Supplementary information

Spatial quantification of clinical biomarker pharmacokinetics through deep learning-based segmentation and signal-oriented analysis of MSOT data

Bianca Hoffmann[a, °], Ruman Gerst[a,b, °], Zoltán Cseresnyés[a], WanLing Foo[c,d], Oliver Sommerfeld[c,d], Adrian T. Press[c,d,e], Michael Bauer[c,d], Marc Thilo Figge[a,d,f, *]

[a] Research Group Applied Systems Biology, Leibniz Institute for Natural Product Research and Infection Biology - Hans Knöll Institute (HKI), Beutenbergstr. 11a, 07745 Jena, Germany

[b] Faculty of Biological Sciences, Friedrich Schiller University Jena, Bachstr. 18k, 07743 Jena, Germany

[c] Department of Anesthesiology and Intensive Care Medicine, Jena University Hospital, Am Klinikum 1, 07747 Jena, Germany

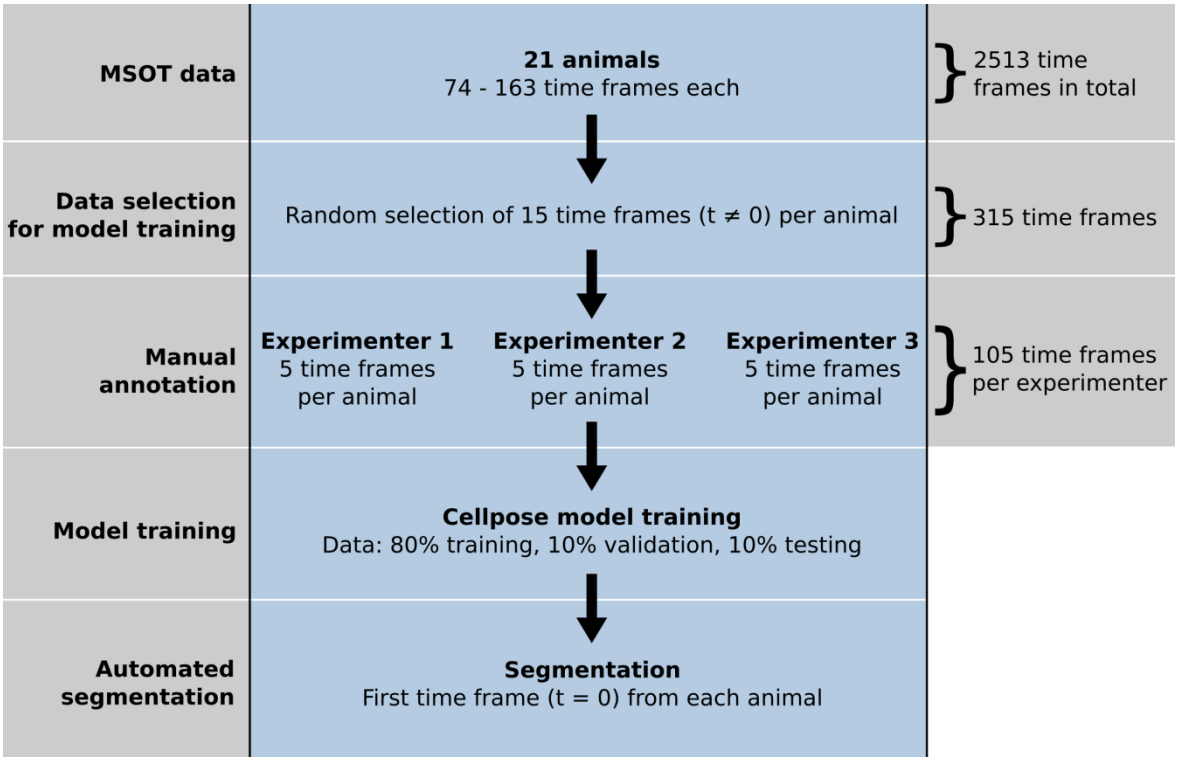
[d] Center for Sepsis Control and Care, Jena University Hospital, Am Klinikum 1, 07747 Jena, Germany

[e] Faculty of Medicine, Friedrich Schiller University Jena, Am Klinikum 1, 07747 Jena, Germany

[f] Institute of Microbiology, Faculty of Biological Sciences, Friedrich Schiller University Jena, Neugasse 25, 07743 Jena, Germany

[°] Authors contributed equally

* Correspondence: thilo.figge@leibniz-hki.de



Supplementary Figure A1: Schematic of training of deep learning-based segmentation.
Each of the 21 MSOT image stacks consisted of up to 163 time frames with a total of 2513 time frames. For each animal, 15 time frames were selected randomly, excluding the first time frame (t = 0), and distributed among the three experimenters. These 315 time frames were used as ground truth for the training of the Cellpose neural network and were divided into 80% for training, 10% for testing and 10% for validation. The derived model was then used for the final segmentation of the first time frame (t = 0) of each MSOT image stack.

Supplementary Table A2: Software and libraries used in *Mcat.*

Software	Version	URL
Apache Commons	2.6.0	https://commons.apache.org/
Apache Maven	3.6.3	http://maven.apache.org/
ImageJ	1.52r	https://imagej.net/
ImageScience	3.0.0	https://imagescience.org/
Imglib2	5.1.0	https://imagej.net/
FeatureJ	2.0.0	https://imagescience.org/meijering/software/featurej/
Flexmark	0.18.5	https://github.com/vsch/flexmark-java
Jackson (JSON)	2.11.2	https://github.com/FasterXML/jackson
Java	1.8	https://java.com/
JFreeChart	1.5.0	http://jfree.org/
JFreeSVG	3.4	http://jfree.org/
JGraphT	1.3.1	https://jgrapht.org/
IJP-Toolkit	2.1.2	https://github.com/ij-plugins/ijp-toolkit
MorphoLibJ	1.4.1	https://github.com/ijpb/MorphoLibJ/
MTrackJ	1.5.4	https://imagescience.org/meijering/software/mtrackj/
MultiStackRegistration	1.46.2	https://github.com/miura/MultiStackRegistration
RandomJ	2.0.0	https://imagescience.org/meijering/software/randomj/
SciJava	27.0.1	https://scijava.org/
Slf4J	1.7.9	http://www.slf4j.org/
SwingX	1.6.5-1	https://mvnrepository.com/artifact/org.swinglabs/swingx

Supplementary Information A3: Descriptive statistics

Signal-oriented analysis

AUC values

$\mu_{Sham} = 0.9$; $\sigma_{Sham} = 0.19$

$\mu_{PCI} = 1.46$; $\sigma_{PCI} = 0.35$

Tissue-oriented analysis

AUC of mean signal intensity values

$\mu_{Sham} = -0.3$; $\sigma_{Sham} = 0.64$

$\mu_{PCI} = 0.38$; $\sigma_{PCI} = 1.07$

AUC of maximum signal intensity values

$\mu_{Sham} = 46.73$; $\sigma_{Sham} = 15.36$

$\mu_{PCI} = 52.83$; $\sigma_{PCI} = 12.93$

AUC of 95^{th}-percentile signal intensity values

$\mu_{Sham} = 16.15$; $\sigma_{Sham} = 3.96$

$\mu_{PCI} = 19.3$; $\sigma_{PCI} = 4.03$

Robustness of signal-oriented analysis towards slight changes of ROIs

AUC values deep learning-based segmentation

$\mu_{Sham} = 0.9$; $\sigma_{Sham} = 0.18$

$\mu_{PCI} = 1.45$; $\sigma_{PCI} = 0.34$

AUC values experimenter 1

$\mu_{Sham} = 0.86$; $\sigma_{Sham} = 0.19$

$\mu_{PCI} = 1.45$; $\sigma_{PCI} = 0.36$

AUC values experimenter 2

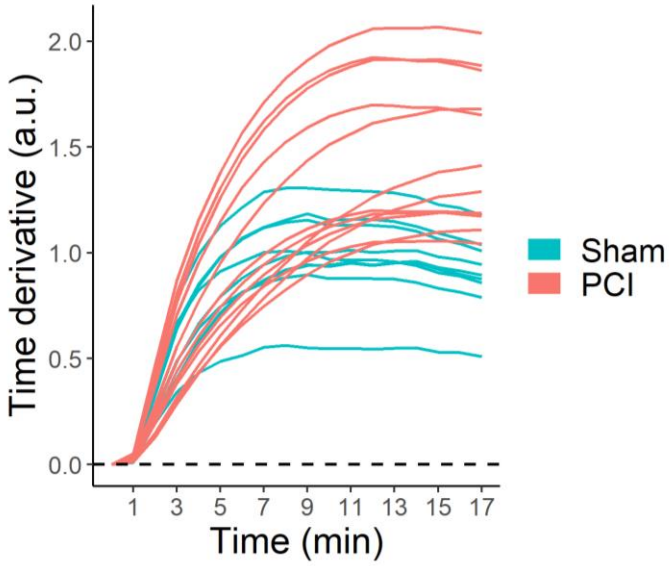
$\mu_{Sham} = 0.88$; $\sigma_{Sham} = 0.2$

$\mu_{PCI} = 1.46$; $\sigma_{PCI} = 0.36$

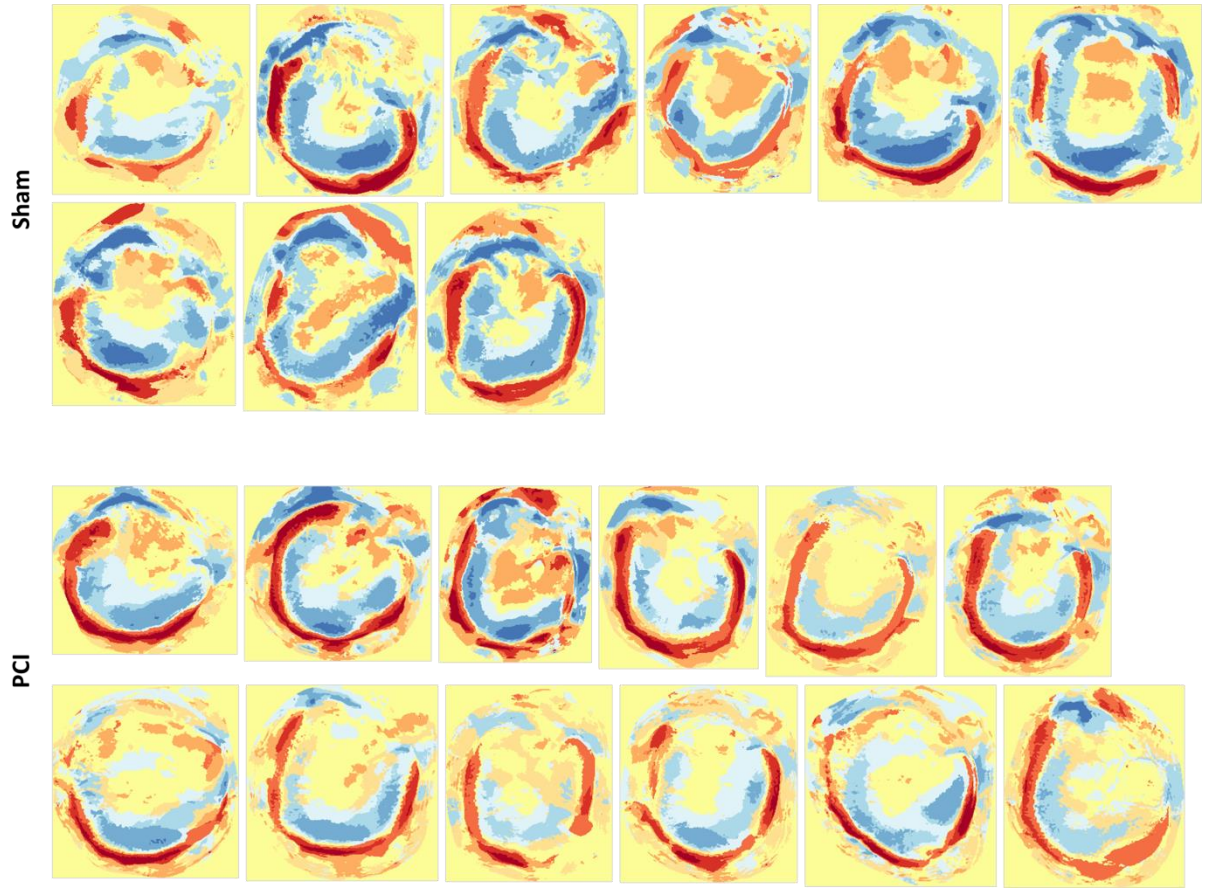
AUC values experimenter 3

$\mu_{Sham} = 0.85$; $\sigma_{Sham} = 0.19$

$\mu_{PCI} = 1.4$; $\sigma_{PCI} = 0.34$

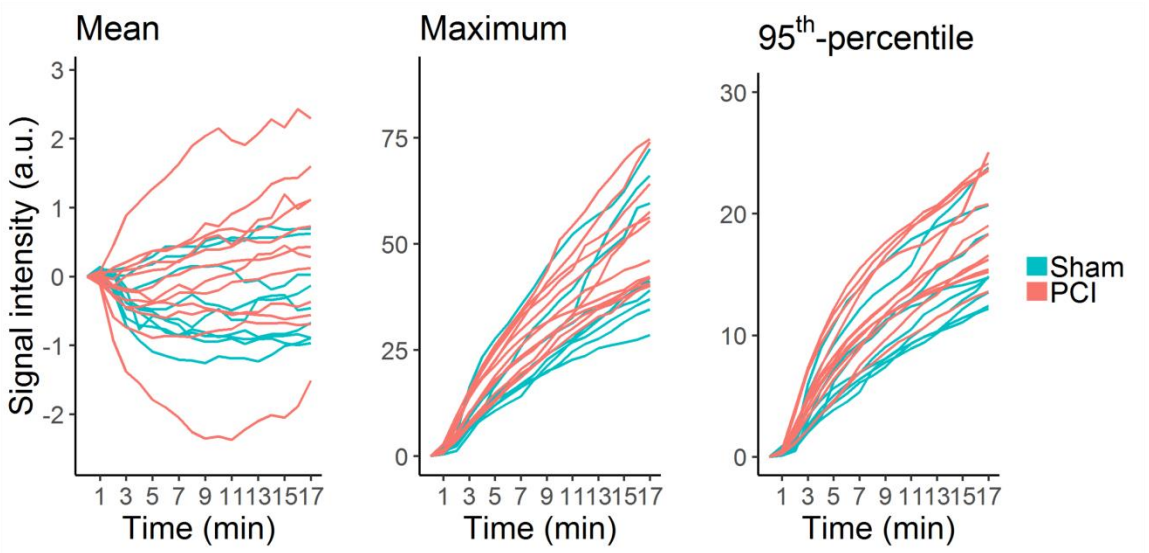


Supplementary Figure A4: Integrated signal of WACs from the signal-oriented analysis.
The time derivative values of the WACs were integrated to obtain the original kinetic shape of the curves for the signal-oriented analysis. The curves of Sham and PCI animals can visually be well distinguished. Seven of the twelve PCI animals show a delayed signal increase and lower plateau values.



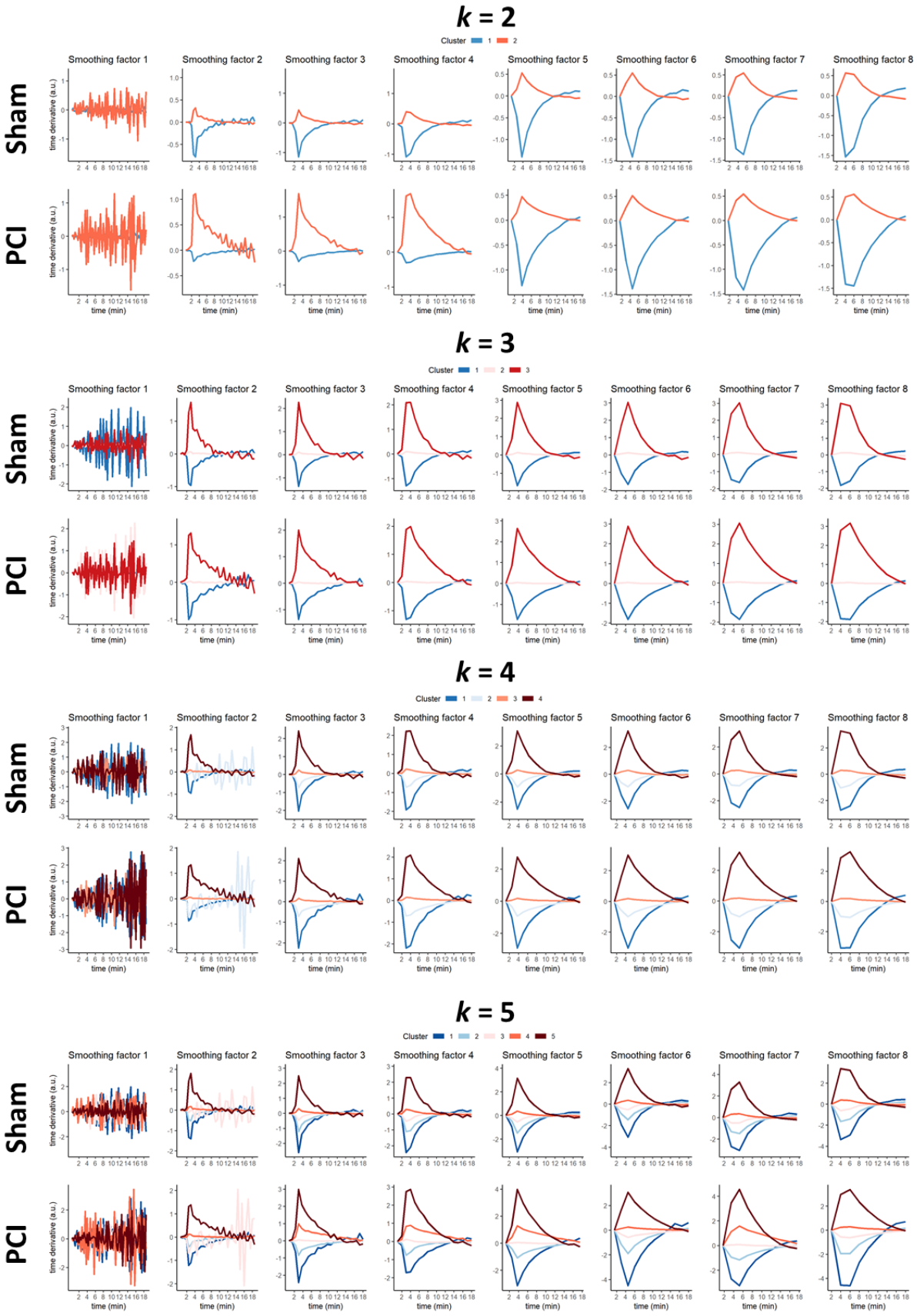
Supplementary Figure A5: Spatial distribution of kinetic clusters.

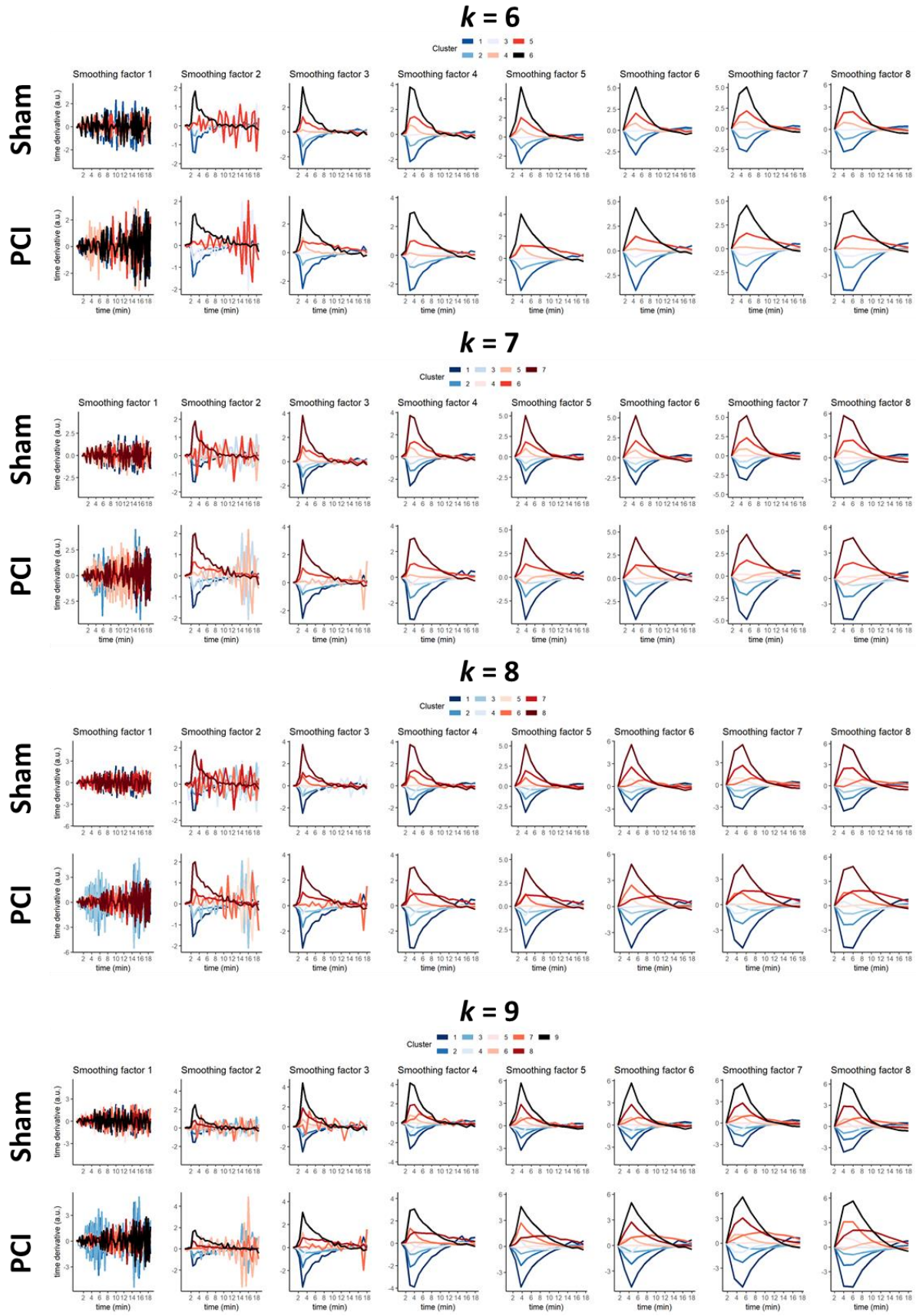
Colour-coded images reveal the biodistribution of ICG. The two top rows show all colour-coded images for Sham animals, the two bottom rows for PCI animals. The biodistribution is visually not different between Sham and PCI animals.

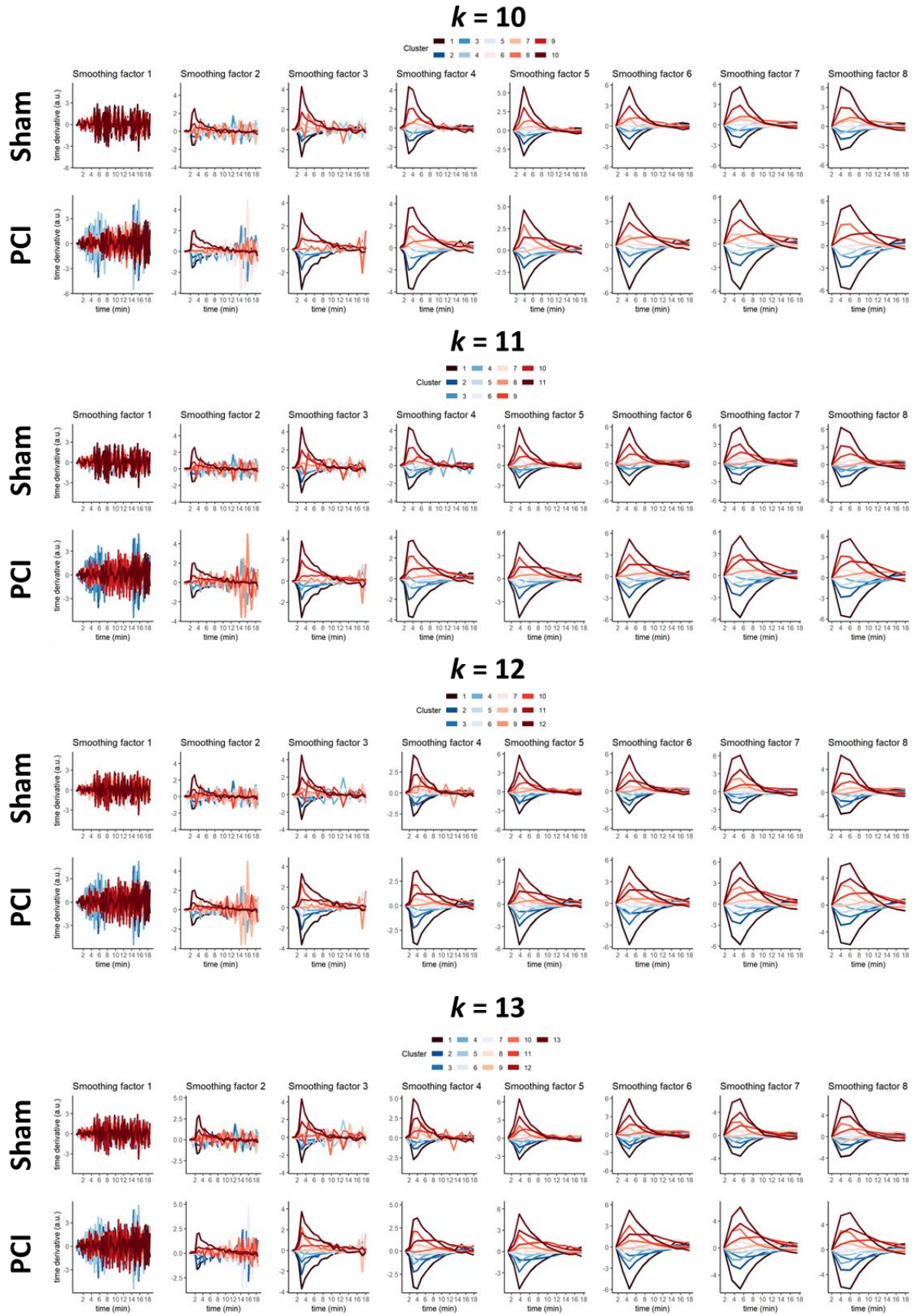


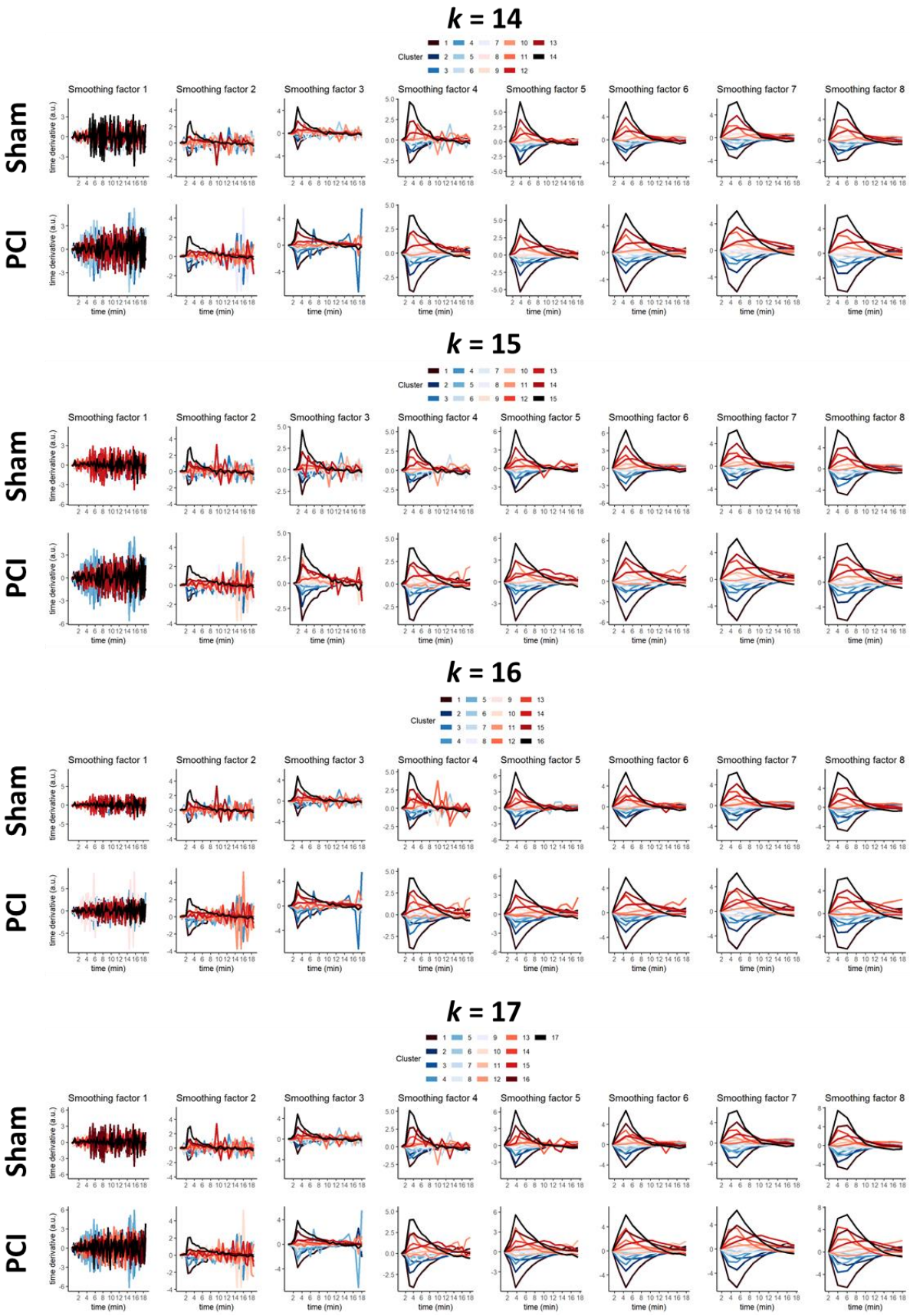
Supplementary Figure A6: Integrated signal of time derivative values from tissue-oriented analysis.

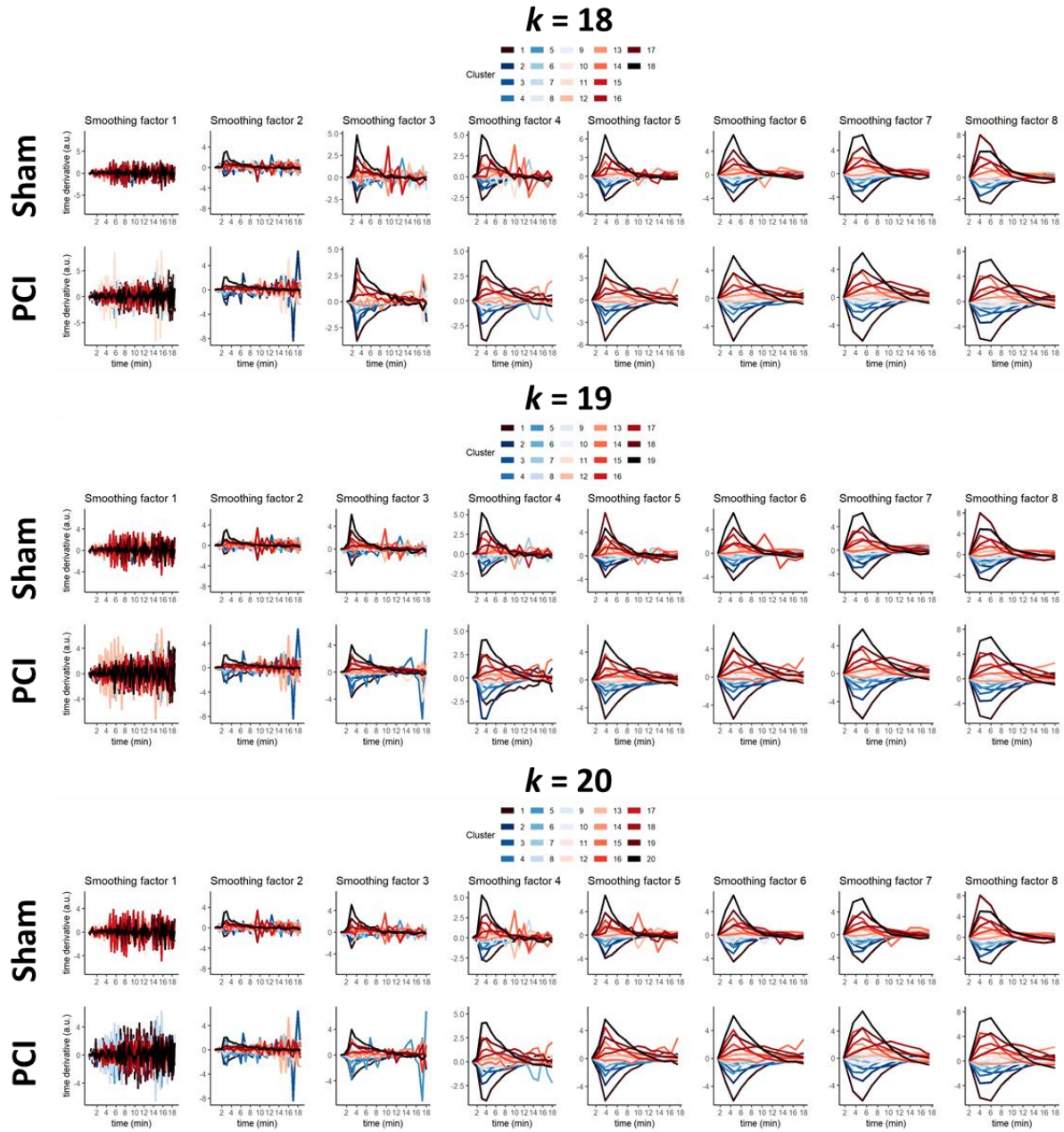
The time derivative values were integrated for the three different intensity features of the tissue-oriented analysis to obtain the original kinetic shape of the ICG signal uptake. The time curves of Sham and PCI animals are strongly overlapping, making visual distinction impossible.


k = 2
Cluster 1 2
k = 3
Cluster 1 2 3
k = 4
Cluster 1 2 3 4
k = 5
Cluster 1 2 3 4 5
Sham
PCI
Smoothing factor 1
Smoothing factor 2
Smoothing factor 3
Smoothing factor 4
Smoothing factor 5
Smoothing factor 6
Smoothing factor 7
Smoothing factor 8
time derivative (a.u.)
time (min)


k = 6
Cluster 1 2 3 4 5 6
Smoothing factor 1
Smoothing factor 2
Smoothing factor 3
Smoothing factor 4
Smoothing factor 5
Smoothing factor 6
Smoothing factor 7
Smoothing factor 8
Sham
PCI
time derivative (a.u.)
time (min)
k = 7
Cluster 1 2 3 4 5 6 7
k = 8
Cluster 1 2 3 4 5 6 7 8
k = 9
Cluster 1 2 3 4 5 6 7 8 9


k = 10
Cluster 1 2 3 4 5 6 7 8 9 10
Sham
PCI
Smoothing factor 1
Smoothing factor 2
Smoothing factor 3
Smoothing factor 4
Smoothing factor 5
Smoothing factor 6
Smoothing factor 7
Smoothing factor 8
time derivative (a.u.)
time (min)
k = 11
Cluster 1 2 3 4 5 6 7 8 9 10 11
k = 12
Cluster 1 2 3 4 5 6 7 8 9 10 11 12
k = 13
Cluster 1 2 3 4 5 6 7 8 9 10 11 12 13

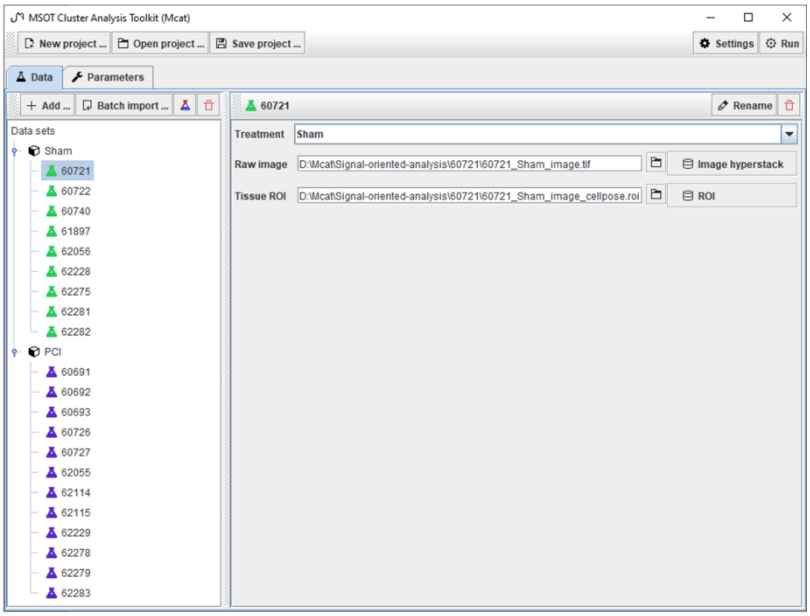

k = 14
Cluster 1 2 3 4 5 6 7 8 9 10 11 12 13 14
k = 15
Cluster 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
k = 16
Cluster 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16
k = 17
Cluster 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17
Sham
PCI
Smoothing factor 1
Smoothing factor 2
Smoothing factor 3
Smoothing factor 4
Smoothing factor 5
Smoothing factor 6
Smoothing factor 7
Smoothing factor 8
time derivative (a.u.)
time (min)



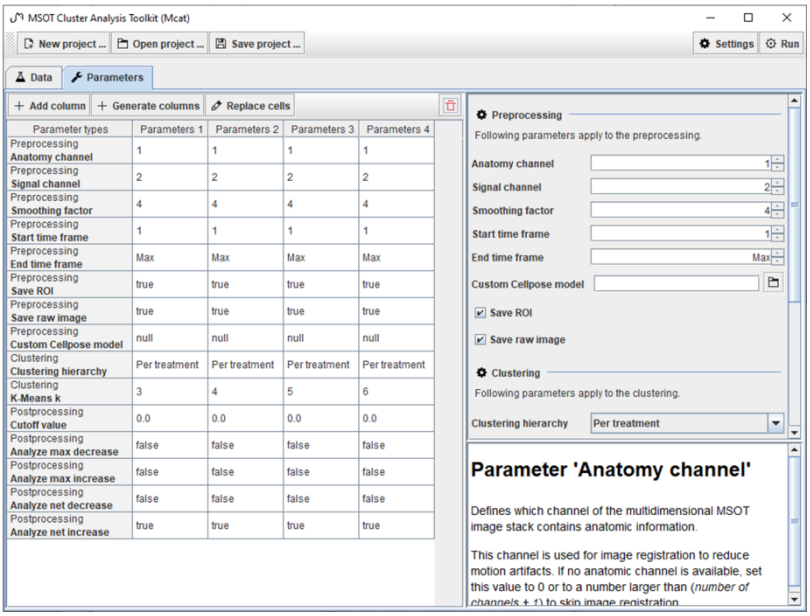
Supplementary Figure A7: Curves of kinetic clusters for Sham and PCI treatment with different values of k-means *k*.

The curves of the kinetic clusters as found by pixel-wise clustering with varying values for *k* are shown for the smoothing factor *s* ranging from 1 to 8 and k-means *k* ranging from 2 to 20. Low values of *s* show strong breathing artefacts in the curves, while increasing *s* leads to decreasing temporal resolution. Large values of *k* result in curves with apparent signal fluctuations also for larger values of *s*, which are expected due to overfitting.

a



b



Supplementary Figure A8: Graphical user interface of *Mcat* software toolkit.

The main graphical user interface is divided into a *Data* **(a)** and a *Parameters* **(b)** view. In the *Data* view, data can be imported manually or by a batch importer functionality, and imported sample data can be edited. Analysis parameters can be adjusted in the *Parameters* view. A short description of each parameter is shown in the bottom right part of the view when hovering over a parameter name.

Supplementary Information A9: Documentation of *Mcat* parameters

The MSOT clustering analysis toolkit (*Mcat*) comprises three main steps: (i) pre-processing of the MSOT image stacks, (ii) pixel-wise clustering of kinetics and (iii) post-processing of the clustering result. For these steps, the following parameters can be adjusted:

i) Pre-processing

anatomic channel

Defines which channel of the MSOT image stack contains the anatomic information (first channel is indexed as 1). This channel is used for image registration to reduce motion artefacts. If no anatomic channel is available, set this value to 0 to skip image registration.

signal channel

Defines which channel of the MSOT image stack contains the signal that should be analysed. This parameter has to be set to a value between 1 and the number of image channels.

smoothing factor s

Defines how strongly image data is smoothed in the time domain. This parameter is used to reduce breathing artefacts by performing downsampling and averaging over *s* consecutive time frames. Set this value to 1 if the image data shall not be smoothed. Smaller values lead to increased computation time and possibly strong fluctuations in the extracted kinetic curves, while larger values decrease computation time while reducing the time resolution.

start time frame and end time frame

Define which time frames are taken into account. *Start time frame* and *end time frame* have to be in the range of 0 and number of time frames. The stack will be cropped to the time range [*start time frame*, *end time frame*].

save ROI and save raw image

Define if raw data is saved to the output folder. If *save ROI* is enabled, the original or derived ROI file will be stored in the output folder. If *save raw* image is enabled, the original raw image will be stored in the output folder.

custom Cellpose model

If a custom Cellpose model shall be used for the segmentation, specify the path to the model file here. Don’t rename Cellpose model files as the name encodes important model information.

ii) Clustering

k-means k

Defines how many kinetic clusters are extracted. This parameter controls how many clusters are used when performing k-means clustering. Smaller values can lead to bad approximation of the

true kinetics present in the image, while larger values can lead to overfitting to confined regions with possibly strongly fluctuating signal intensities.

clustering hierarchy

Defines how samples are grouped when performing k-means clustering. This parameter can be set to one of the three options: *per subject*, *per treatment* or *all in one*.

If *per subject* is used, clustering is performed for each subject individually, resulting in individual kinetic clusters for each subject. This setting can be used to see if subjects show different kinetics. If *per treatment* is used, all subjects from one treatment are grouped together and clustering is performed per treatment. This setting can be used to compare treatments against each other. If *all in one* is used, all subjects are grouped together for clustering. This setting can be used to obtain main kinetic clusters for the whole dataset and examine how the clusters are distributed across samples.

iii) Post-processing

cutoff value

Defines which part of the kinetic curves is used for calculation of area under the curve (AUC) statistics. This parameter can be set to a floating point number between 0 and 1 and controls which portion at the beginning of the curves is excluded from AUC calculation. If this parameter is set to 0, the whole curve is taken into account. If it is set to 0.5, the first half of the curve is omitted. This option can for example be useful to analyse initial dye uptake and dye excretion separately or to limit a study to certain phases of signal development.

analyse max decrease

Defines if the kinetic cluster with the maximum signal decrease should be used for AUC calculation. If it is enabled, a weighted curve is calculated for each subject with the weight being the corresponding pixel abundance value for the kinetic cluster with the maximum signal decrease. These curves are then normalised by the total number of pixels of the respective subjects and the AUC is calculated, taking into account the cutoff value.

analyse net decrease

Defines if all kinetic clusters with a signal net decrease should be used for AUC calculation. If it is enabled, a weighted curve is calculated for each subject with the weights being the corresponding pixel abundance values for all kinetic clusters with a signal net decrease. These curves are then normalised by the total number of pixels of the respective subjects and the AUC is calculated, taking into account the cutoff value.

analyse max increase

Defines if the kinetic cluster with the maximum signal increase should be used for AUC calculation. If it is enabled, a weighted curve is calculated for each subject with the weight being the corresponding pixel abundance value for the kinetic cluster with the maximum signal

increase. These curves are then normalised by the total number of pixels of the respective subjects and the AUC is calculated, taking into account the cutoff value.

analyse net increase

Defines if all kinetic clusters with a signal net increase should be used for AUC calculation. If it is enabled, a weighted curve is calculated for each subject with the weights being the corresponding pixel abundance values for all kinetic clusters with a signal net increase. These curves are then normalised by the total number of pixels of the respective subjects and the AUC is calculated, taking into account the cutoff value.

Supplementary Information A10: Application of DL-based segmentation and signal-oriented analysis to different MSOT data and photoabsorber

To evaluate the applicability of our DL-based segmentation for other MSOT data and our signal-oriented analysis for a photoabsorber other than ICG, we used Mcat to analyse MSOT data that we derived from another study.

In brief, eleven FVB/N mice of mixed gender and aged older than eight weeks were anesthetised and shaved thoroughly for the whole abdomen area using shaver and commercial hair removal cream. After insertion of a tail-vein catheter, mice were placed in a Multispectral Optoacoustic Tomography inVision 256-TF machine (iTheraMedical, Germany) equipped with laser wavelength from 680-980 nm. Mice were anaesthetised with 1.5 to 2 % of isoflurane vaporised in oxygen throughout the preparation and imaging process. After 2 min of baseline image acquisition, 30 µg polyplex micelles loaded with BHQ3 labelled siRNA were injected intravenously through the tail vein catheter. Imaging was continued for 45 min at 6 wavelengths (680 nm, 700 nm, 720 nm, 760 nm, 800 nm, 900 nm) and captured two cross-sectional frames, i.e. liver and kidney. The acquired raw MSOT images were reconstructed by model-based backprojection (filter range: 50kHz to 6.5MHz) using the proprietary software ViewMSOT v3.8.1.04 (iTheraMedical, Munich, Germany), and spectrally unmixed into 4 channels (water, BHQ3, deoxygenated blood, oxygenated blood) by linear regression.

As a proof of principle, we performed the automated segmentation of the liver in the water channel with the inbuilt DL model and compared it to manual segmentation. We then applied the signal-oriented analysis to extract five kinetic clusters for the channel representing oxygenated blood for both DL-based and manual segmentation. Except for the channel of interest, all other settings were kept at their default values.

The comparison of DL-based segmentation and manual segmentation was carried out as described in the main manuscript in section 3.1. The resulting Dice scores for all animals are shown in Fig. A10a. The Dice score for nine out of the eleven animals was comparable to the concordance found between experimenters 2 and 3 in the main manuscript (see Fig. 2 in main manuscript). Two animals reached clearly lower Dice scores. While large parts of the animal outlines were identified correctly in these cases, the upper left region was not segmented properly. This area typically has low resolution of the animal outline because it spatially corresponds to the 90-degree “dead space”, where no ultrasonic detectors are placed due to the need to provide physical access to the sample. Representative examples of manual and DL-based segmentation are shown in Fig. A10b, including one of the cases with a lower Dice score at the very right.

The results of the signal-oriented analysis for DL-based and manual segmentation are shown in Supplementary Figs. A10c and A10d. The extracted kinetic clusters were virtually identical for both DL-based and manual segmentation. Also the AUC values that were calculated for the kinetic clusters reflecting a signal net increase (see section 3.2 of main manuscript for details on calculation) were found to be very similar (Wilcoxon rank sum test $p = 0.89$, Hedges' $g = 0.018$).

These findings demonstrate the general applicability of our inbuilt DL-based segmentation for similar MSOT images and highlight that our signal-oriented analysis does not rely on exact ROI definition but provides results that are robust against the various ways of determining the mouse ROI.

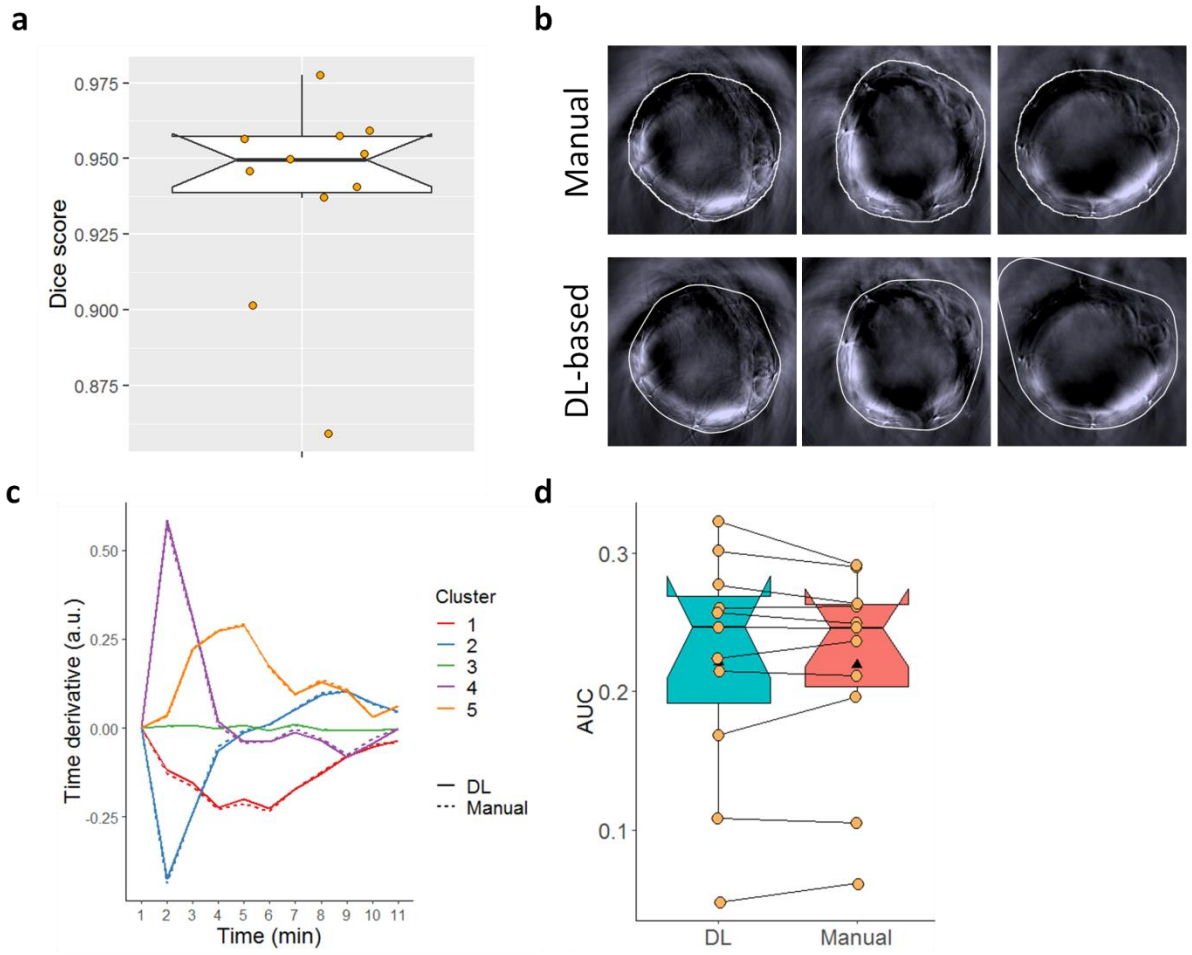


Figure A10: Application of DL-based segmentation and signal-oriented analysis to other MSOT data.

(a) The Dice score between manual and DL-based segmentation shows a high concordance for nine out of the eleven animals. **(b)** This is also supported by the visual comparison of representative example segmentations. In two cases, the upper left animal outline was not segmented properly with the inbuilt DL model (one example shown at the very right). **(c)** The comparison of extracted kinetic clusters and **(d)** corresponding AUC values of curves reflecting a net increase shows similar results for DL-based and manual segmentation (statistical comparison of AUC values: $p = 0.89$, $g = 0.018$).