# Topology Characterises Tumour Vasculature

**Helen M. Byrne**, **Heather A. Harrington**, **Ruth Muschel**, **Gesine Reinert**, **Bernadette J. Stolz** and **Ulrike Tillmann**, Centre for Topological Data Analysis, University of Oxford

## Introduction

**Understanding how the spatial structure** of blood vessel networks relates to their function in healthy and abnormal biological tissues could improve diagnosis and treatment for diseases such as cancer. Angiogenesis, the formation of new blood vessels from existing ones, is the process by which a tumour establishes its own supply of nutrients so that it may grow and spread to other parts of the body. The mathematical community has been studying the mechanisms of tumour-induced angiogenesis for more than 35 years [1–3].

New imaging techniques generate multiple, high-resolution images of the same tissue region, and show how vessel networks evolve during disease onset and treatment. Such experimental advances have created an exciting opportunity for discovering new links between vessel structure and disease through the development of geometric and topological algorithms that can analyse these rich data sets.

Here we explain how topological data analysis (TDA) can be used to study the structure of a vessel network. TDA is a growing field in the mathematical and computational sciences that consists of algorithmic methods for identifying global and multi-scale structures in high-dimensional data sets that may be noisy and incomplete [4–6].

### Tumour vessel networks

The structure of a vessel network can reveal the presence of an underlying disease or how a patient may respond to treatment. For example, tumour vasculature loses its hierarchical patterning, has different fractal dimensions, tortuous (i.e. bent) and enlarged vessels, and different paths for blood flow, which can inhibit nutrient and drug delivery [7]. In addition, oxygen availability, which determines treatment responsiveness, depends on the vessel network structure [8]. Studies of vessel networks use metrics such as inter-vessel spacing, number of branching points, vessel length density, tortuosity and fractal dimensions [9]. While these metrics can be related to tumour progression and treatment response [10], their values may be sensitive to the algorithms used to convert the raw images to network-based descriptions. Further, most measures characterise networks at a single spatial scale, although patterns may emerge at multiple spatial scales. TDA offers a promising and rigorous alternative for relating the structure of vessel networks obtained from raw images to their function and the disease status of the perfused tissue.

### Mathematical models of artificial vessel networks

A large number of mathematical models have been developed over the past 30 years to study tumour-induced angiogenesis [1–3] and to simulate artificial vessel networks. For example, now we have multi-scale agent-based models that simulate angiogenesis and vascular tumour growth [11–13]. These models have been used to investigate how vessel networks respond to anti-angiogenic and other vascular targeting agents. For our purposes, we can simulate the models under different conditions and obtain data, which we can then compare with experimental observations.

We remark that tumour-induced angiogenesis networks are inherently multi-scale ranging from arteries, to arterioles, to the smallest single-cell capillaries. Therefore, it is natural to apply a multi-scale method to analyse such data sets.

In this article, we present preliminary results where we apply TDA to tumour-induced vascular networks. First, we introduce topological data analysis and then describe the state-of-the-art angiogenesis data that are now being generated. Then we introduce angiogenesis data, give preliminary results and end with a short discussion.

## Topological data analysis

Data from biological processes, which are observed as physical objects, such as the vascular networks of interest here, are inherently spatial networks. It is, thus, natural to employ data science methods to analyse the geometric features of these networks.

### Mathematics

One of the most well developed tools in TDA is persistent homology, which is also our preferred method here [4–6]. We will describe below a generic procedure for associating a data set with its persistent homology barcode, its topological signature. As we will see, in applications this can flexibly be adjusted to suit different problems.

### From point clouds to topological spaces

Given a point cloud $P$ in Euclidean space $\mathbb{R}^N$ we can define its $\epsilon$-neighbourhood $N_\epsilon(P)$ to be the union of balls $B_\epsilon(p)$ of radius $\epsilon$ around all points $p$ in $P$:

$$N_\epsilon(P) = \bigcup_{p\in P} B_\epsilon(p), \qquad \epsilon \geq 0.$$

The space $N_\epsilon(P)$ has a topology that varies with $\epsilon \geq 0$ from a totally discrete space for $\epsilon = 0$ to one large featureless blob resembling a large ball for large $\epsilon$. It is the spaces between these extremes that are of interest. We seek to determine how their characteristic topological features change as $\epsilon$ increases.

### From topological spaces to combinatorial data

Topological spaces can be approximated by combinatorial data, so-called simplicial complexes $K = \{K_n\}_{n\geq 0}$. These are higher-dimensional analogues of graphs. Indeed, given a point cloud $P$ and $\epsilon \geq 0$, we construct the associated Vietoris–Rips complex $VR_\epsilon(P)$ by first building a graph with vertices $P$ and edges $(p_0, p_1)$ for all pairs of points in $P$ of distance $d(p_0, p_1) \leq \epsilon$. We then add an $n$-simplex for each complete subgraph on $n+1$ vertices. The combinatorial data are, thus, given by

$$VR_\epsilon(P) = \bigcup_{n\geq 0} VR_\epsilon(P)_n,$$
$$VR_\epsilon(P)_n = \{(p_0, \ldots, p_n) \mid d(p_i, p_j) \leq \epsilon \text{ for all } i, j\},$$

where $VR_\epsilon(P)_n$ lists all $n$-simplices of our simplicial complex.

**From combinatorial data to linear algebra**

Consider a simplicial complex $K$, and let $\mathbb{F}$ be a field; this could be the real or complex numbers but often for computational topology the field with two elements $\mathbb{F}_2 = \{0, 1\}$ given by mod 2 arithmetic is chosen. Define the $n$-chains of $K$ to be the vector space $\mathbb{F}[K_n]$ with basis the $n$-simplices, and a linear map $\partial_n : \mathbb{F}[K_n] \to \mathbb{F}[K_{n-1}]$ that takes an $n$-simplex to its boundary, the (alternating) sum of its faces. For the Vietoris–Rips complex, this is simply

$$\partial_n(p_0, \ldots, p_n) = \sum_{n \geq i \geq 0} (-1)^i (p_0, \ldots, \hat{p}_i, \ldots, p_n)$$

where $\hat{p}_i$ denotes that $p_i$ is removed.

A straightforward computation shows that applying the boundary operator twice gives the zero map: $\partial_n \circ \partial_{n+1} = 0$. This important algebraic identity reflects the geometric fact that the boundary of a boundary is always empty. Homology measures the difference between the cycles ($n$-chains with zero boundary) and the boundaries ($n$-chains that are boundaries of $(n+1)$-chains). For each $n \geq 0$, define the $n$th homology group:

$$H_n(K) := \frac{\ker \partial_n}{\operatorname{im} \partial_{n+1}}$$
$$b_n := \dim H_n(K) = \dim \ker \partial_n - \dim \operatorname{im} \partial_{n+1}.$$

The number $b_n$ is also called the $n$th Betti number: $b_0$ is the number of connected components of $K$ and, for $K$ a graph, $b_1$ is the total number of simple circuits. For more general $K$, the dimensions of the higher homology groups give a count of how many higher dimensional cavities there are. For example, the boundary of an $(n+1)$-simplex has a non-trivial homology group in dimension $n$ with Betti number $b_n = 1$.

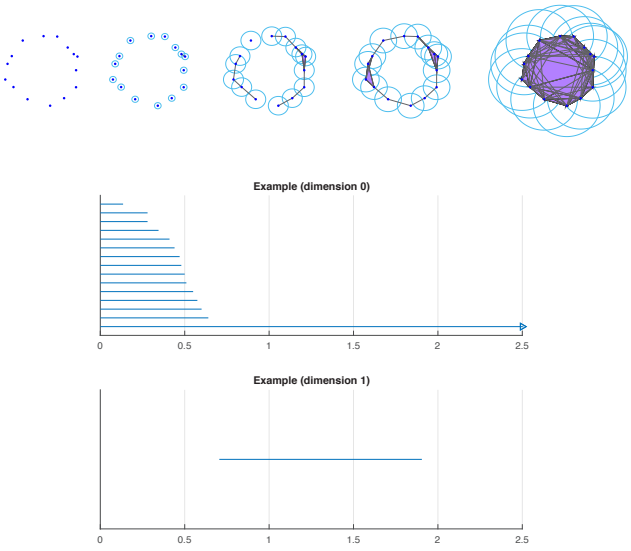



*Figure 1: An example of a Vietoris–Rips filtration with examples of the complex at $\epsilon = 0, 0.1, 0.55, 1, 2$ (top row). The corresponding barcodes in dimension 0 and dimension 1 are given where the horizontal axis gives the value of $\epsilon$.*

**Filtering and functoriality**

An important property of homology is functoriality: a map between simplicial complexes induces a map between their homologies. Now, for each $\epsilon' < \epsilon$, we have an inclusion of the corresponding Vietoris–Rips complexes $VR_{\epsilon'}(P) \subset VR_\epsilon(P)$ and, by functoriality, an induced map in homology. (Note, it is important here that we deal with homology groups and not just Betti numbers!) Thus, with growing $\epsilon$, we can track elements in the homology groups of the corresponding complexes $VR_\epsilon(P)$. A non-zero element that first appears (is born) at $\epsilon_{\text{birth}}$ and is mapped to zero (dies) at $\epsilon_{\text{death}}$ is represented by the interval (a bar) $[\epsilon_{\text{birth}}, \epsilon_{\text{death}})$. In $H_1$, this will correspond to the formation of a new circuit in the underlying graph at $\epsilon_{\text{birth}}$ and filling of that circuit with 2-dimensional simplices at $\epsilon_{\text{death}}$. The $n$th persistent homology $PH_n(P)$ of $P$ is the system of homology groups $H_n(VR_\epsilon(P))$ for all $\epsilon > 0$ with the induced maps.

There are two important results concerning persistent homology, which we describe in the following two theorems.

**Existence of barcode theorem:** *For each homological dimension $n$, compatible bases can be found so that for each $\epsilon$, the dimension of $H_n VR_\epsilon(P)$ is given by the number of corresponding bars containing $\epsilon$. In other words, the $n$th persistent homology can be represented faithfully by a barcode [14]; see Figure 1 for an example of a barcode.*

**Stability theorem:** *If two point clouds $P$ and $Q$ are close to each other (in terms of the Hausdorff distance), then the corresponding $n$-dimensional barcodes are also close (in terms of the bottleneck distance). In other words, the persistent homology transform $PH_n$ from point clouds to bars is continuous [15].*

**Adapting persistent homology to different problems**

Not all data sets come in the form of a point cloud and different applications demand a change in the general set-up. The main feature of persistent homology is that it can track the changing topology as the geometric object of interest is filtered in a suitable way. Above we considered a sequence of increasing Vietoris–Rips complexes as the radius of the little balls grew slowly until they eventually filled the whole space. Below we study larger and larger parts of the system of blood vessels in a tumour by growing the radius of vision (so just one ball) from the tumour centre.

## Statistical analysis in topological data analysis

Once a Vietoris–Rips complex or a related TDA summary of the data, such as a barcode, is obtained, statistics can be used to detect deviations from what would be expected. Although for most summaries such as Betti numbers, theoretical results about what to expect at random are available (see, for example, [16]), for most applications, these have assumptions that are too general and theoretical results are yet to be derived from more restrictive models.

In the absence of such results, simulations are employed for statistical inference. Suppose that data from a candidate model can be generated and the TDA summaries calculated. These summaries can then be used to indicate what is to be expected under the model. An observed summary outside (or close to the boundary of) the range of these simulated summaries indicates a deviation from the expectation. For vascular networks, theoretical models have been investigated (for example, in [11]) but depending on the complexity of the simulations, simpler models based on branching processes may be more appropriate. If it is not possible to simulate from appropriate models, then statistical machine learning offers a model-free approach to classifying data through using TDA summaries as features in an automated learning method such as random forests.

## Computation

Over the last decade, considerable effort has been spent developing and improving the computation of persistent homology. The first paper on the standard persistent homology algorithm was introduced in [14]. Since then, multiple software libraries have been developed that enable different types of data to be analysed, for example, point clouds, networks or images. The libraries are often specialised for certain filtrations. The data structures of the input and output differ for each library. Many of the libraries have software for computing statistical summaries of topological outputs. For a review of the software and a tutorial, see [17].

## Angiogenesis data

High-resolution 3D images showing vessel network evolution in mouse tumours undergoing treatment (with radiotherapy and vascular targeting agents) are being generated by Professor Ruth Muschel's research group (see Figure 2(a) for an example), while Professor Mike Brady's group (biomedical engineering) are developing new image segmentation tools to extract vessel structures from these images [6] (see Figure 2(b) for an example of the structures extracted). The imaging data here were obtained from tumours generated from murine colon carcinoma cells in mice genetically engineered to have fluorescent endothelium. Using video 2-photon microscopy, the entire tumour and its blood vessels were visualised daily after tumour induction.

The mice were subjected to different treatment regimens once the tumours reached a specified size: (i) controls, (ii) treatments known to increase vessel sprouting, (iii) treatments known to decrease vessel sprouting, (iv) treatment by single-dose irradiation ($1 \times 15\,\text{Gy}$) and (v) treatment by fractionated-dose irradiation ($5 \times 3\,\text{Gy}$). In addition to examining the structure of the vessels, vascular function was also evaluated. While statistical techniques can characterise facets of the networks (e.g. vessel lengths, radii and tortuosity), the analysis and interpretation of their topological features and how these vary across spatial scales remain open problems that TDA is ideally suited to address.

## Preliminary results

We characterise the unique features of tumour blood vessels, in particular the loops and the high degree of tortuosity, using persistent homology. We exploit the fact that our data are already in the form of a network rather than a point cloud.

Loops can be captured by persistent homology in dimension 1 of any filtration that is built using the structure of the vessel network data. Tortuosity has previously been successfully quantified when studying persistent homology in dimension 0 of a filtration that can be imagined as a stepwise sliding of a plane over a biological network. In the first filtration step, the entire network is situated on one side of the plane. As the plane moves, it starts intersecting the network until eventually the whole network is on the other side of the plane. The side of the plane that is initially empty thereby gives rise to a sequence of embedded objects that can be interpreted as a filtration of the network. This approach has been used to quantify the tortuosity of brain arteries [18] and the geometric structure of airways [19]. Kanari et al. [20] use a similar approach to classify the branching patterns of neurons using radial distances from the neuronal tree root, i.e. considering a sphere with decreasing radius around the root. These examples motivate the approach adopted here.

In contrast to brain arteries and neurons, tumour blood vessels are not tree-like objects and they do not have a natural orientation. By exploiting the fact that tumours are often viewed as spherical objects, we root our filtration in the tumour centre. Since we perform our analysis on the tumour blood vessels rather than the tumour itself, we approximate the tumour centre by the centre of mass of the blood vessel point data. We then search the neighbourhood of the centre point, increasing the radial distance stepwise and include all vessel points within the radius. If two points that are connected by an edge in the blood vessel data are within the given radius, we add the edge to our filtration.

Figure 3 shows a schematic of the radial filtration on blood vessel data. Based on this filtration, we study the topology of the growing network at every filtration step capturing tortuosity in dimension 0 and loops in dimension 1. Since the filtration

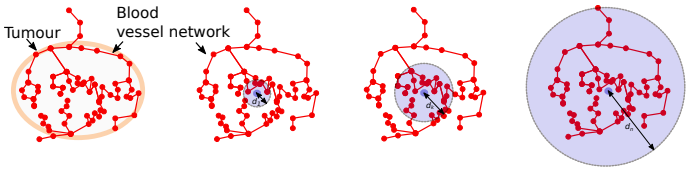
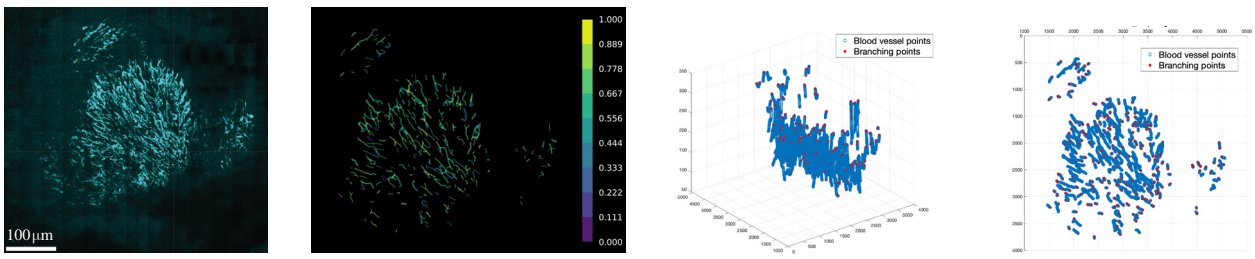



*Figure 3: Schematic of the radial filtration of a tumour blood vessel network. On the $k$th filtration step, we include all vessel nodes and edges that are fully contained in the purple ball of radius $d_k$ around the centre of mass of the vessel points.*




(a) (b) (c) (d)

*Figure 2: Example images of tumour blood vessels. (a) Tumour blood vessels as seen from above under the microscope. Slices of the tumour are imaged to allow full 3D reconstruction.[1] (b) View from above on 3D skeleton of tumour blood vessels coloured according to a measure of tortuosity.[2] (c) Vessel points we extracted from the 3D skeleton. Vessel points at the beginning or end of a branch (branching points) are shown in red. We interpret all vessel points as nodes and physical connections between them as edges in a blood vessel network, which is the input into our analysis. (d) Perspective corresponding to the 2D image shown in (a) and (b) on vessel points we extracted from the skeleton image.*

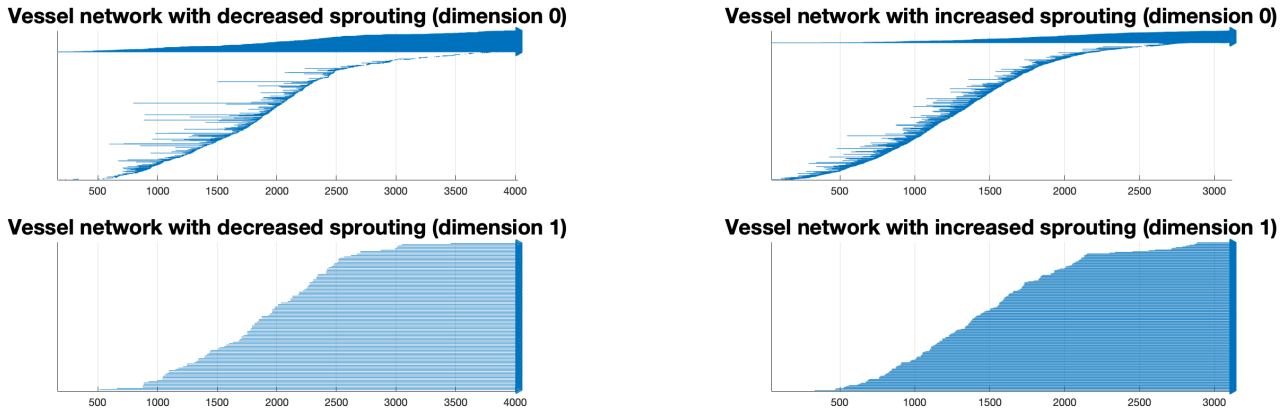



*Figure 4: Example barcodes from the radial filtration of networks subjected to two different treatment conditions. The vessel networks were imaged 3 days after treatment was administered. Each bar in the barcodes represents one topological feature as the radius increases (horizontal axis). The bar begins at the distance to the tumour centre at which the feature is first recorded and ends when that feature disappears in the filtration.*

generates topological information with respect to the tumour centre, we also obtain information about the heterogeneity of these characteristics within a single network.

In Figure 4 we present barcodes obtained from the radial filtration of tumour blood vessel networks that were subjected to different treatment regimes: (a) treatment to decrease vessel sprouting and (b) treatment to increase vessel sprouting.

We observe differences in the barcodes in both dimensions. We note also that the barcodes end at different filtration points, i.e. around 4000 for the decreased sprouting case and around 3000 for the increased sprouting case. These values reflect biological differences in the sizes of the two tumours. In dimension 0, where we expect to capture tortuosity, we find short-lived connected components, i.e. short bars, interspersed with longer-lived connected components, i.e. more persistent bars. In dimension 1, we see more loops in the network exposed to treatment that increases sprouting than the network exposed to treatment that decreases sprouting, as expected. This indicates that barcodes can capture the effects of treatment on both the sprouting behaviour and tortuosity of vessel networks.

## Discussion and outlook

In this article we have presented preliminary results that show how TDA can be used to quantify changes in the morphology of tumour vascular networks following exposure to treatments that alter the rate at which new vessels form. By employing TDA, we could analyse the vessels at multiple scales from the tumour centre, without having to select a threshold for network analysis. Thus, TDA is also a method for detecting parameter sensitivity in topological features, which may be useful for other multi-scale data sets. While at this stage differences in the spatial distribution of tortuosity or loops are not apparent, in future work, we will investigate these observations further by statistically analysing the full data set.

Moreover, we will use our existing multi-scale models of vascular tumour growth to generate artificial networks for analysis with TDA [11–13]. In this way, we aim to establish whether observed patterns (such as the persistence of cycles in the topological description of a network) can be related to the specific biophysical mechanisms relevant to the construction of vessel networks and investigate how the barcode associated with a particular tumour changes as the tumour and its vasculature evolve. We will also simulate the experimental image acquisition process by adding the noise associated with multi-photon microscopy to the artificial networks. An analysis of synthetic networks will be used to assess the robustness of TDA to image-processing artefacts.

In the longer term, the application of TDA to clinical images and spatial biomedical networks may reveal new relationships between network structure and treatment response. For example, in cancer, there is no consensus on vessel normalisation theory [21], which proposes that vascular targeting agents transiently improve tumour response to therapy by reverting an aberrant tumour vasculature to one resembling healthy tissue. TDA may be used to determine if characteristics more typical of healthy vasculature are being restored during treatment, thereby improving the diagnostic potential of existing imaging methods. In the future, we will also investigate TDA methods for other applications, such as wound healing, retinal pathology and cardiac disease.

## Acknowledgements

We would like to thank Russell Bates, James Grogan, Bostjan Markelc, Jakob Kaeppler and Nicola Richmond for helpful discussions. BJS gratefully acknowledges EPSRC and MRC grant (EP/G037280/1) and F. Hoffmann–La Roche AG for funding her doctoral studies. HAH gratefully acknowledges funding from a Royal Society university research fellowship. All the authors are members of the Centre for Topological Data Analysis, funded by the EPSRC grant (EP/R018472/1).

## Notes

1. Image taken by Bostjan Markelc and Jakob Kaeppler.
2. Skeleton and tortuosity values extracted by Russell Bates.

### References

1 Peirce, Sh.M. (2008) Computational and mathematical modeling of angiogenesis, *Microcirculation*, vol. 15, no. 8, pp. 739–751.

2 Mantzaris, N.V., Webb, S. and Othmer, H.G. (2004) Mathematical modeling of tumor-induced angiogenesis, *J. Math. Biol.*, vol. 49, no. 2, pp. 111–187.

3 Scianna, M., Bell, C.G. and Preziosi, L. (2013) A review of mathematical models for the formation of vascular networks, *J. Theor. Biol.*, vol. 333, pp. 174–209.

4 Edelsbrunner, H. and Harer, J. (2010) *Computational Topology*, American Mathematical Soc.

5 Carlsson, G. (2009) Topology and data, *Bull. Am. Math. Soc.*, vol. 46, no. 2, pp. 255–308.

6 Zomorodian, A.J. (2005) *Topology for Computing*, Cambridge.

7 Baish, J.W. et al. (1996) Role of tumor vascular architecture in nutrient and drug delivery: an invasion percolation-based network model, *Microvasc. Res.*, vol. 51, no. 3, pp. 327–46.

8 Grimes, D.R. et al. (2016) Estimating oxygen distribution from vasculature in three-dimensional tumour tissue, *J. R. Soc.*, vol. 13, no. 116, 20160070.

9 Tozer, G.M. et al. (2005) Intravital imaging of tumour vascular networks using multi-photon fluorescence microscopy, *Adv. Drug Deliv. Rev.*, vol. 57, pp. 135–152.

10 Jain, R.K. (1988) Determinants of tumor blood flow: a review, *Cancer Res.*, vol. 48, pp. 2641–58.

11 Perfahl, H. et al. (2017) 3D hybrid modelling of vascular network formation, *J. Theor. Biol.*, vol. 414, pp. 254–268.

12 Grogan, J.A. et al. (2017) Microvessel chaste: an open library for spatial modeling of vascularized tissues, *Biophys. J.*, vol. 112, no. 9, pp. 1767–1772.

13 Grogan, J.A. et al. (2016) Predicting the influence of microvascular structure on tumor response to radiotherapy, *IEEE Trans. Biomed. Eng.*, vol. 64, no. 3, pp. 504–511.

14 Zomorodian, A. and Carlsson, G. (2005) Computing persistent homology, *Discret. Comput. Geom.*, vol. 33, no. 2, pp. 249–274.

15 Cohen-Steiner, D., Edelsbrunner, H. and Harer, J. (2007) Stability of persistence diagrams, *Discret. Comput. Geom.*, no. 37, pp. 103–120.

16 Kahle, M. (2014) Topology of random simplicial complexes: a survey, *AMS Contemp. Math.*, vol. 620, pp. 201–222.

17 Otter, N. et al. (2017) A roadmap for the computation of persistent homology, *EPJ Data Sci.*, vol. 6, no. 1, pp. 1–38.

18 Bendich, P. et al. (2016) Persistent homology analysis of brain artery trees, *Ann. Appl. Stat.*, vol. 10, no. 1, pp. 198–218.

19 Belchi, F. et al. (2018) Lung topology characteristics in patients with chronic obstructive pulmonary disease, *Sci. Rep.*, vol. 8, no. 1, 5341.

20 Kanari, L. et al. (2018) A topological representation of branching neuronal morphologies, *Neuroinformatics*, vol. 16, no. 1, pp. 3–13.

21 Jain, R.K. (2005) Normalization of tumor vasculature: an emerging concept in antiangiogenic therapy, *Science*, vol. 307, no. 5706, pp. 58–62.