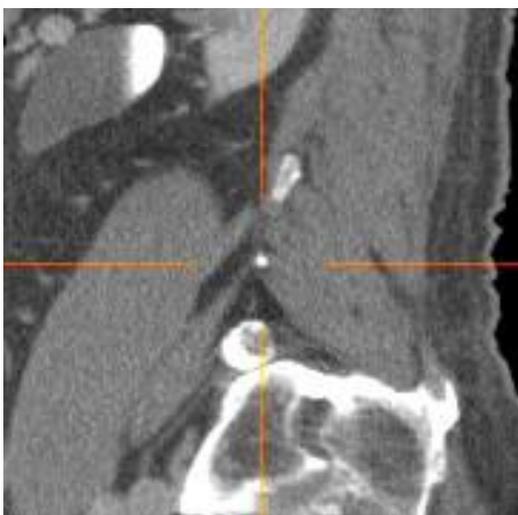
## Appendix B Anatomical landmark list

## 右第 4 腰椎横突起先端 [R\_L4\_TransverseP]

	AX	COR
AX		
SAG		

R\_L-3\_TransverseP

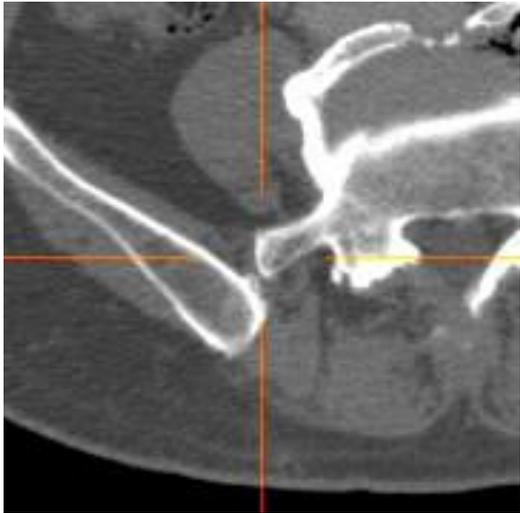
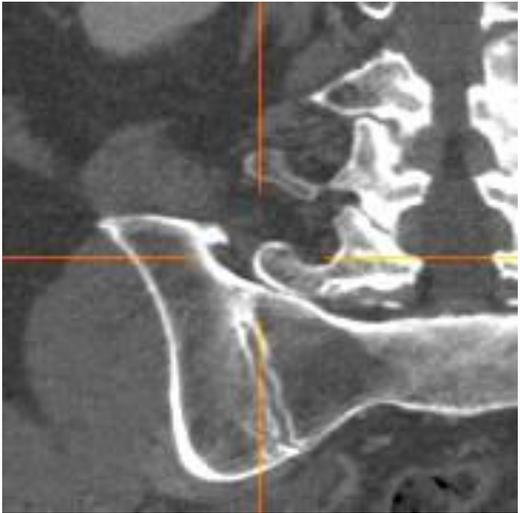
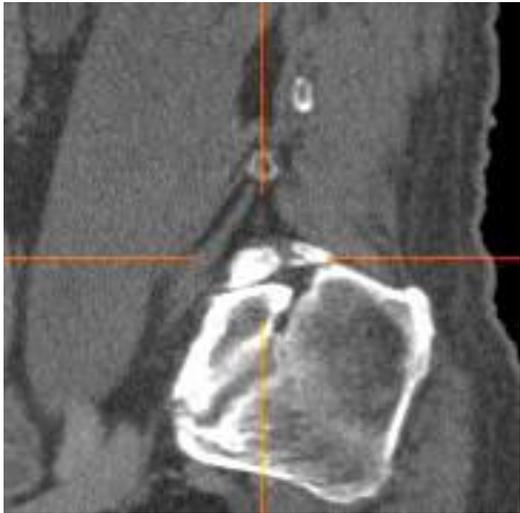
## 左第 4 腰椎横突起先端 [L\_L4\_TransverseP]

	AX	COR
AX		
SAG		

L\_L-3\_TransverseP

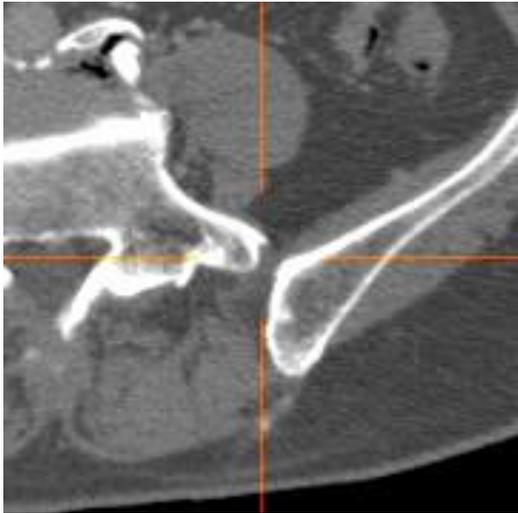
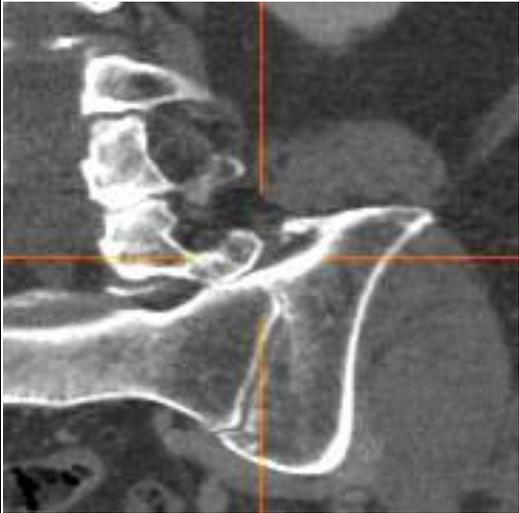
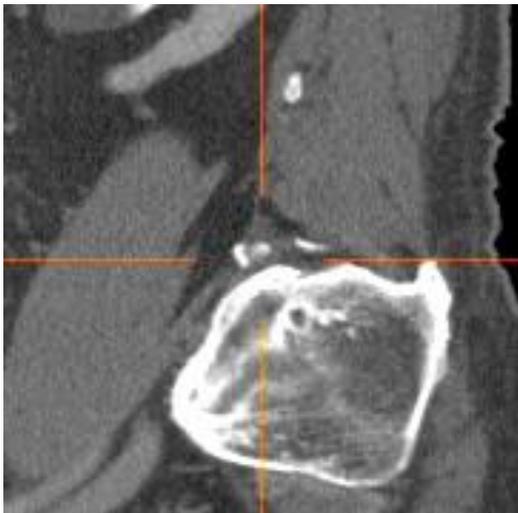
## Appendix B Anatomical landmark list

右第 5 腰椎横突起先端 [R\_L5\_TransverseP]

	AX	COR
AX		
SAG		

R\_L-2\_TransverseP

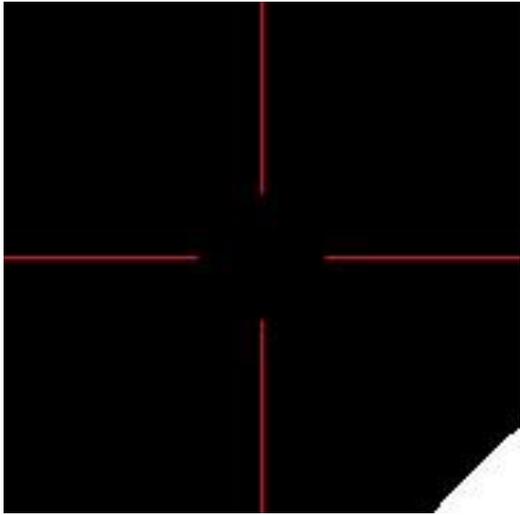
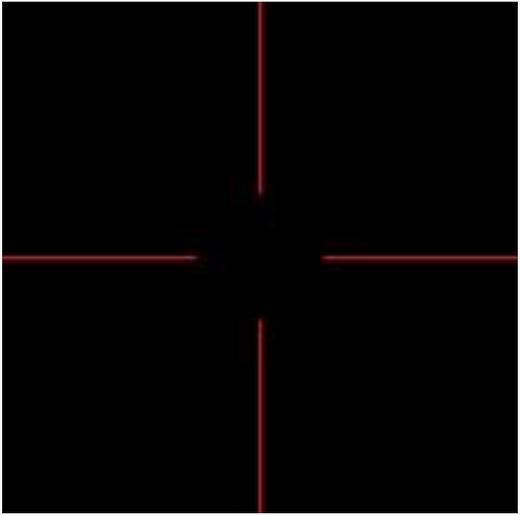
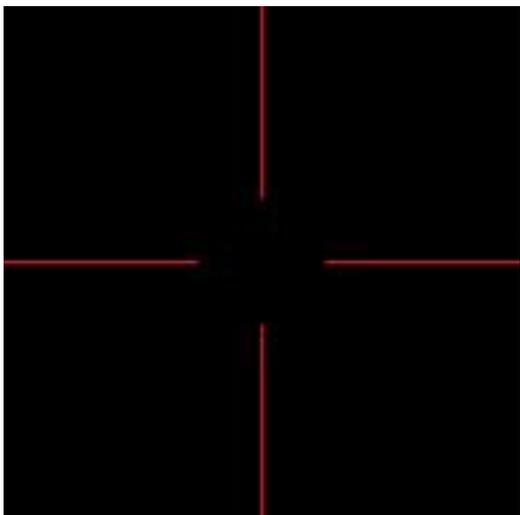
## 左第 5 腰椎横突起先端 [L\_L5\_TransverseP]

	AX	COR
AX		
SAG		

L\_L-2\_TransverseP

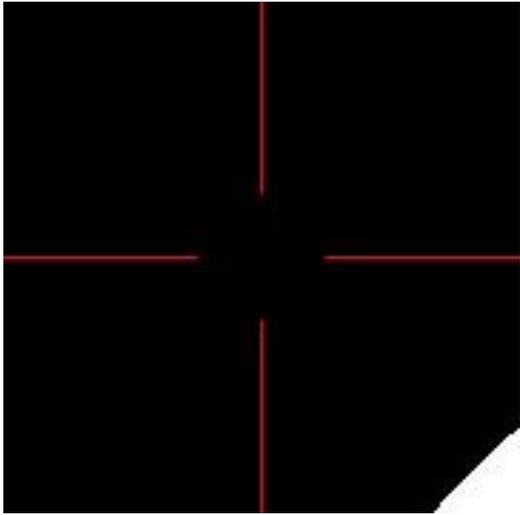
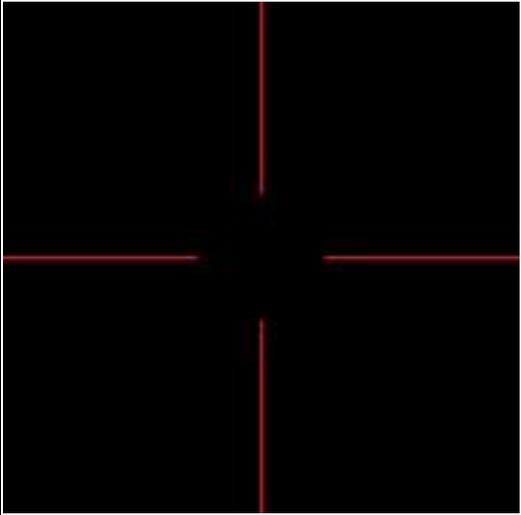
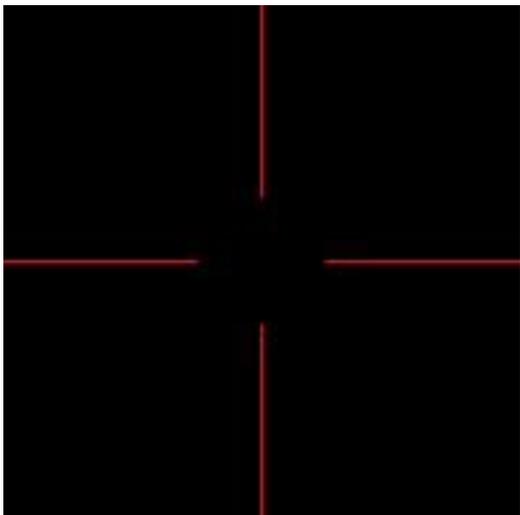
## Appendix B Anatomical landmark list

## 右第 6 腰椎横突起先端 [R\_L6\_TransverseP]

	AX	COR
AX		
SAG		

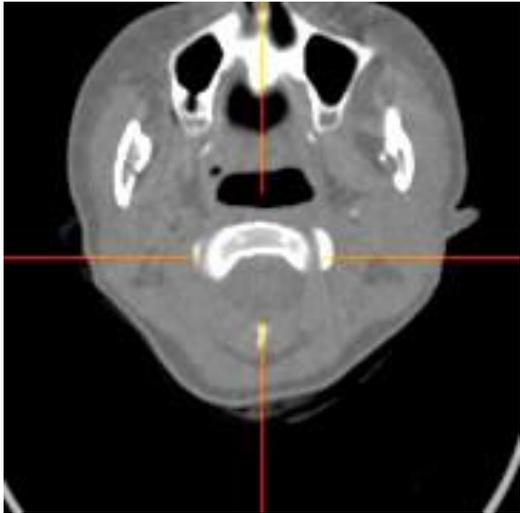
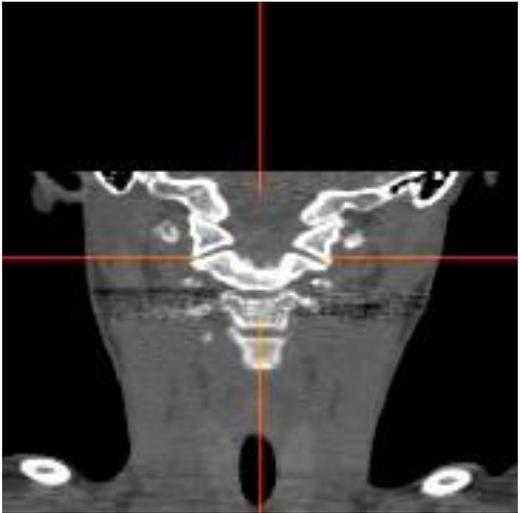
(破格時のみ)

## 左第 6 腰椎横突起先端 [L\_L6\_TransverseP]

	AX	COR
AX		
SAG		

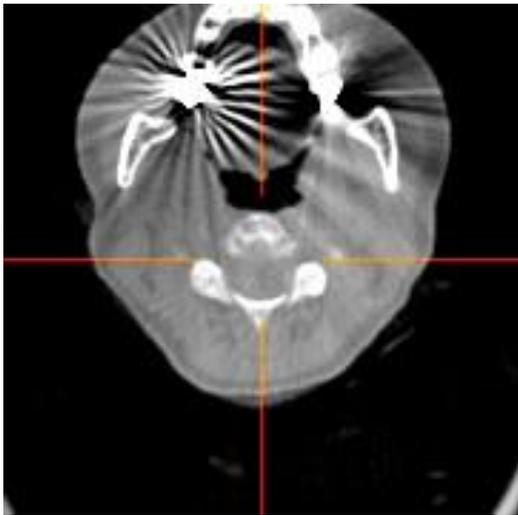
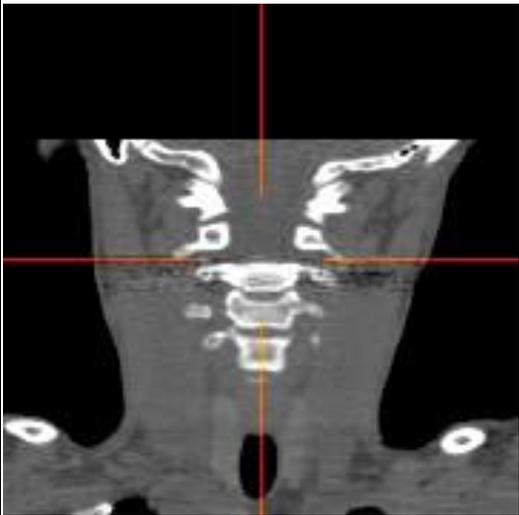
(破格時のみ)

## C1 下椎間後縁正中 [C1\_Intervert\_Post] (C1/2 intervertebral disk posterior)

	AX	COR
AX		
SAG		

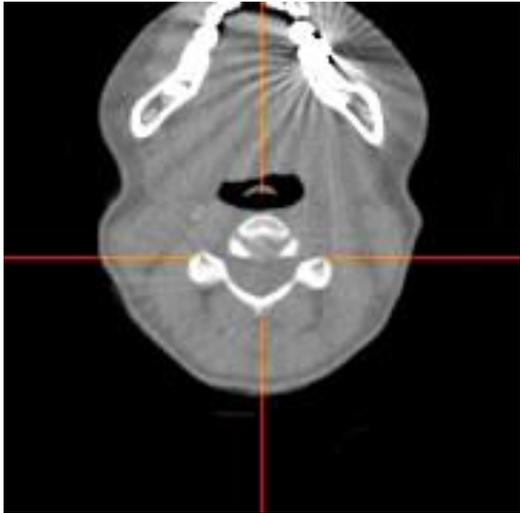
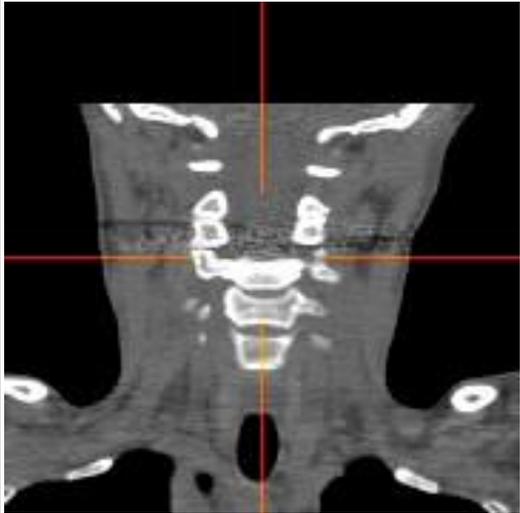
C1/2 椎間腔の後縁正中(脊柱管の前縁正中)を点として取る。骨棘は極力無視すること。

## C2 下椎間後緣正中 [C2\_Intervert\_Post]

	AX	COR
AX		
SAG		

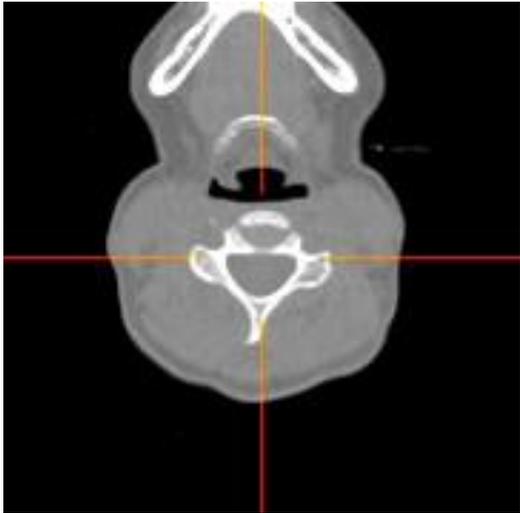
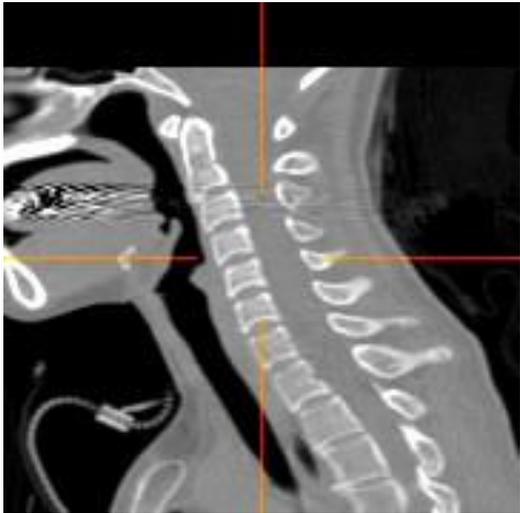
C2/3 椎間腔後緣正中。

## C3 下椎間後縁正中 [C3\_Intervert\_Post]

	AX	COR
AX		
SAG		

C3/4 椎間腔後縁正中。

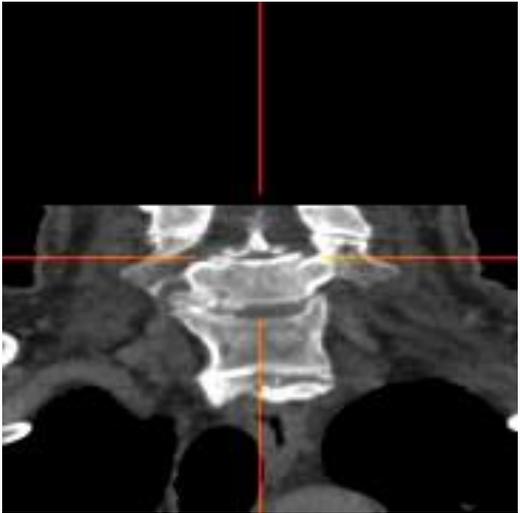
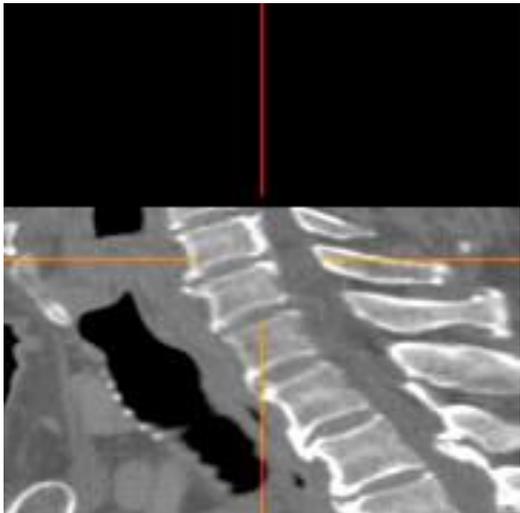
## C4 下椎間後緣正中 [C4\_Intervert\_Post]

	AX	COR
AX		
SAG		

C4/5 椎間腔後緣正中。

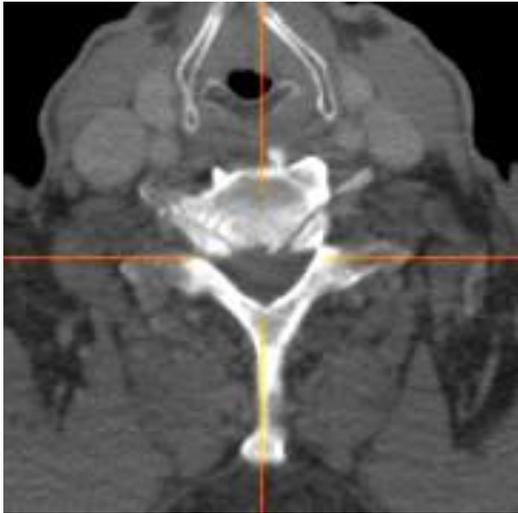
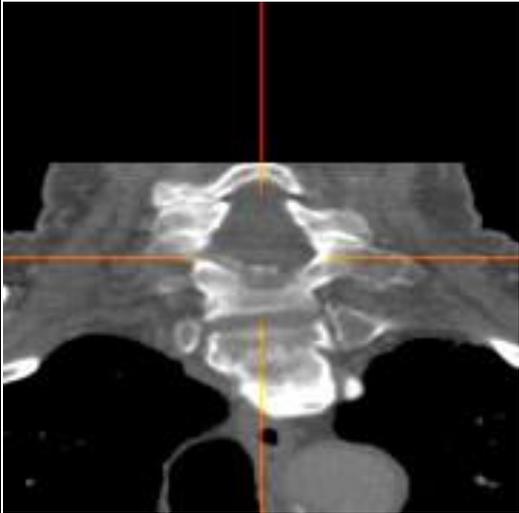
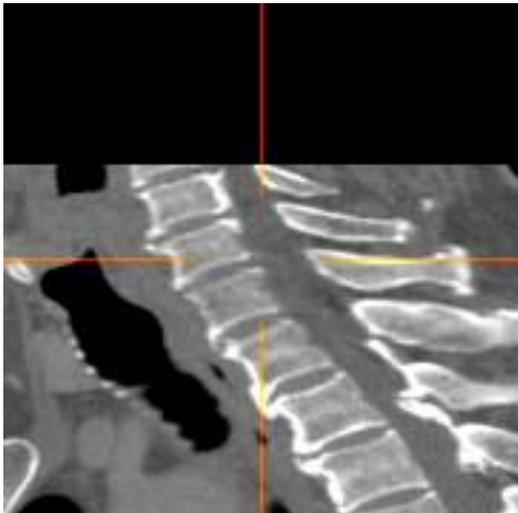
## Appendix B Anatomical landmark list

## C5 下椎間後縁正中 [C5\_Intervert\_Post]

	AX	COR
AX		
SAG		

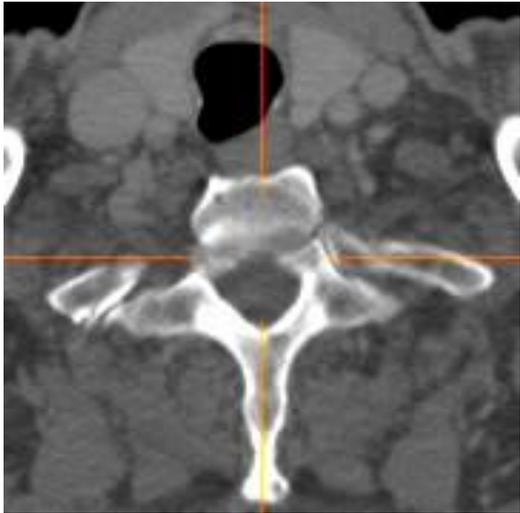
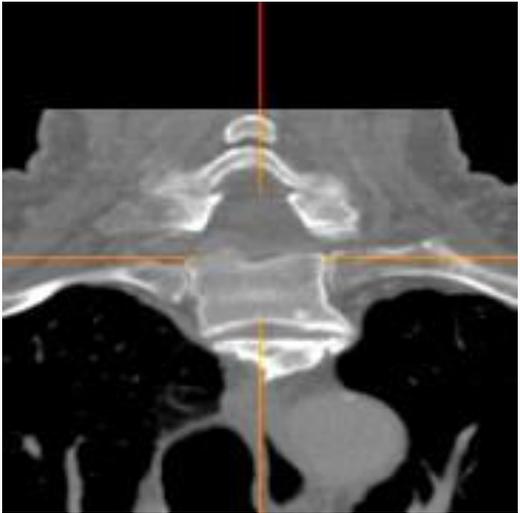
C5/6 椎間腔後縁正中。

## C6 下椎間後縁正中 [C6\_Intervert\_Post]

	AX	COR
AX		
SAG		

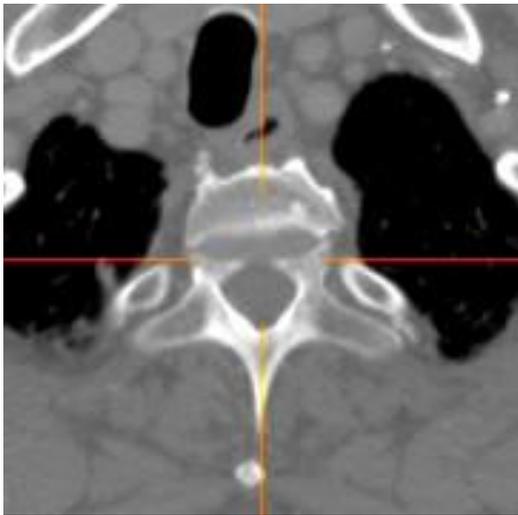
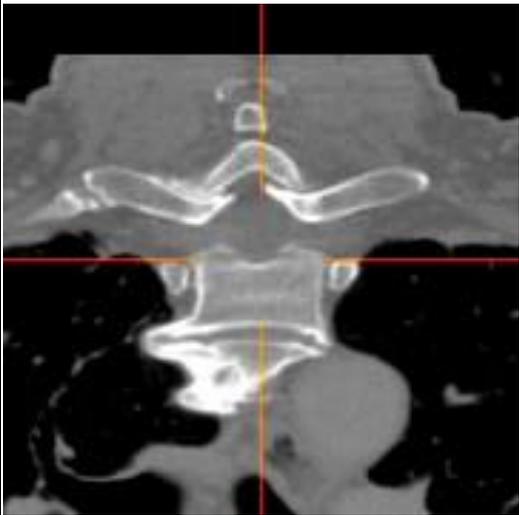
C6/7 椎間腔後縁正中。

## C7 下椎間後縁正中 [C7\_Intervert\_Post]

	AX	COR
AX		
SAG		

C7/Th1 椎間腔後縁正中。

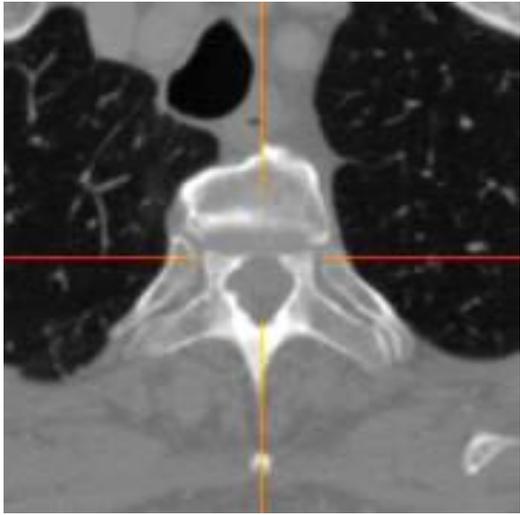
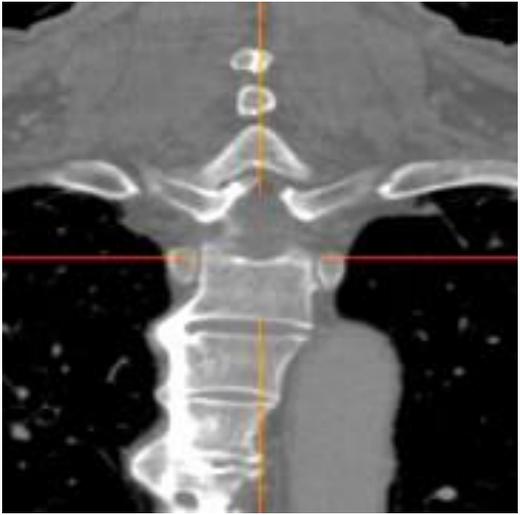
## Th1 下椎間後縁正中 [Th1\_Intervert\_Post]

	AX	COR
AX		
SAG		

Th1/2 椎間腔後縁正中。

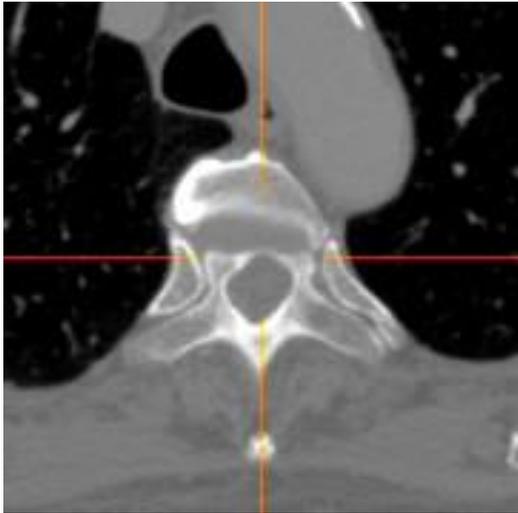
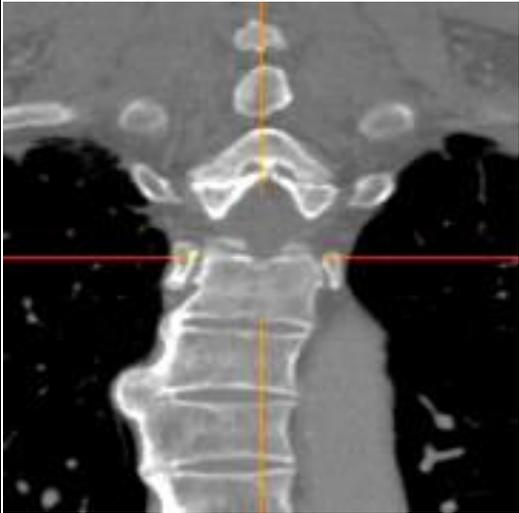
## Appendix B Anatomical landmark list

## Th2 下椎間後縁正中 [Th2\_Intervert\_Post]

	AX	COR
AX		
SAG		

Th2/3 椎間腔後縁正中。

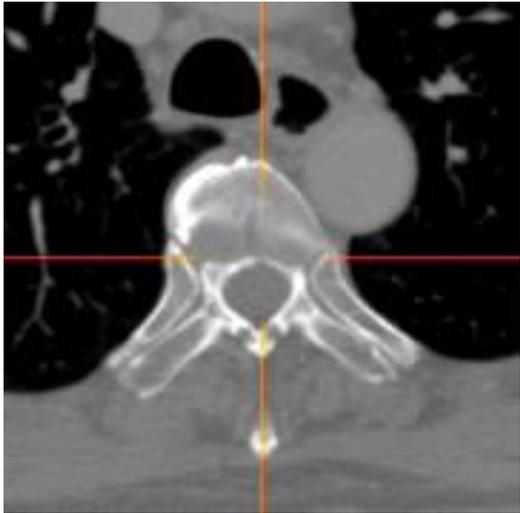
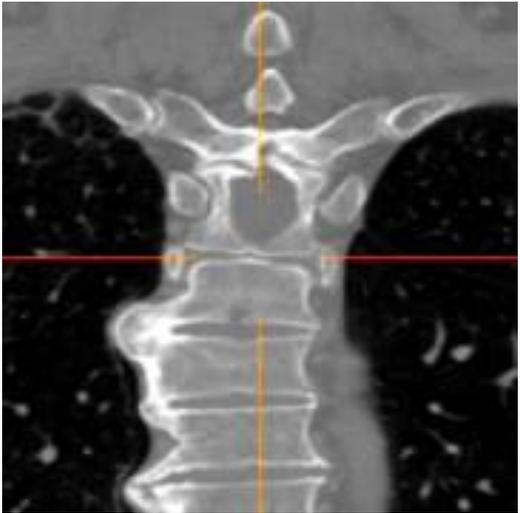
## Th3 下椎間後緣正中 [Th3\_Intervert\_Post]

	AX	COR
AX		
SAG		

Th3/4 椎間腔後緣正中。

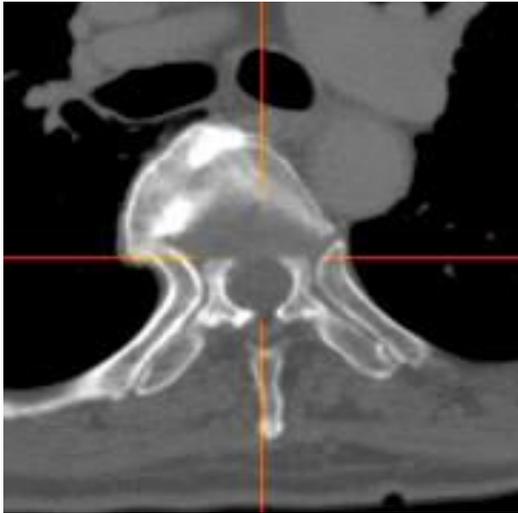
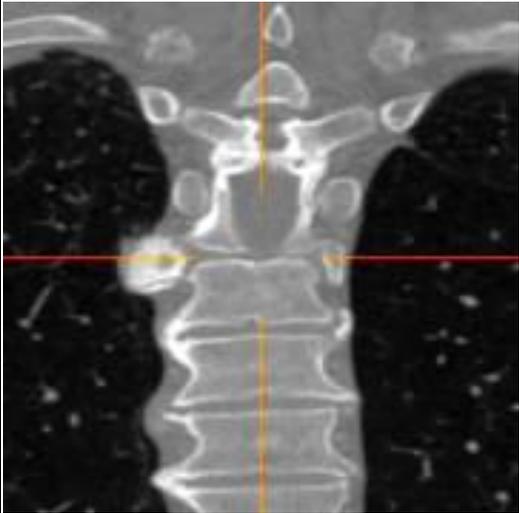
## Appendix B Anatomical landmark list

## Th4 下椎間後縁正中 [Th4\_Intervert\_Post]

	AX	COR
AX		
SAG		

Th4/5 椎間腔後縁正中。

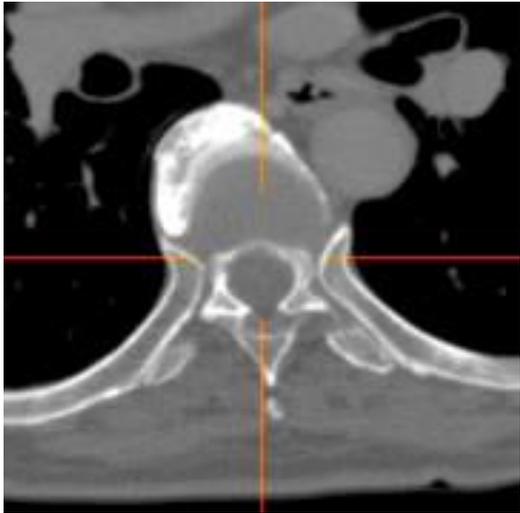
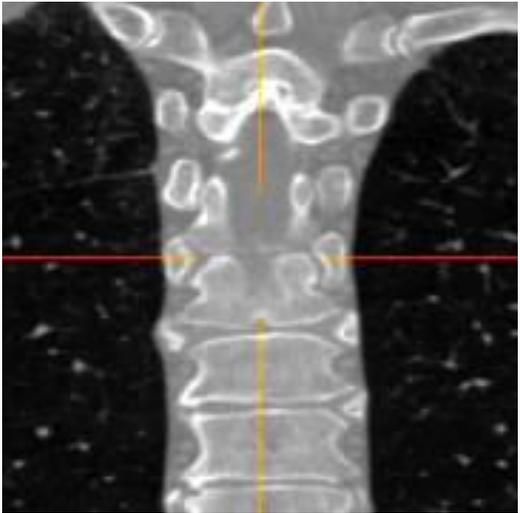
## Th5 下椎間後緣正中 [Th5\_Intervert\_Post]

	AX	COR
AX		
SAG		

Th5/6 椎間腔後緣正中。

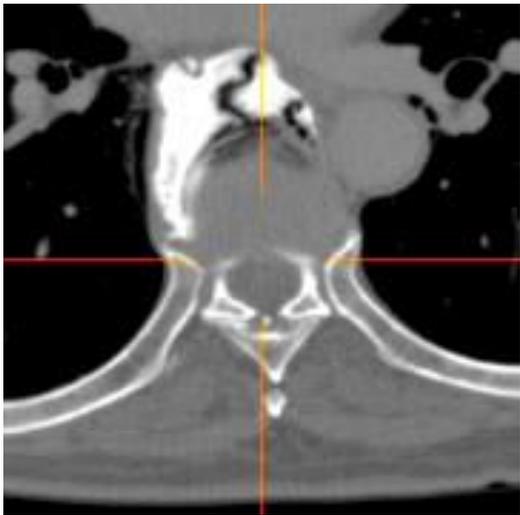
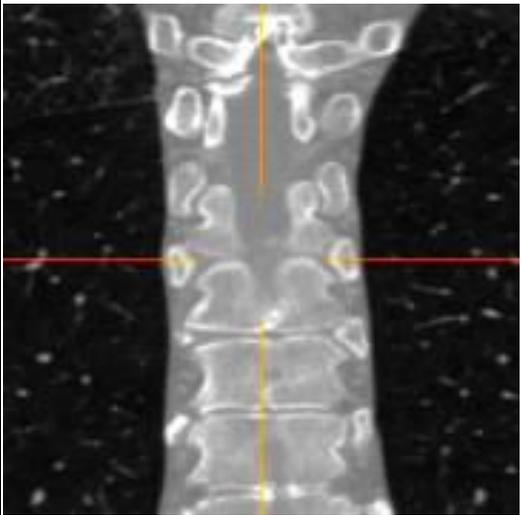
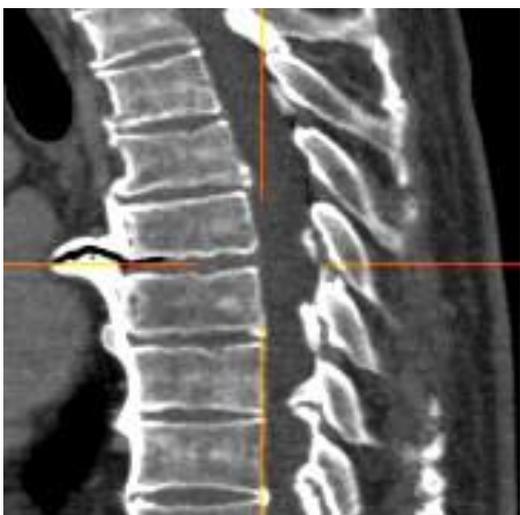
## Appendix B Anatomical landmark list

## Th6 下椎間後縁正中 [Th6\_Intervert\_Post]

	AX	COR
AX		
SAG		

Th6/7 椎間腔後縁正中。

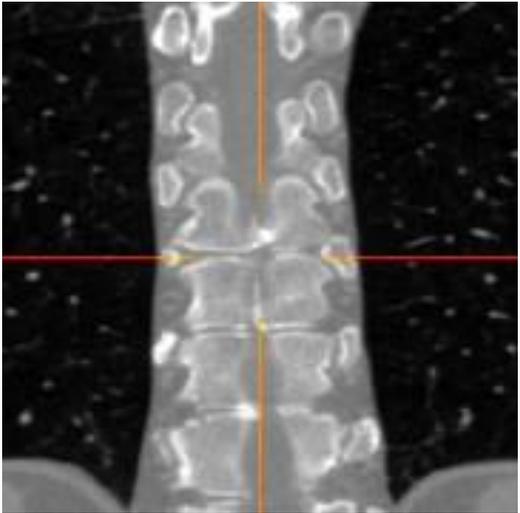
## Th7 下椎間後縁正中 [Th7\_Intervert\_Post]

	AX	COR
AX		
SAG		

Th7/8 椎間腔後縁正中。

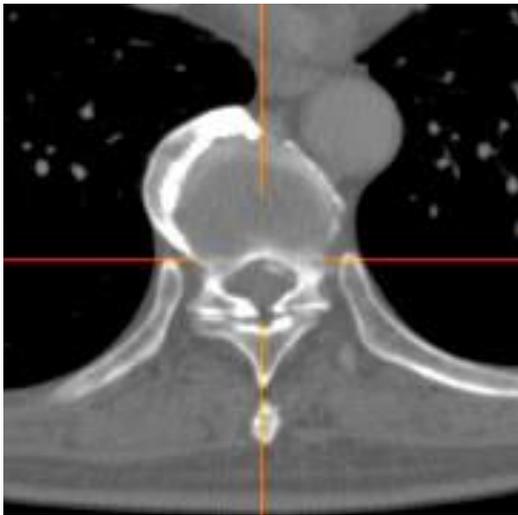
## Appendix B Anatomical landmark list

## Th8 下椎間後縁正中 [Th8\_Intervert\_Post]

	AX	COR
AX		
SAG		

Th8/9 椎間腔後縁正中。

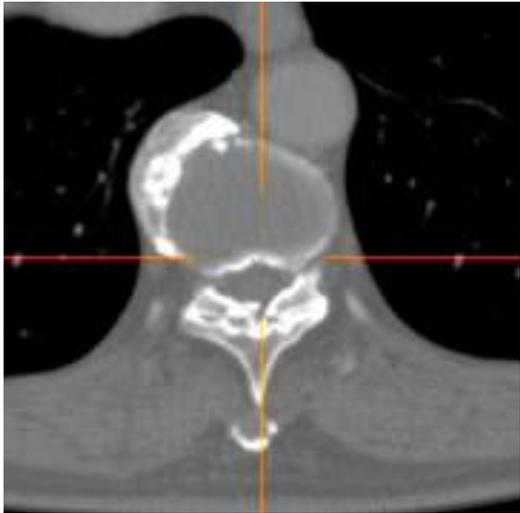
## Th9 下椎間後緣正中 [Th9\_Intervert\_Post]

	AX	COR
AX		
SAG		

Th9/10 椎間腔後緣正中。

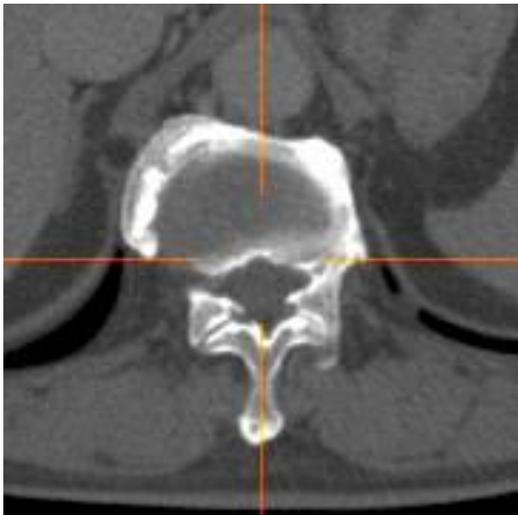
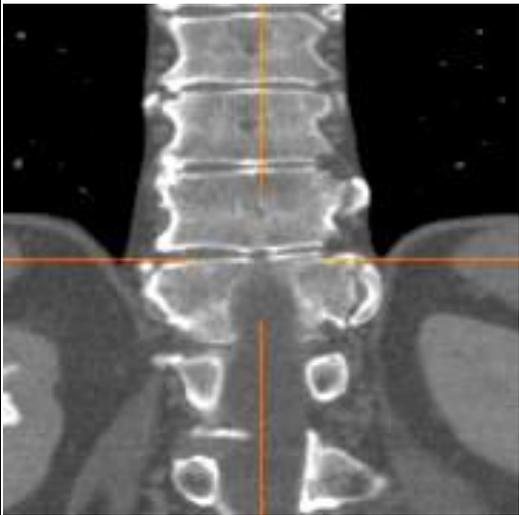
## Appendix B Anatomical landmark list

## Th10 下椎間後縁正中 [Th10\_Intervert\_Post]

	AX	COR
AX		
SAG		

Th10/11 椎間腔後縁正中。

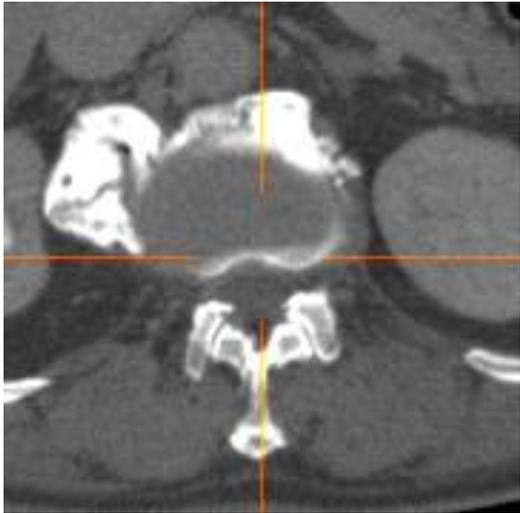
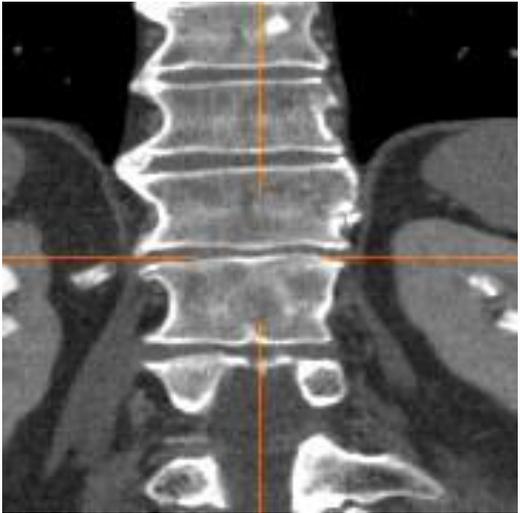
## Th11 下椎間後縁正中 [Th11\_Intervert\_Post]

	AX	COR
AX		
SAG		

Th11/12 椎間腔後縁正中。

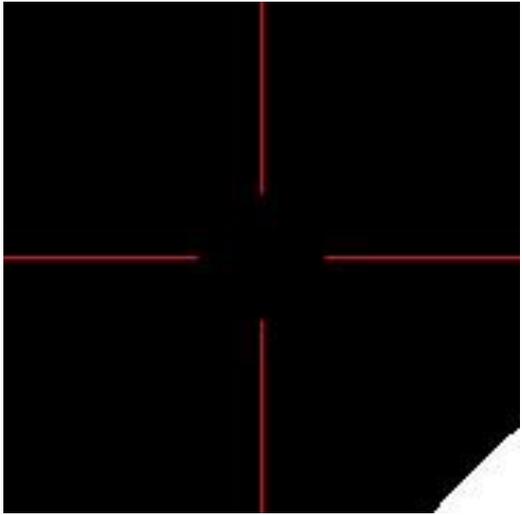
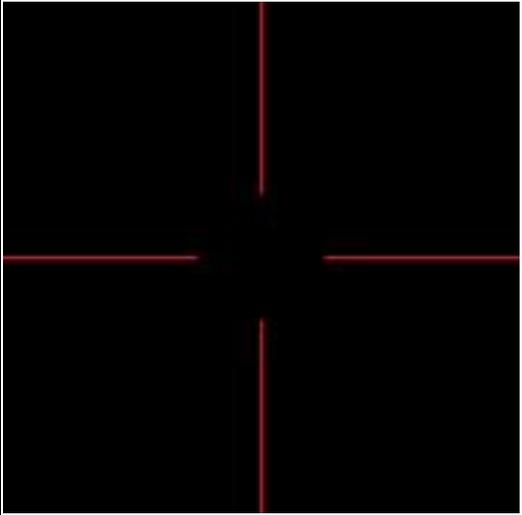
## Appendix B Anatomical landmark list

## Th12 下椎間後縁正中 [Th12\_Intervert\_Post]

	AX	COR
AX		
SAG		

Th12/L1、ただし Th13 が存在するときは Th12/13

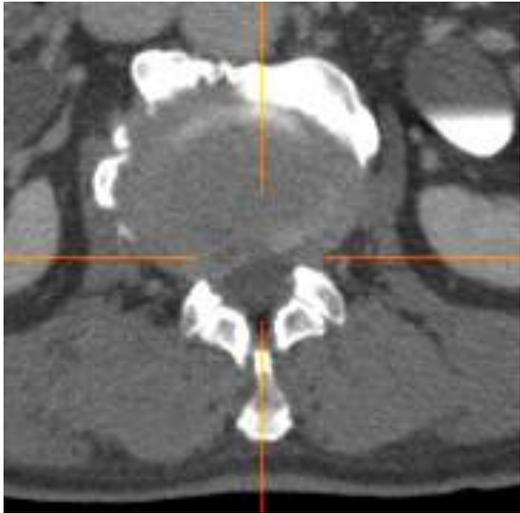
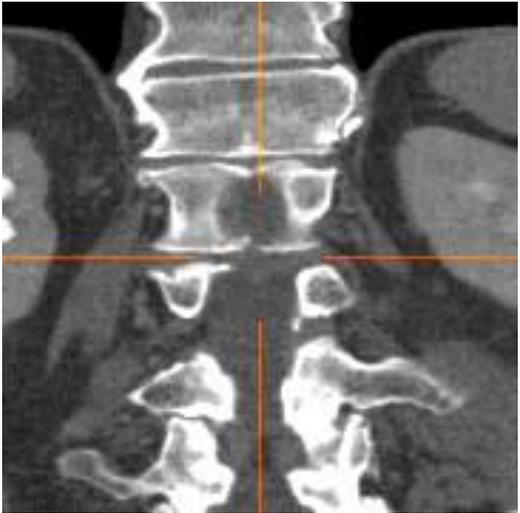
## Th13 下椎間後縁正中 [Th13\_Intervertebr\_Post]

	AX	COR
AX		
SAG		

(破格時のみ)Th13 が存在するときのみ、Th13/L1 をとる。

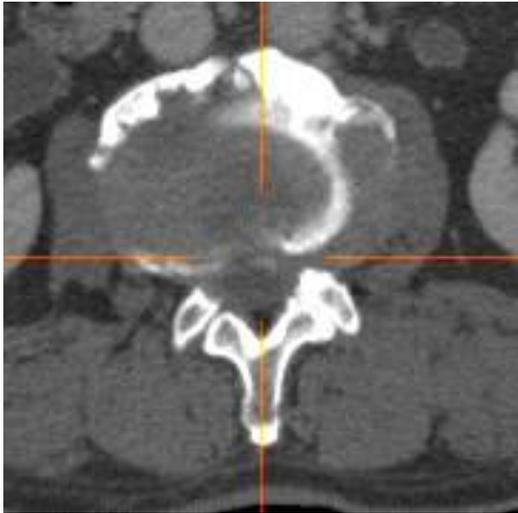
## Appendix B Anatomical landmark list

## L1 下椎間後縁正中 [L1\_Intervert\_Post]

	AX	COR
AX		
SAG		

Th12(もしくは 13)/L1 椎間腔後縁正中。

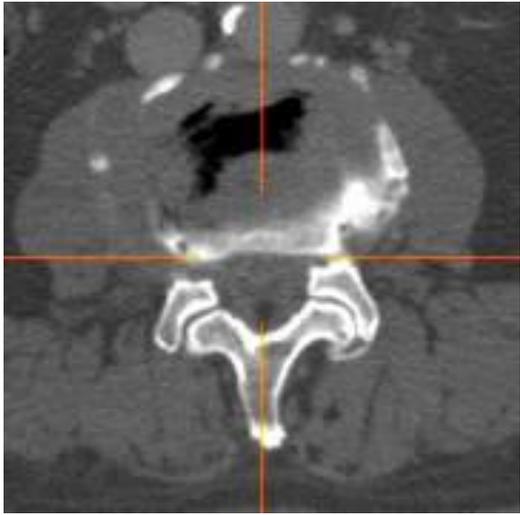
## L2 下椎間後緣正中 [L2\_Intervert\_Post]

	AX	COR
AX		
SAG		

L2/3 椎間腔後緣正中。

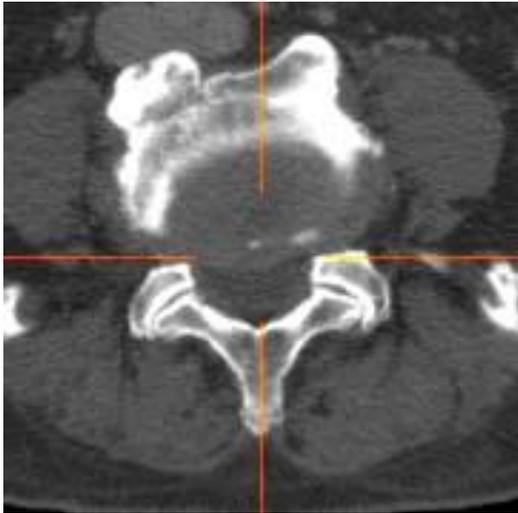
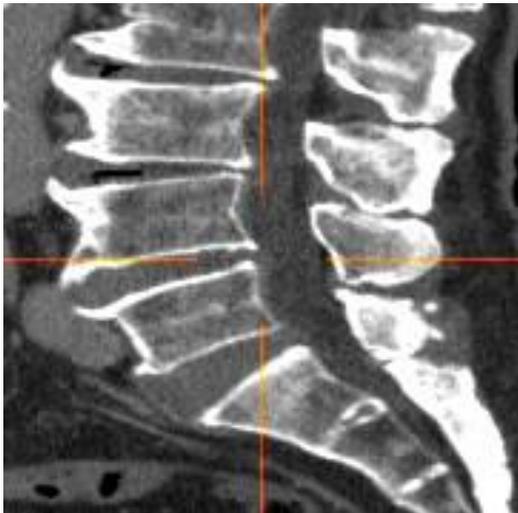
## Appendix B Anatomical landmark list

## L3 下椎間後縁正中 [L3\_Intervert\_Post]

	AX	COR
AX		
SAG		

L3/4 椎間腔後縁正中。

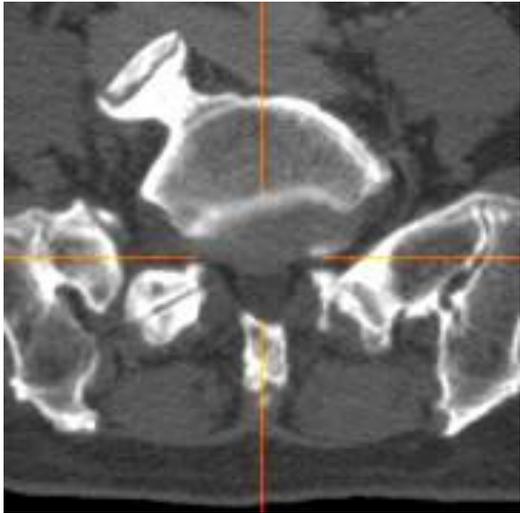
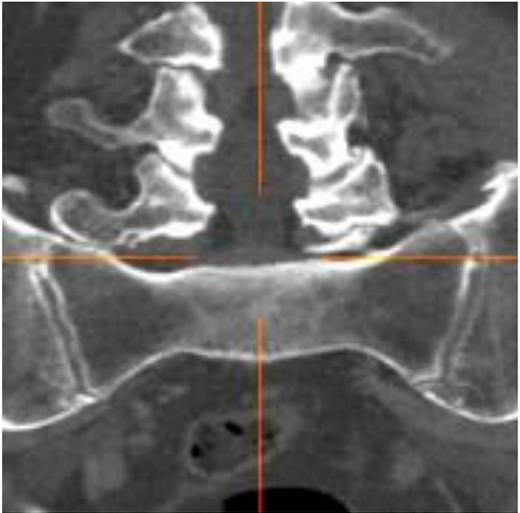
## L4 下椎間後緣正中 [L4\_Intervert\_Post]

	AX	COR
AX		
SAG		

L4/5 椎間腔後緣正中。

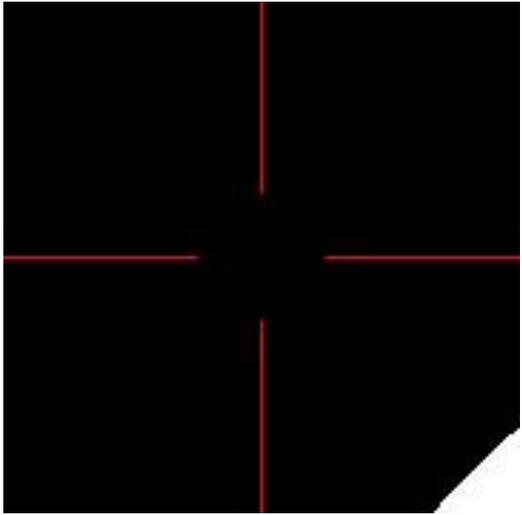
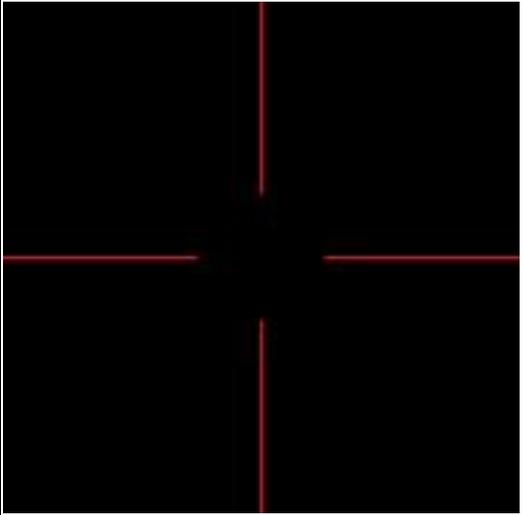
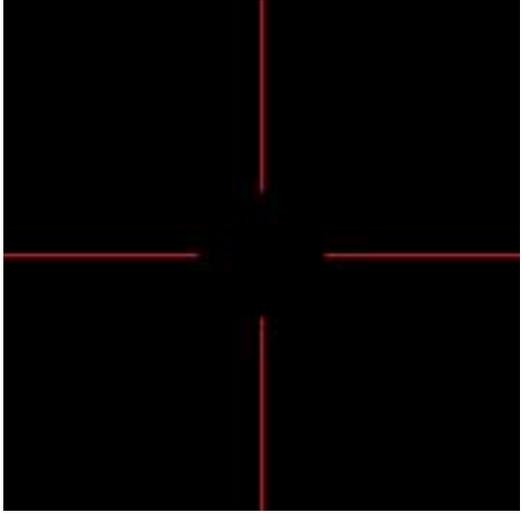
## Appendix B Anatomical landmark list

## L5 下椎間後縁正中 [L5\_Intervert\_Post]

	AX	COR
AX		
SAG		

L5/S1 椎間腔後縁正中。L6 があるときは L5/L6。

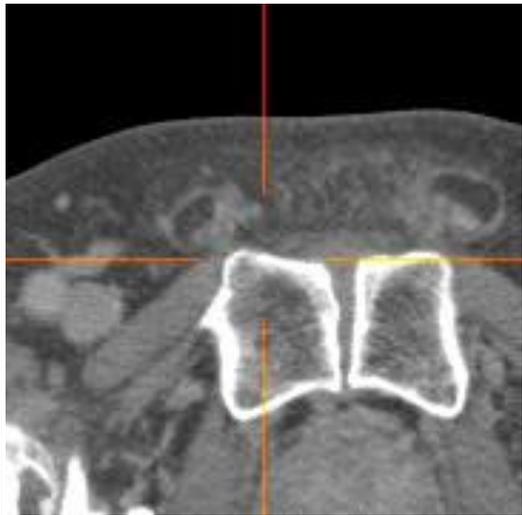
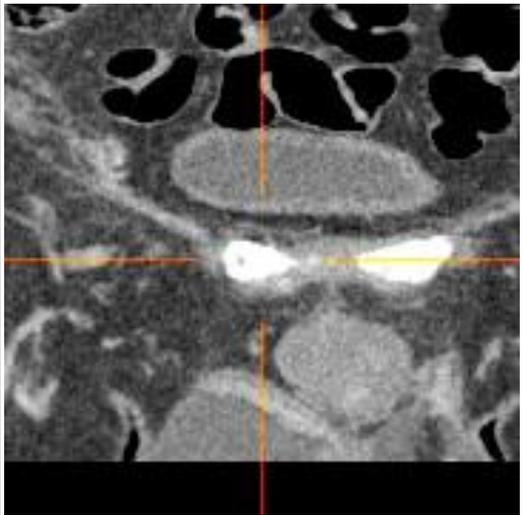
## L6 下椎間後縁正中 [L6\_Intervert\_Post]

	AX	COR
AX		
SAG		

(破格時のみ)L6/S1 椎間腔後縁正中。

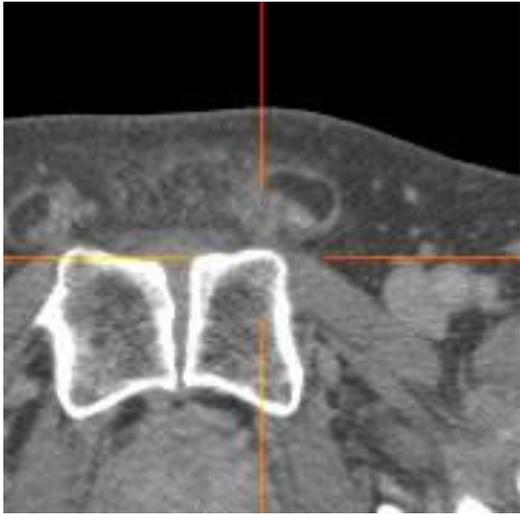
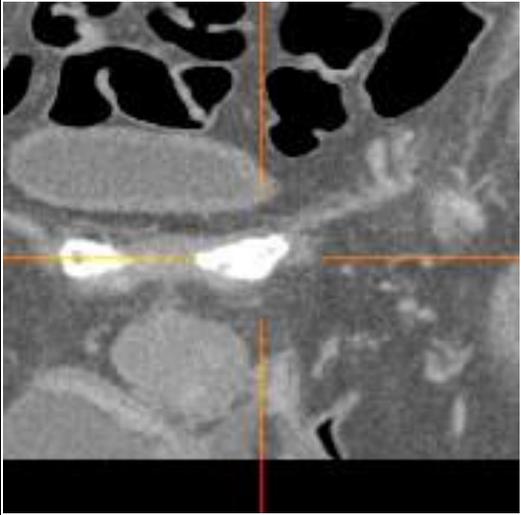
## Appendix B Anatomical landmark list

## 右恥骨結節 [R\_Pubic\_Tubercle]

	AX	COR
AX		
SAG		

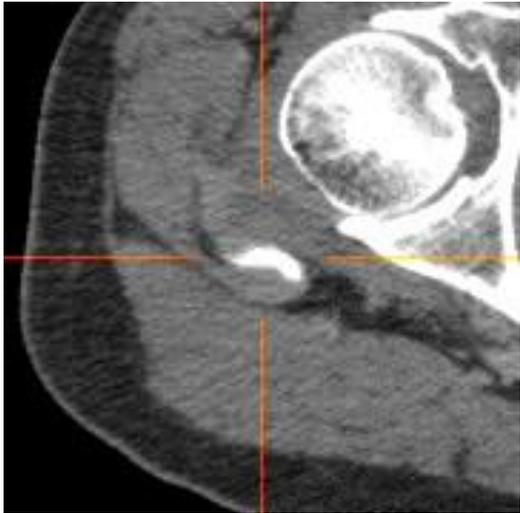
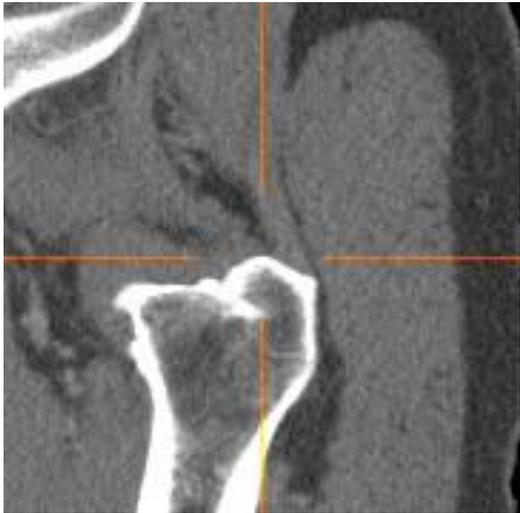
恥骨結合の右、恥骨がやや前方に隆起したところ。ただし対応する構造がないこともあり、そのときはだいたい恥骨結合から1cmほど外側のところの骨皮質をとる。

## 左恥骨結節 [L\_Pubic\_Tubercle]

	AX	COR
AX		
SAG		

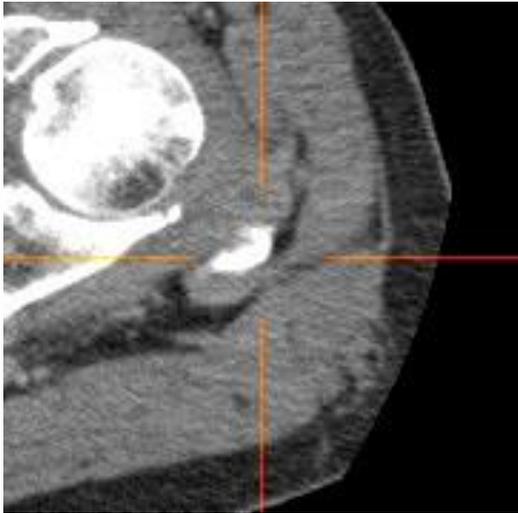
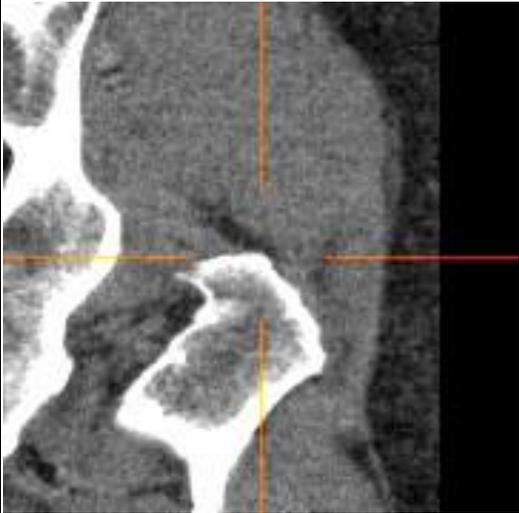
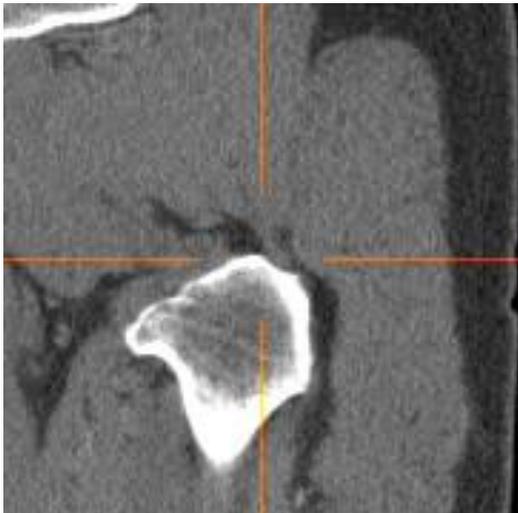
## Appendix B Anatomical landmark list

## 右大転子 [R\_Greater\_Trochanter]

	AX	COR
AX		
SAG		

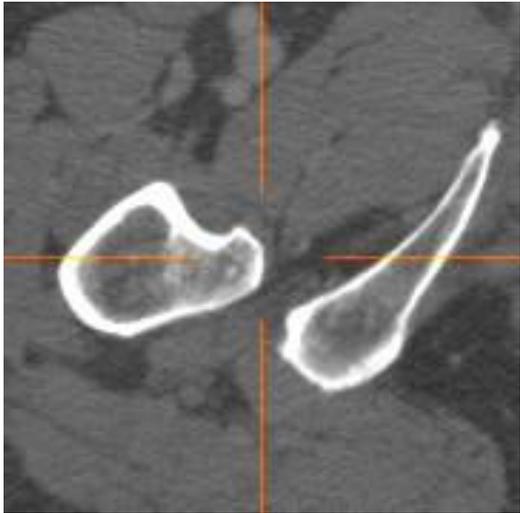
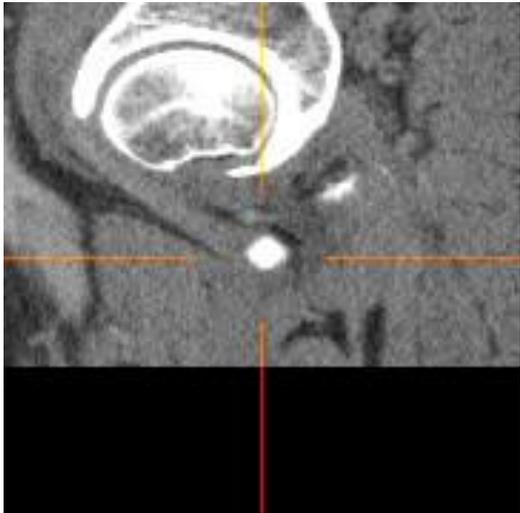
右大転子の頭側端を原則としてはとる。頭側端のなかでどこをとるか迷ったら、突起として先端に近いほうをとる。

## 左大転子 [L\_Greater\_Trochanter]

	AX	COR
AX		
SAG		

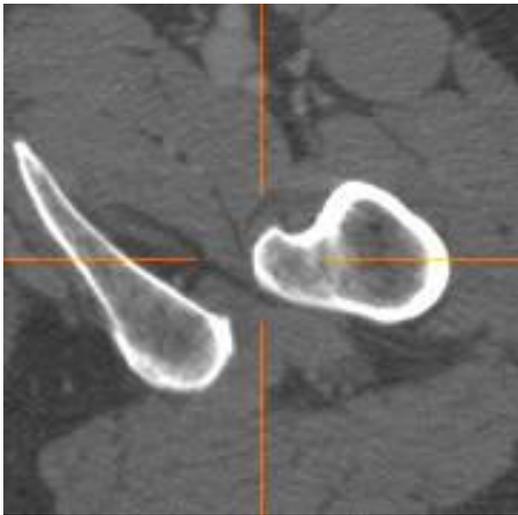
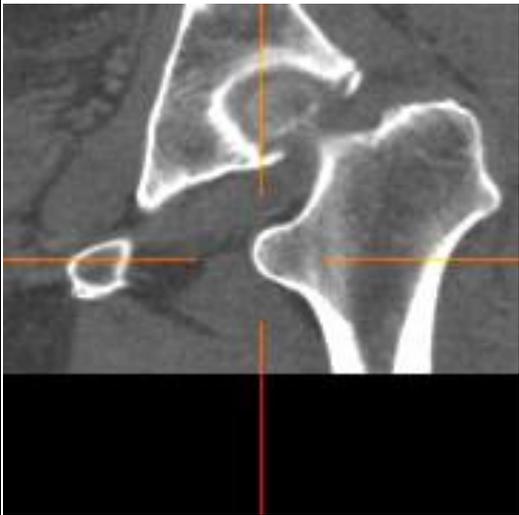
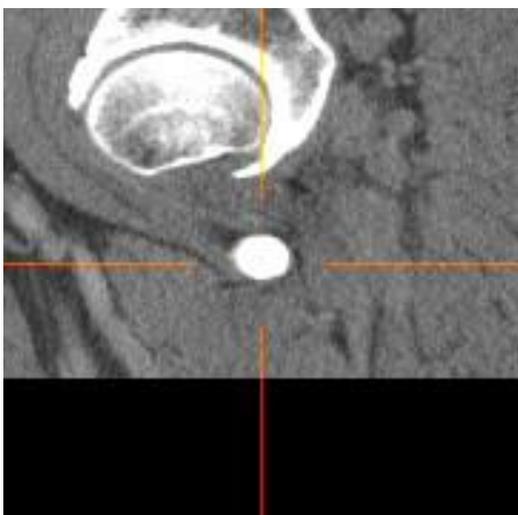
## Appendix B Anatomical landmark list

## 右小転子 [R\_Lesser\_Trochanter]

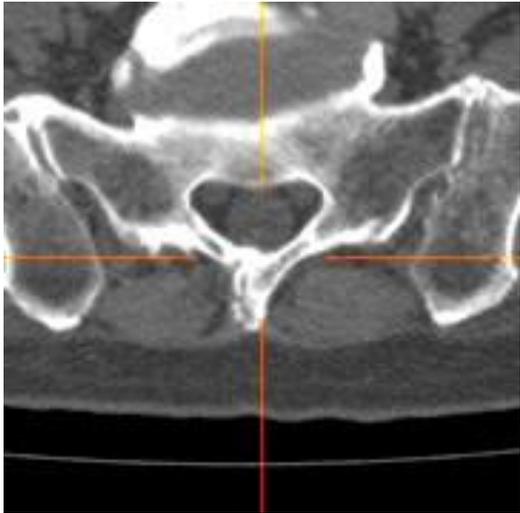
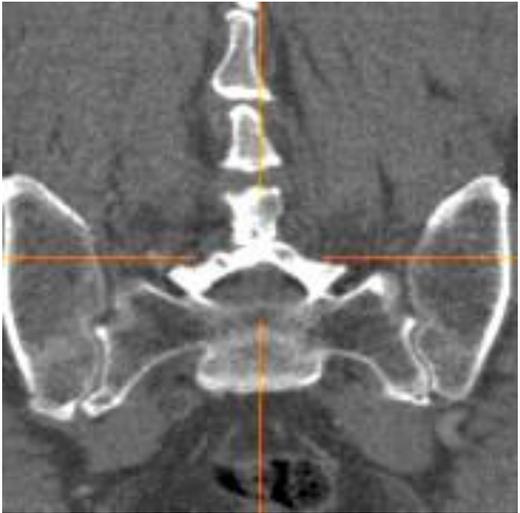
	AX	COR
AX		
SAG		

小転子は小さいので、先端中心の骨皮質を選ぶ。

## 左小転子 [L\_Lesser\_Trochanter]

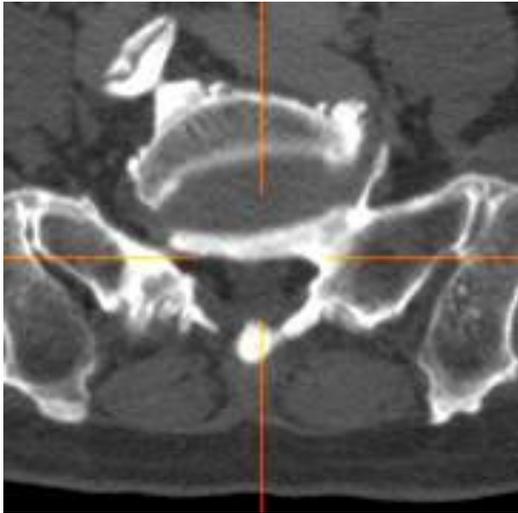
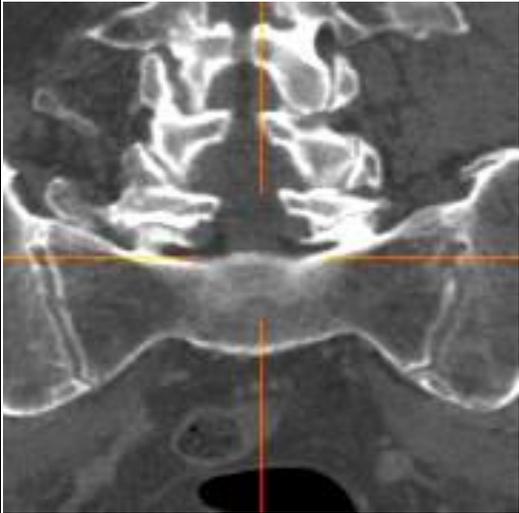
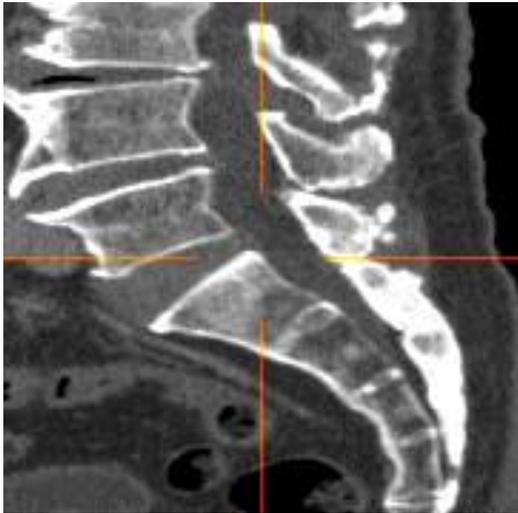
	AX	COR
AX		
SAG		

## 仙骨管後上縁 [SacralCanal\_PostSupTip]

	AX	COR
AX		
SAG		

仙骨管(仙骨に囲まれた、脊柱管と連続する腔)の後壁の上前縁。必ず仙骨正中でとること。左右に分裂している(癒合していない)ことも多いが、そのときは分裂していないレベルまで尾側に下がったところを取る。

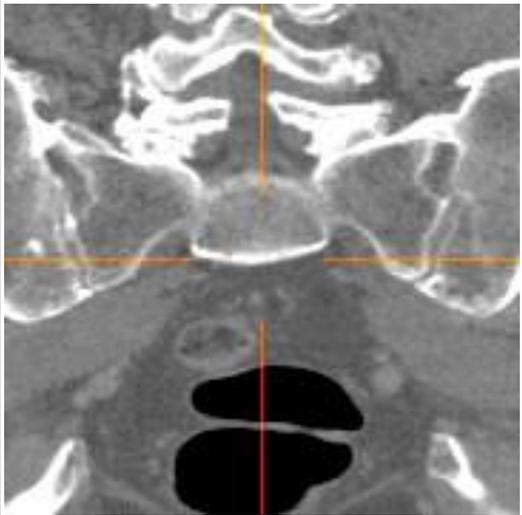
## 仙骨体後上縁 [SacralBody\_PostSupTip]

	AX	COR
AX		
SAG		

仙骨管の前壁上縁、仙骨体の後上縁。L5/S 椎間板後縁正中とほぼ同一の点だが、こちらは骨皮質上をとる。

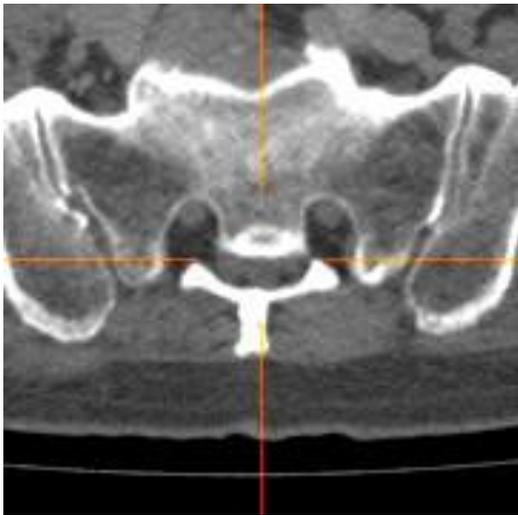
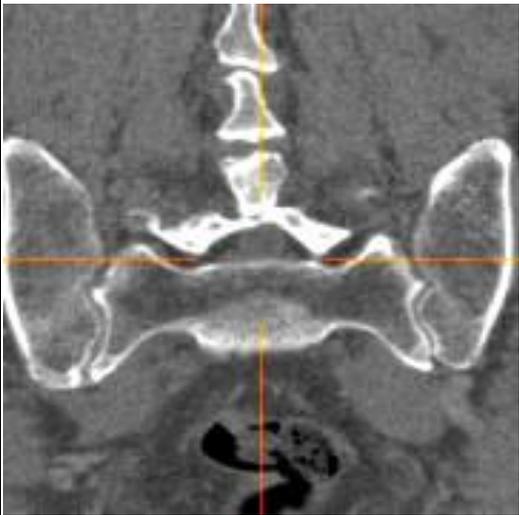
## Appendix B Anatomical landmark list

## 第 1-2 仙椎椎間板正中前縁 [Sacrum Intervert\_Ant]

	AX	COR
AX		
SAG		

S1/2 の椎間板(もしくは椎間板様の構造)の正中前縁。骨癒合していることが多く、そのときは骨皮質上をとる。さもなくば椎間板前端となる。

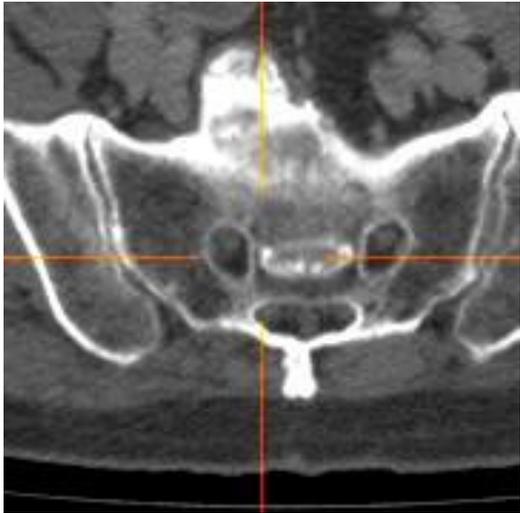
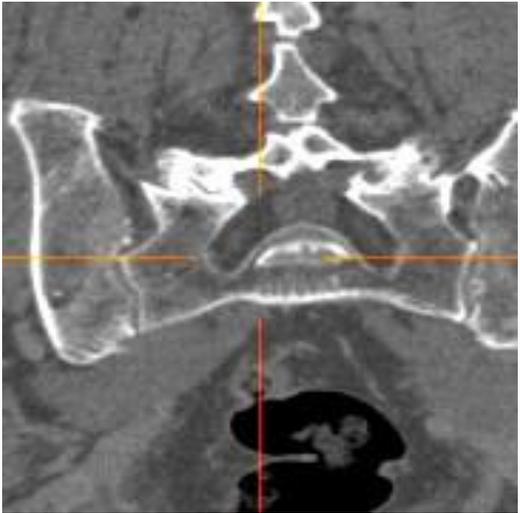
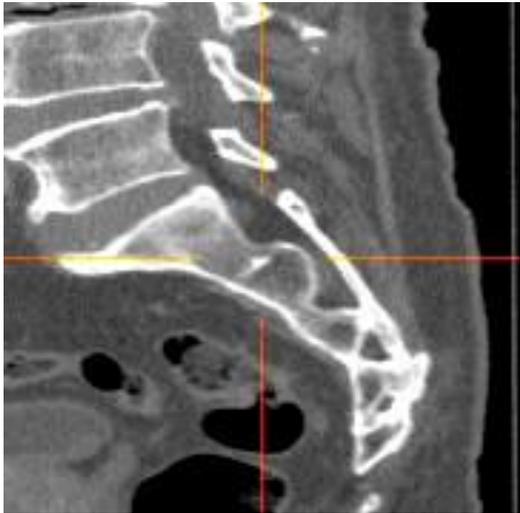
## 第 1-2 仙椎椎間板正中後縁 [Sacrum\_Intervert\_Post]

	AX	COR
AX		
SAG		

同じく S1/2 椎間板(様構造)の正中後縁で、仙骨管前縁。

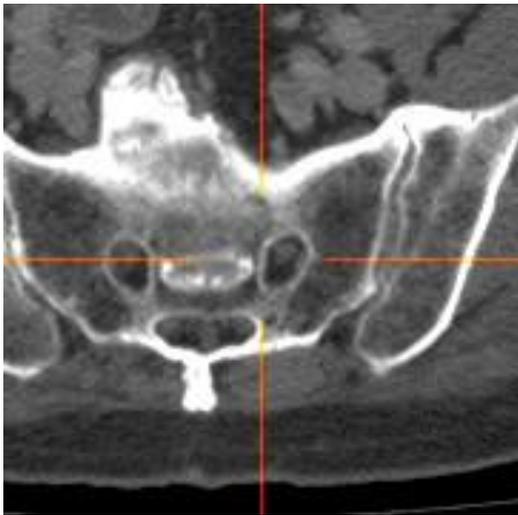
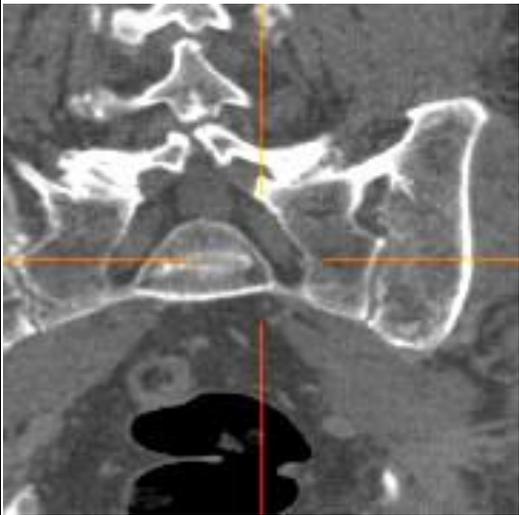
## Appendix B Anatomical landmark list

## 右第 1-2 仙椎椎間板外側縁 [R\_Sacrum\_Intervert]

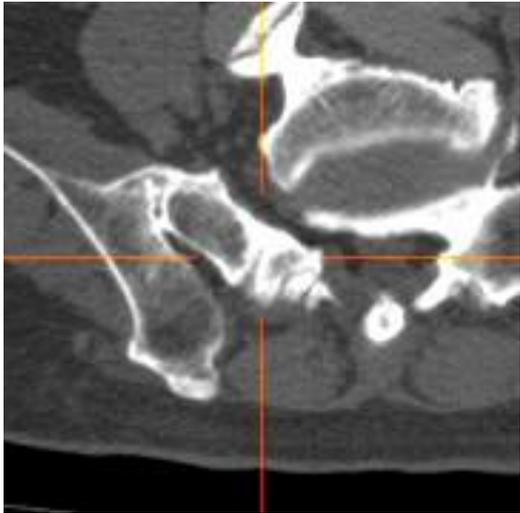
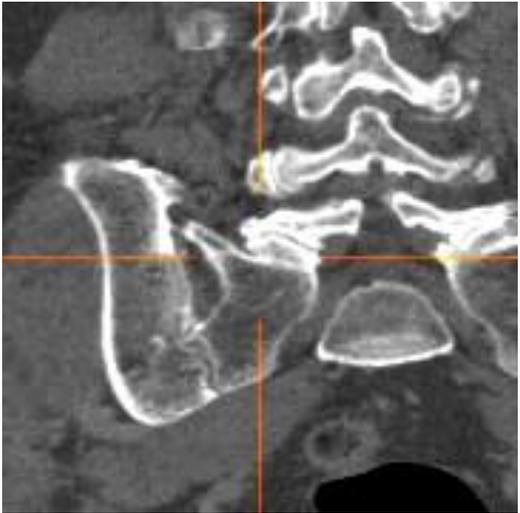
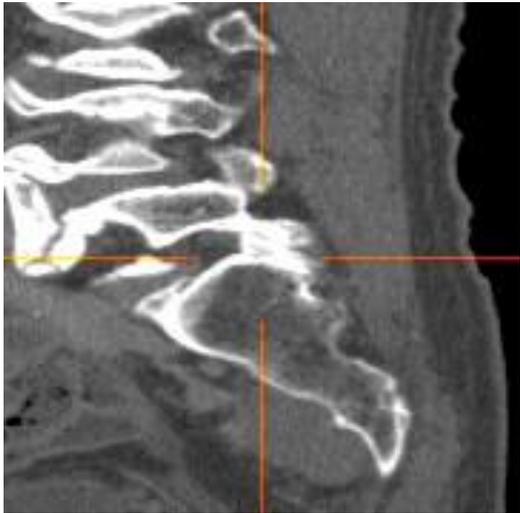
	AX	COR
AX		
SAG		

S1/2 の椎間板(様構造)の右縁。S1 神経根の通る神経孔が椎間板を左右に挟んでメガネ状に並んでいる水平断スライスを選び、その右神経孔と S1/2 椎間板との接点をとること。

## 左第 1-2 仙椎椎間板外側緣 [L\_Sacrum\_Intervert]

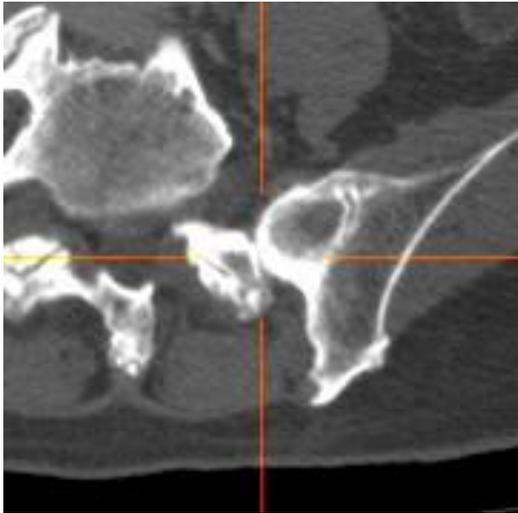
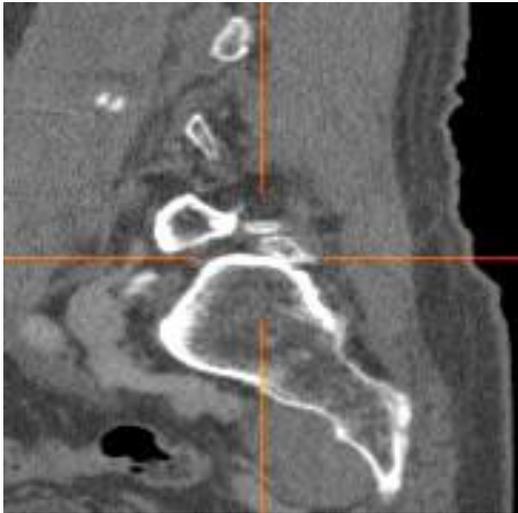
	AX	COR
AX		
SAG		

## 右 SuperiorSacralNotch [R\_Sup\_Sacral\_Notch]

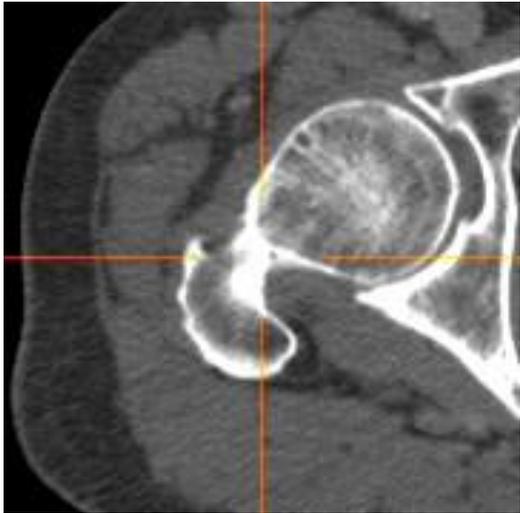
	AX	COR
AX		
SAG		

仙骨の仙骨体(椎体にあたる部分)と右の仙腸関節部の間で、冠状断でみて仙骨が尾側にくぼんでいるところの、その鞍点をとる。

## 左 SuperiorSacralNotch [L\_Sup\_Sacral\_Notch]

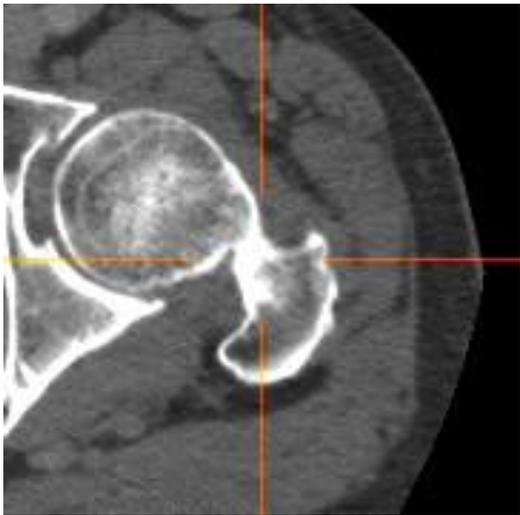
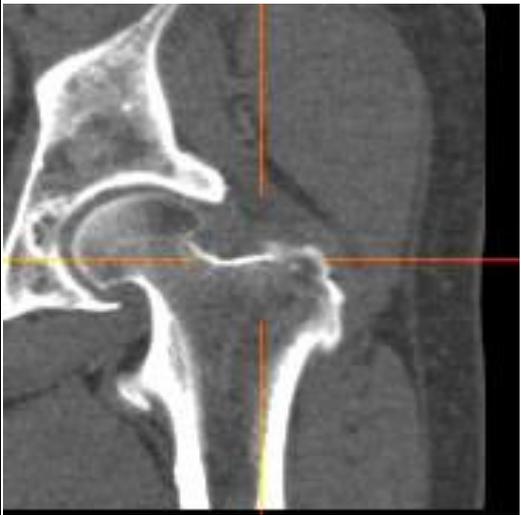
	AX	COR
AX		
SAG		

## 右転子窩 [R\_TrochantericFossa]

	AX	COR
AX		
SAG		

右大転子と小転子の間の峠となる鞍点。

左转子窝 [L\_TrochantericFossa]

	AX	COR
AX		
SAG	