

Appendix A. Supplementary data

Cytogenotoxic effects of polycyclic aromatic hydrocarbons complex mixture in human peripheral blood, lung A549 and liver HepG2 Cells: Translation of a real-scenario exposure to *in vitro*

Luka Kazensky¹, Marija Jelena Lovrić Štefiček², Vilena Kašuba¹, Matjaž Novak³, Karolina Belingar³, Katarina Matković¹, Marko Gerić¹, Jasmina Rinkovec², Ivana Jakovljević², Katarina Baralić⁵, Danijela Đukić-Ćosić⁵, Mirta Milić¹, Želimir Jelčić⁶, Gordana Pehnec², Bojana Žegura^{3,4,*} & Goran Gajski^{1,*}

¹*Division of Toxicology, Institute for Medical Research and Occupational Health, Zagreb, Croatia*

²*Division of Environmental Hygiene, Institute for Medical Research and Occupational Health, Zagreb, Croatia*

³*Department of Genetic Toxicology and Cancer Biology, National Institute of Biology, Ljubljana, Slovenia*

⁴*Biotechnical Faculty, University of Ljubljana, Ljubljana, Slovenia*

⁵*Department of Toxicology “Akademik Danilo Soldatović”, Faculty of Pharmacy, University of Belgrade, Belgrade, Serbia*

⁶*TAPI R&D, PLIVA Croatia Ltd., Zagreb, Croatia*

***Correspondence:** Goran Gajski, Institute for Medical Research and Occupational Health, Division of Toxicology, Ksaverska cesta 2, 10000 Zagreb, Croatia; ggajski@imi.hr & Bojana Žegura, National Institute of Biology, Department of Genetic Toxicology and Cancer Biology, Večna pot 121, 1000 Ljubljana, Slovenia; bojana.zegura@nib.si

Quantitative comet assay image analysis

Contents

Quantitative comet assay image analysis	5
DNA damage.....	5
Image analysis.....	6
Collage of single images.....	8
4h	8
24h	10
Examples of single and image collage 3D images	11
4h	11
24h	13
Image analysis features: Scatter percentile distribution plots of	15
Correlation scatter plots of properties	19
Textural analysis: Haralick features.....	20
Textural analysis	20
Fractal analysis	22
Monofractals: Characterizing space	22
Range Fractal indices of separate subImages	23
<i>Control K</i>	23
<i>Exposure L</i>	25
<i>Exposure M</i>	27
<i>Exposure H</i>	29
<i>Positive control PK</i>	31
Range Fractal indices of Image Collage for 4h and 24 h.....	33
<i>Fractal indices for different exposures, L, M, H</i>	35
<i>Fractal indices for control (K), positive control (PK) and different exposures, L, M, H</i>	36
Multi-scale Focal fractal dimension.....	37
Multifractals: A theory of measures	41
Multifractals	42
The generalized dimension DQ spectra for separate images.....	44
The multifractal spectra for separate images.....	45
The length of the left (multifractal) and right (monofractal) sides of the multifractal spectrum.....	47

Euler number in COMET ASSAY images	47
Comparison with Betti Numbers	48
Complementary Features	48
Results	48
Minkowski scalars	52
Isoperimetric Ratio	52
Examples:	53
Circle:	53
Square:	53
Minkowski Structure Metrics qs	54
Shannon Entropy of Collage Images.....	58
Shannon Entropy of red, green and blue (RGB) colors in Collage Images.....	62
Image Vectorscope Plot	65
Image Vectorscope Plots and 3D chromaticity of these plots.....	66
Graph morphology of comet assays images	71
Morphological Graph	73
Results.....	73
Correlation dimension	74
Results.....	74
Graph Density	77
Graph Distance Matrix, entropy.....	79
Results.....	79
Topological measures, Centrality measures.....	82
Empirical global clustering coefficient.....	83
Results.....	83
Theoretical global clustering coefficient.....	85
Results.....	85
Degree-based Centrality Metrics	87
Results.....	87
Betweenness centrality	88
Results.....	88
Betweenness centrality	88
<i>Shannon Entropy of Normalized</i>	89
Edge Betweenness Centrality.....	90

Eigenvector centrality	94
Results.....	95
Information centrality	97
Results.....	97
Eccentricity	103
Results.....	103
Assortativity	106
Results.....	106
Adjacency matrix.....	109
Results.....	109
Weighted adjacency matrix.....	118
Results.....	118
Incidence Matrix.....	124
Results.....	124
Joint Degree Matrix	125
Results.....	126
Kirchhoff Matrix	131
Entropy of a Kirchhoff matrix	131
Results.....	131
Radiality centrality	135
Results.....	135
Voronoi entropy for a nearest-neighbour graph.....	138
Results.....	138
Voronoi entropy for a r-nearest-neighbour graph	139
Voronoi entropy for a k-nearest-neighbour graph.....	140
Persistent homology.....	142
Persistent Life=Death-Birth	144
Persistence diagrams: Death vs. Birth.....	144
Persistence diagrams: Life vs. Birth	145
Persistence diagrams	147
Barcodes	150
Shannon persistent entropy	151
Discussion.....	154
Conclusions.....	155

References 156

This study investigates the multifractal properties of comet assays from the human donor, focusing on potential morphological changes related to pollution induced by PAH and spatial differences within the Comet assays. Using multifractal spectra, we analysed various parameters, including generalized dimensions and Hölder exponents, to characterize the Comet assays' morphology. Despite the detailed analysis, no significant correlation was found between morphology and pollution induced by PAH. However, clear morphological differences were seen between pollution induced by PAH after 4h and 24h. Comet assays after 24h displayed higher morphological complexity and greater local homogeneity, while those after 4h exhibited more scaling laws and higher local heterogeneity.

Further studies are recommended to confirm these findings and to explore the full potential of multifractal analysis in comet assays research.

The calculation of the multifractal spectrum of comet assays images is ineffective due to artifacts that can lead to an unreliable result. The search for methodological approaches to using multifractal image analysis in relation to the assessment of the comet assays is relevant.

Quantitative comet assay image analysis

Quantitative comet assay analysis^{1,2,3} is carried out in 2D images⁴. Initial attempts at comet assay measurement utilised a manual approach and were tedious and time consuming. However, comet assay characterisation based on visual analysis suffers from user bias and reproducibility. In the last two decades, there has been significant development in automated quantitative comet assay analysis methodology. Methods of comet assay analysis require ideal imaging based on size (e.g., microscopes with digital cameras).

DNA damage

DNA damage is a common cellular feature seen in cancer and neurodegenerative disease⁵. Comet assays^{6,7,8} are a biochemical tool to measure DNA damage in single cells through microscopy assessment based on the migration of broken DNA strands towards a positive electrode, which creates a quantifiable “tail”. Comet visualization allows the observation of stained DNA either as dots (representing intact DNA) or as comets (indicating damaged DNA).

The main obstructive point is the image analysis of COMET ASSAY images is that the COMET ASSAY parameters: Tail Intensity, Tail Length, Tail Moment and Total Area do not correlate well with the exposure level and time. The positive control (PK) COMET ASSAY parameters differ from the other experiments by factor 1.3 up to 6.7; those, we did not take these values in analyses.

From Figure 1-Figure 2 we can conclude that there is some effect of exposure level at 4h.

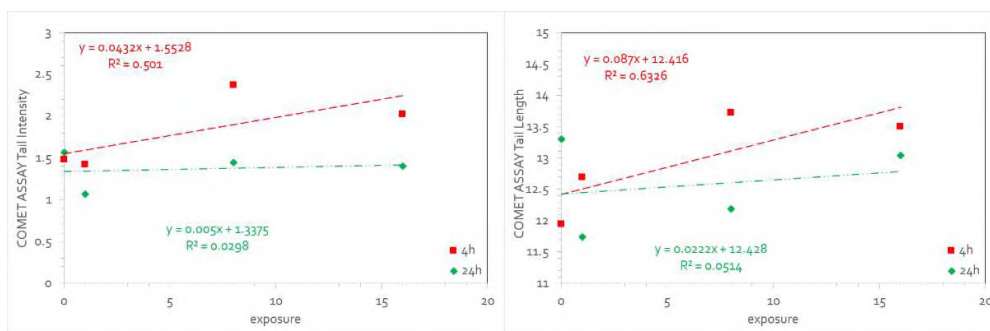


Figure 1 The COMET ASSAY Tail Intensity versus exposure (left) and COMET ASSAY Tail Length versus exposure (right) at 4h (red) and 24h (green).

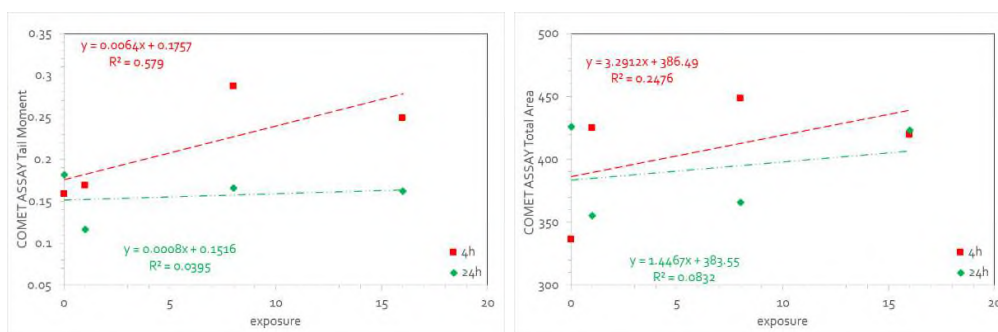


Figure 2 The COMET ASSAY Tail Moment versus exposure (left) and COMET ASSAY Total Area versus exposure (right) at 4h (red) and 24h (green).

Image analysis

After identifying parts in the 10 images, their properties were measured with the Mathematica function **ComponentMeasurements**. This function may produce a dataset listing, 83 properties. Examples of properties are *Centroid*, "centre of mass coordinates", *Circularity*, "ratio of equivalent disk perimeter to the perimeter length", *Holes*, "number of holes in each part", and *Shape*, "shape of the component given as a binary image". (Quotes are from documentation of the Mathematica function **ComponentMeasurements**.) For the following analysis we discarded all but scalar properties, i.e., we kept only properties whose value is a single number. Thus, *Centroid* and *Shape* were discarded at this step, leaving me with 61 properties. We also discarded properties that didn't vary. For example, *Holes* was discarded at this stage because it was 0 for all images. This left us with 48 measurements. Many of these 48 measurements were highly correlated with others. Measurements were easily interpretable. For example, 24 properties, had properties that in effect measure the size of the objects. We also discarded *Orientation*.

Elongation and *CaliperElongation* are zero for a perfect circle. Thus, in many images these properties have values near zero, and a small change in the way

elongation is measured can produce large relative change in these small numbers.

The eccentricity of a best-fit ellipse, denoted by **e**, is a measure of its deviation from a perfect circle, calculated as the ratio of the distance from the centre to a focus (**c**) to the length of the semi-major axis (**a**), or by the formula $\mathbf{e} = \sqrt{1 - (\mathbf{b}/\mathbf{a})^2}$ where **b** is the semi-minor axis length. The value of **e** ranges from 0 to 1; an eccentricity of 0 indicates a circle, while a value closer to 1 signifies a more elongated and flattened ellipse. Using the focus (**c**) and semi-major axis (**a**): $\mathbf{e} = \mathbf{c}/\mathbf{a}$.

Table 1

Area Radius Coverage	fraction of pixels within the equivalent disk radius
Area	approximate area, where each pixel area is weighted by its neighbourhood configuration
Perimeter Length	total length of outer pixel sides
Parameters of the oriented best-fit ellipse	
Length	largest axis of the best-fit ellipse
Width	smallest axis of the best-fit ellipse
Semi Axes	lengths of the semiaxes of the best-fit ellipse
Orientation	angle between the largest axis and the horizontal axis
Elongation	elongation, computed as
Eccentricity	eccentricity of the best-fit ellipse
Convex Area	area of the convex hull polygon
Convex Perimeter Length	length of the convex hull polygon
Skew	asymmetry in intensity distribution
Circularity	$2\pi r/p$, with polygonal length <i>p</i> and equivalent disk radius <i>r</i>

Collage of single images

Collage of single images, a composite picture including up to 10 fields of view (single images), was constructed by aligning each field of view according to a simple algorithm. One field of vision includes a matrix of 720x 511 pixels, *i.e.*, separate columns and rows.

4h



Figure 3 KPAH4K



Figure 4 KPAH4L



Figure 5 KPAH4M



Figure 6 KPAH4H

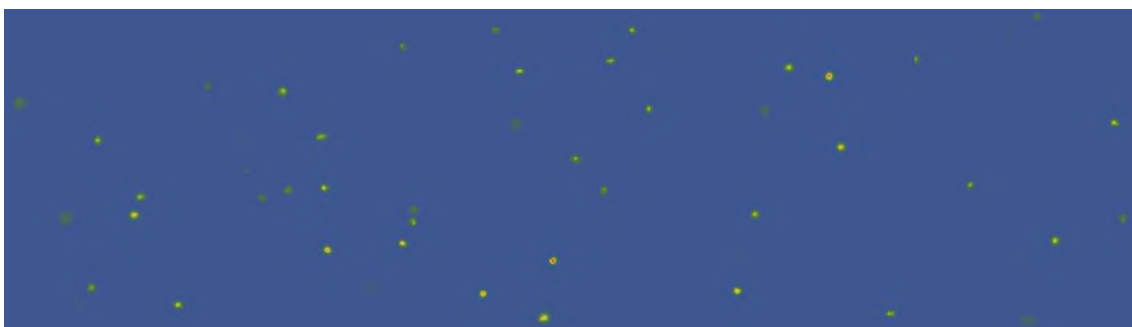


Figure 7 KPAH4PK

24h

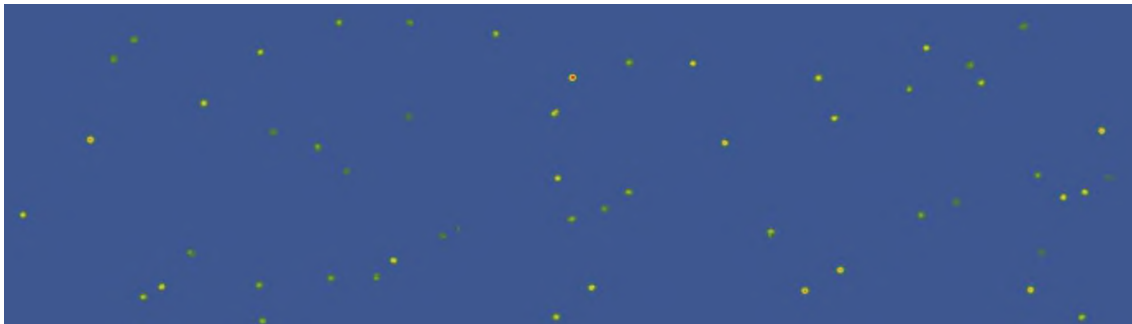


Figure 8 KPAH24K

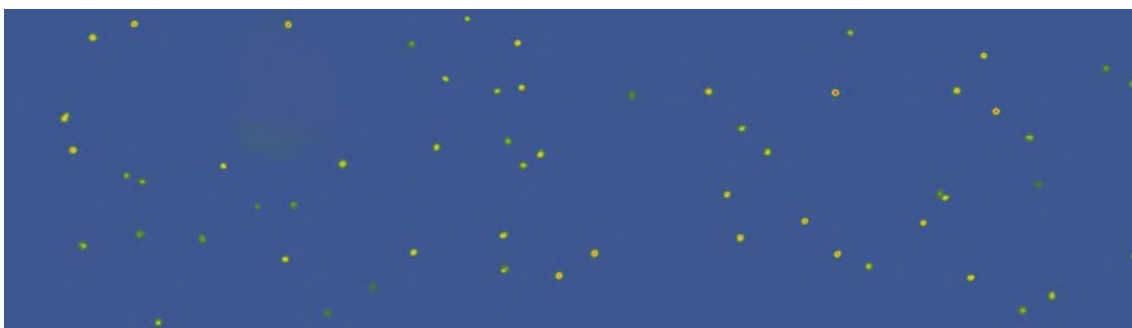


Figure 9 KPAH24L



Figure 10 KPAH24M



Figure 11 KPAH24H

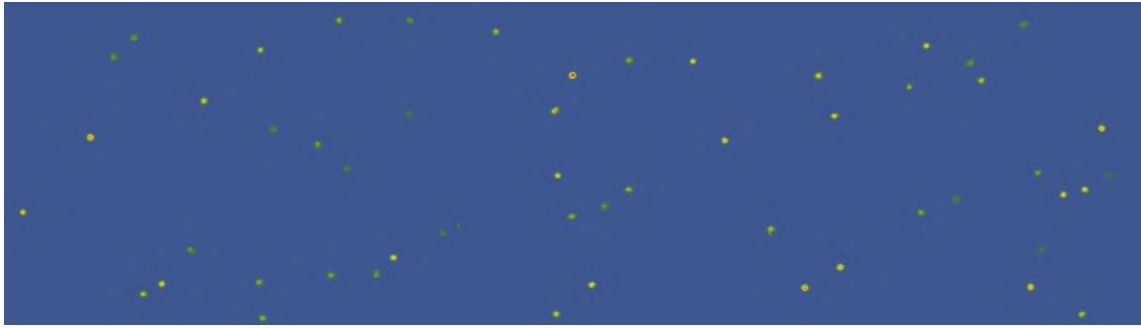


Figure 12 KPAH24PK

Examples of single and image collage 3D images

4h

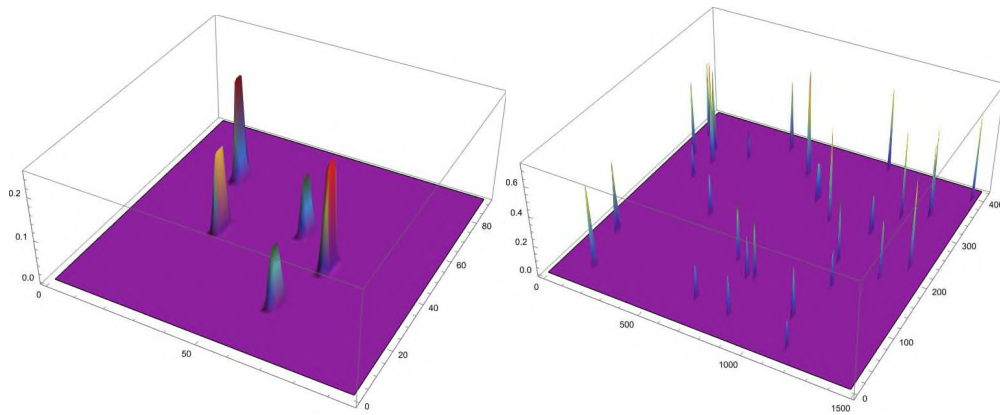


Figure 13 KPAH4K, single image (left) and image collage (right)

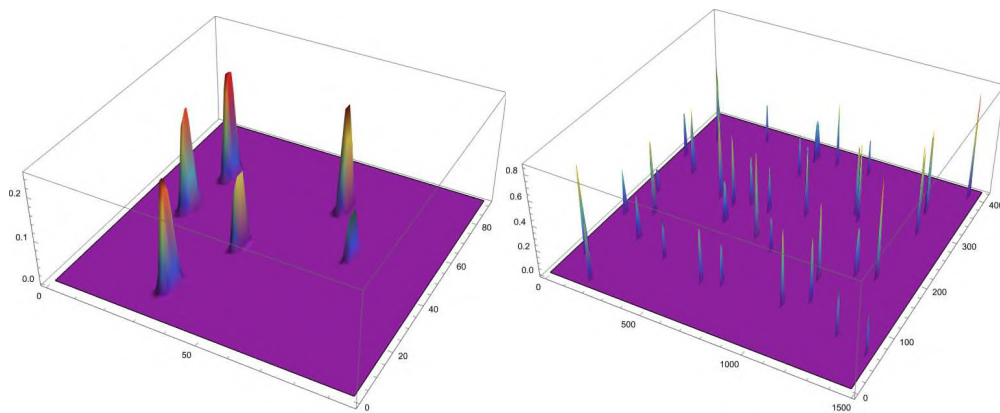


Figure 14 KPAH4L, single image (left) and image collage (right)

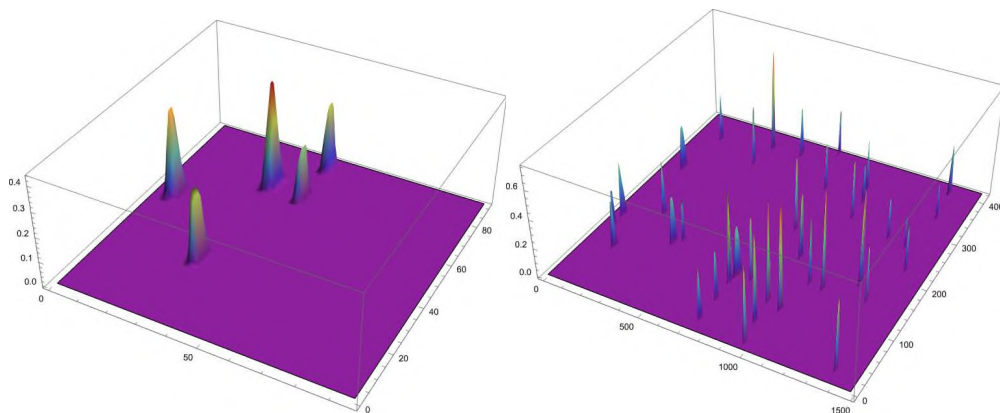


Figure 15 KPAH4M, single image (left) and image collage (right)

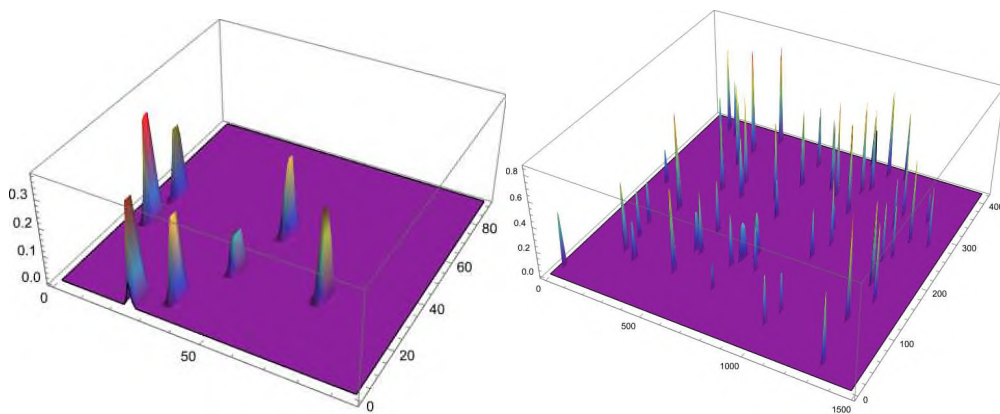


Figure 16 KPAH4H, single image (left) and image collage (right)

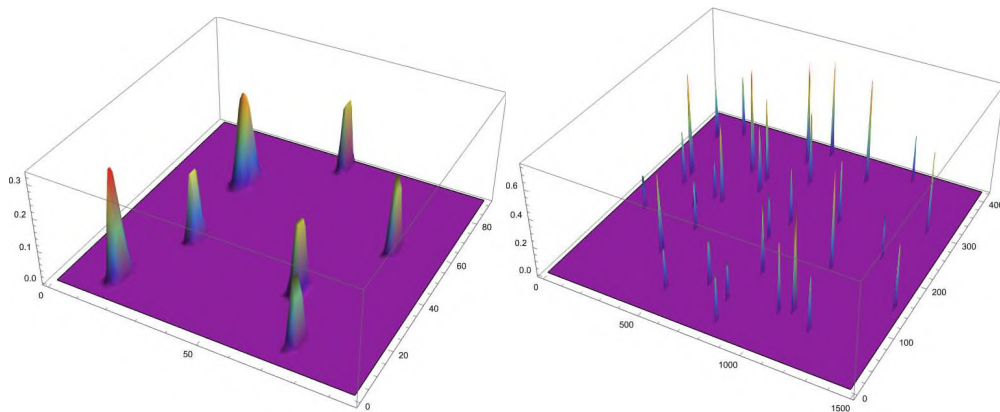


Figure 17 KPAH4PK, single image (left) and image collage (right)

24h

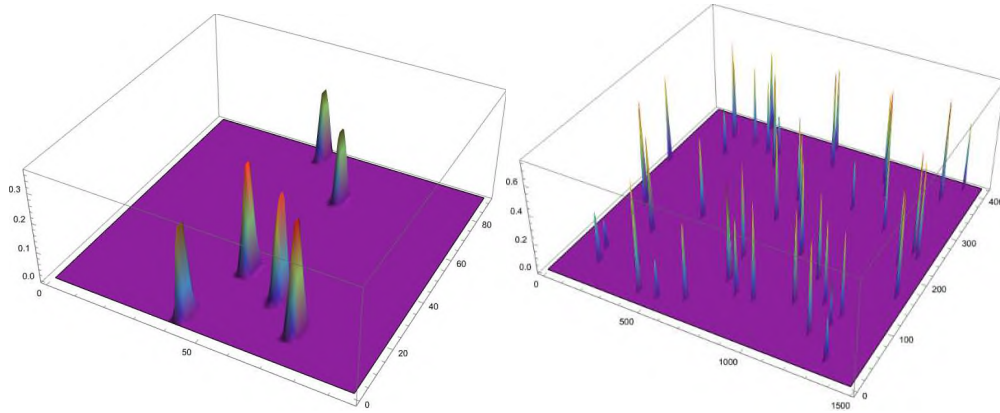


Figure 18 KPAH24K, single image (left) and image collage (right)

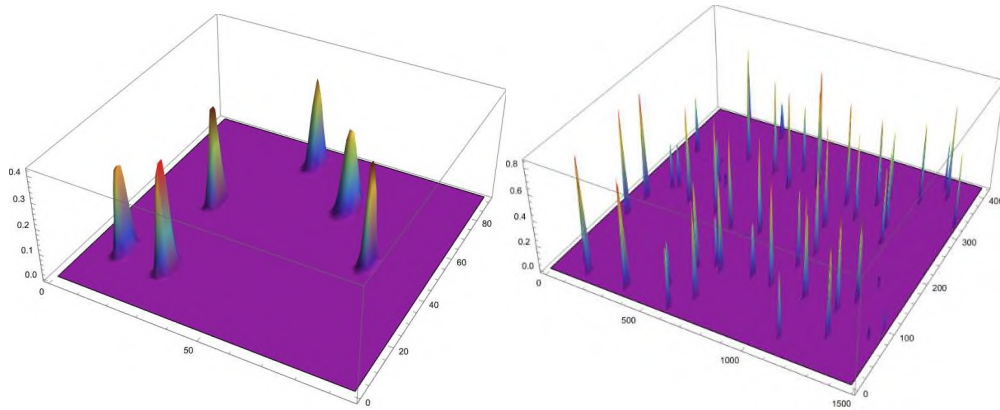


Figure 19 KPAH24L, single image (left) and image collage (right)

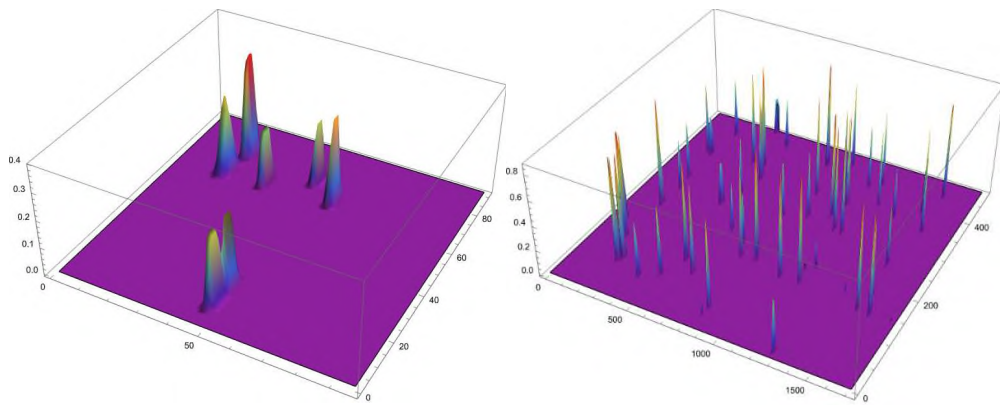


Figure 20 KPAH24M, single image (left) and image collage (right)

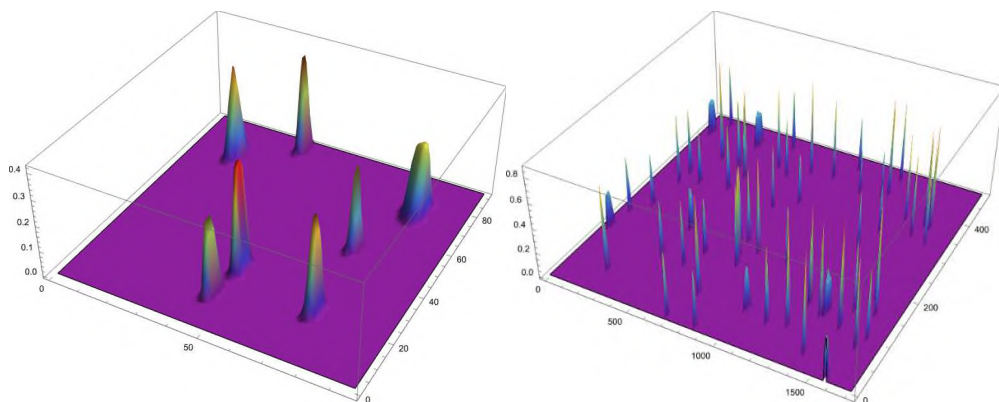


Figure 21 KPAH24H, single image (left) and image collage (right)

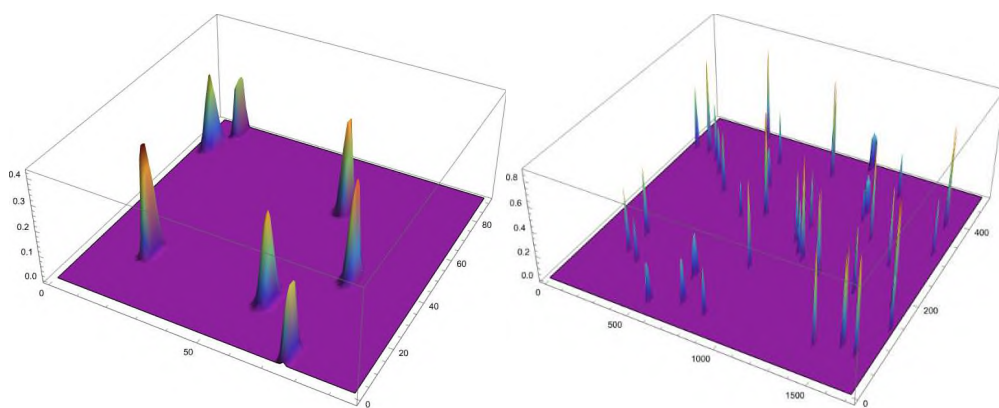


Figure 22 KPAH24PK, single image (left) and image collage (right)

Image analysis features: Scatter percentile distribution plots of

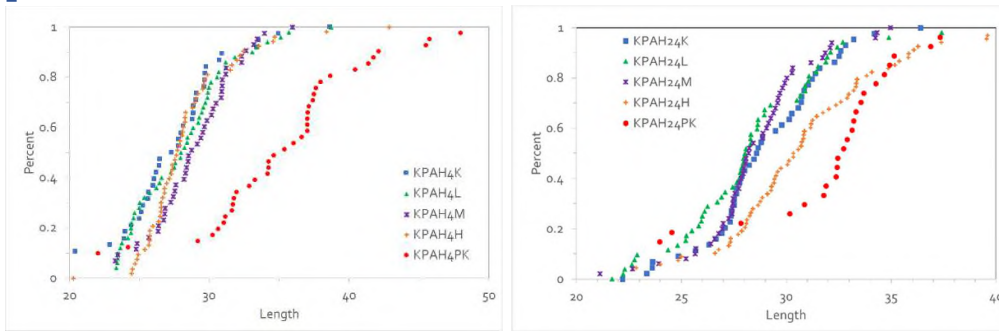


Figure 23 Percentile distribution of Lengths in Image Collages for 4h (left) and 24 h (right)

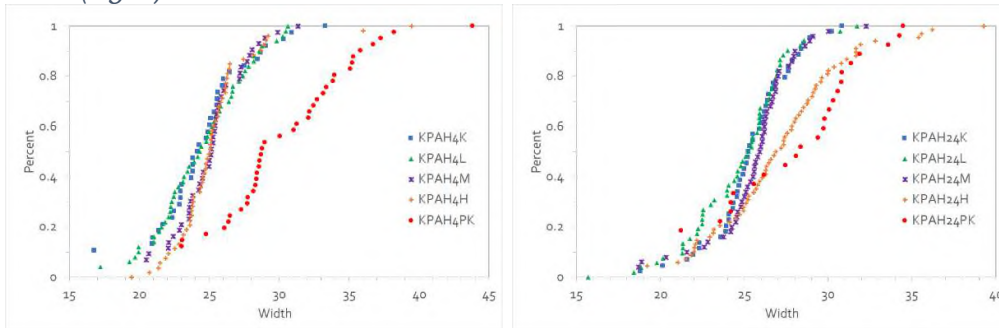


Figure 24 Percentile distribution of Widths in Image Collages for 4h (left) and 24 h (right)

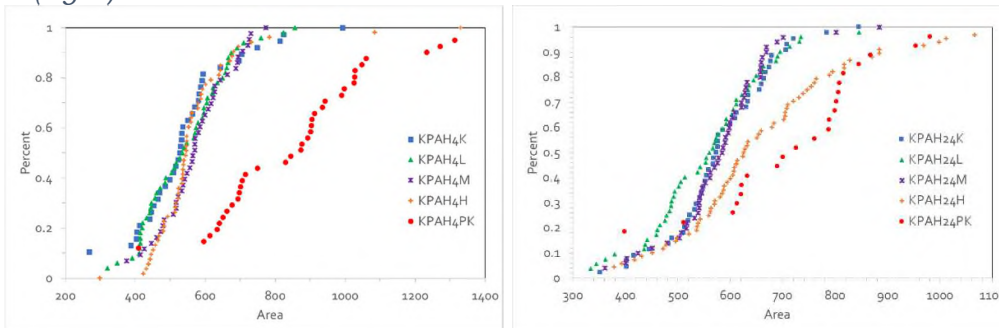


Figure 25 Percentile distribution of Areas in Image Collages for 4h (left) and 24 h (right)

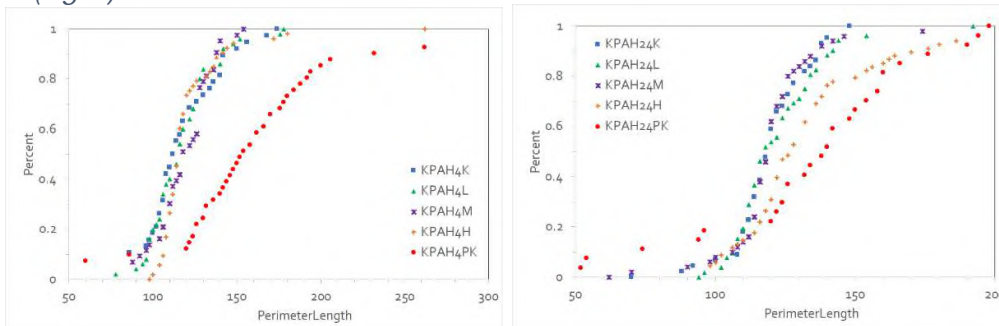


Figure 26 Percentile distribution of Perimeter Lengths in Image Collages for 4h (left) and 24 h (right)

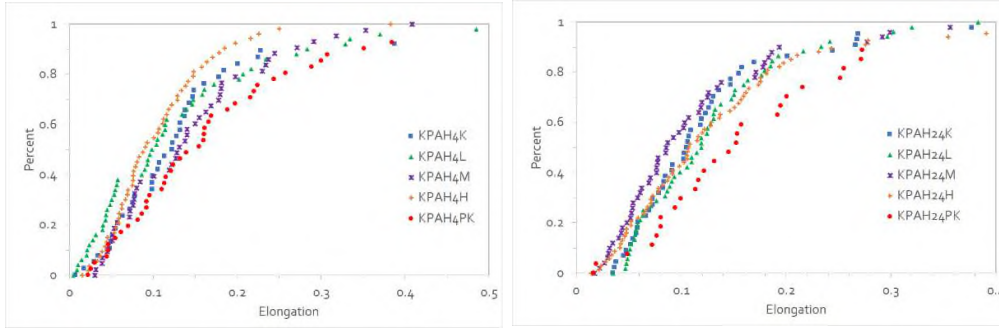


Figure 27 Percentile distribution of Elongations in Image Collages for 4h (left) and 24 h (right)

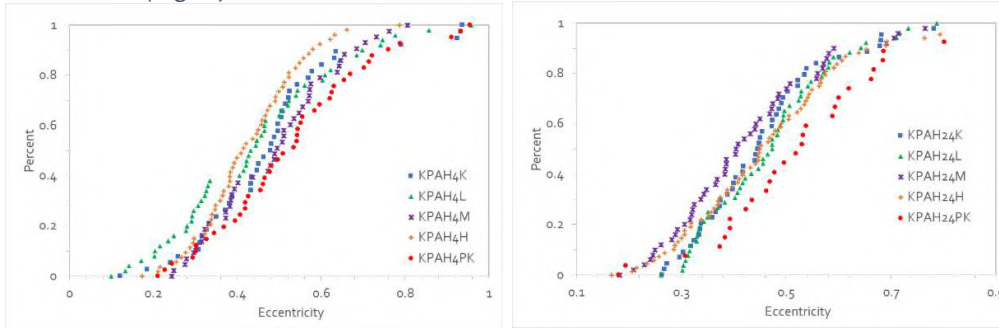


Figure 28 Percentile distribution of Eccentricities in Image Collages for 4h (left) and 24 h (right)

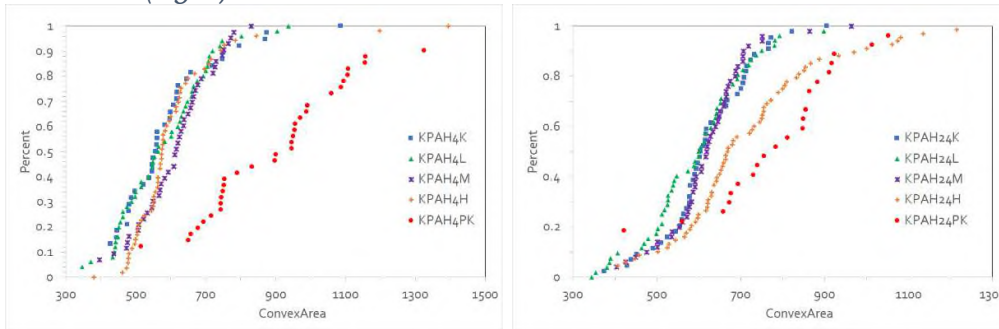


Figure 29 Percentile distribution of Convex Areas in Image Collages for 4h (left) and 24 h (right)

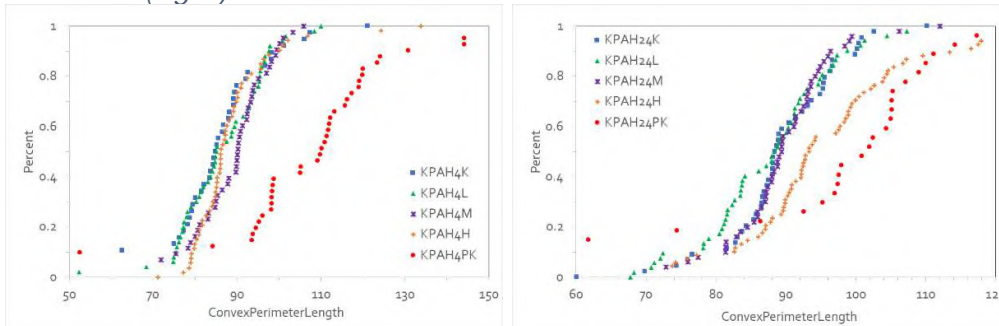


Figure 30 Percentile distribution of Convex Perimeter Lengths in Image Collages for 4h (left) and 24 h (right)

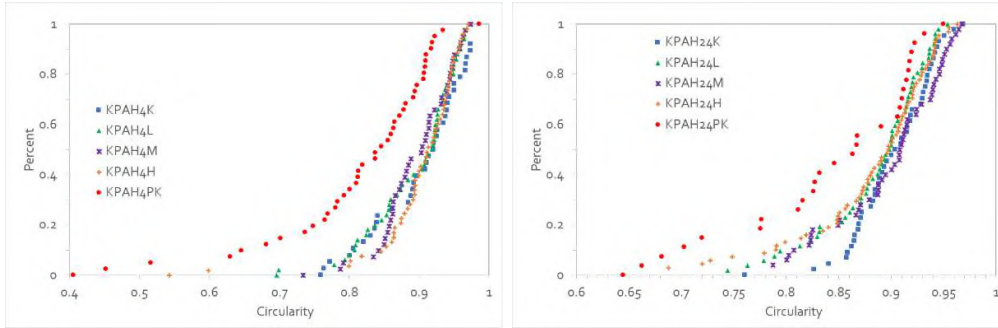


Figure 31 Percentile distribution of Circularities in Image Collages for 4h (left) and 24 h (right)

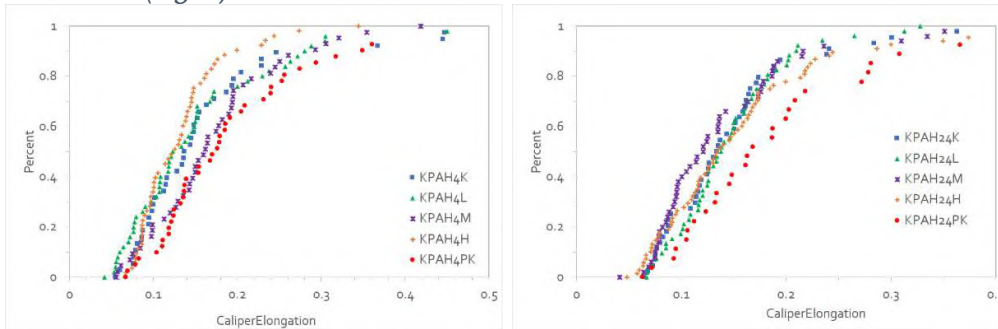


Figure 32 Percentile distribution of Caliper Elongations in Image Collages for 4h (left) and 24 h (right)

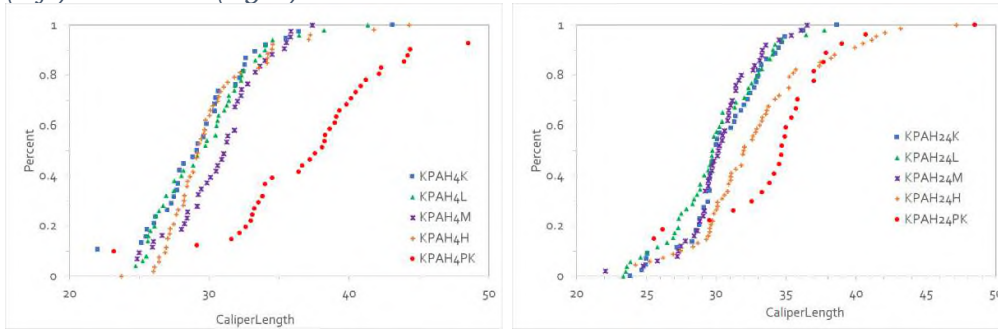


Figure 33 Percentile distribution of Caliper Lengths in Image Collages for 4h (left) and 24 h (right)

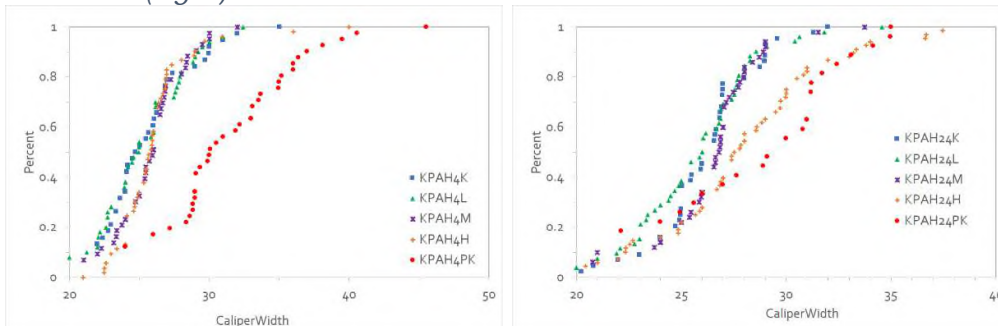


Figure 34 Percentile distribution of Caliper Widths in Image Collages for 4h (left) and 24 h (right)

Figure 35

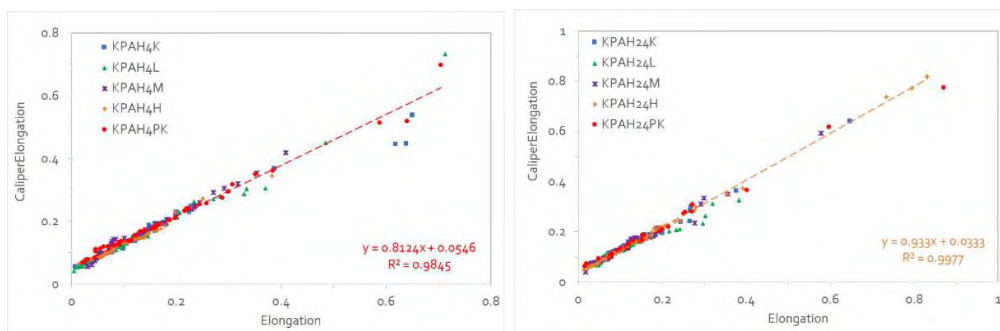


Figure 36 Correlations of Percentile distribution of Caliper Elongations vs. Elongations in Image Collages for 4h (left) and 24 h (right)

Correlation scatter plots of properties

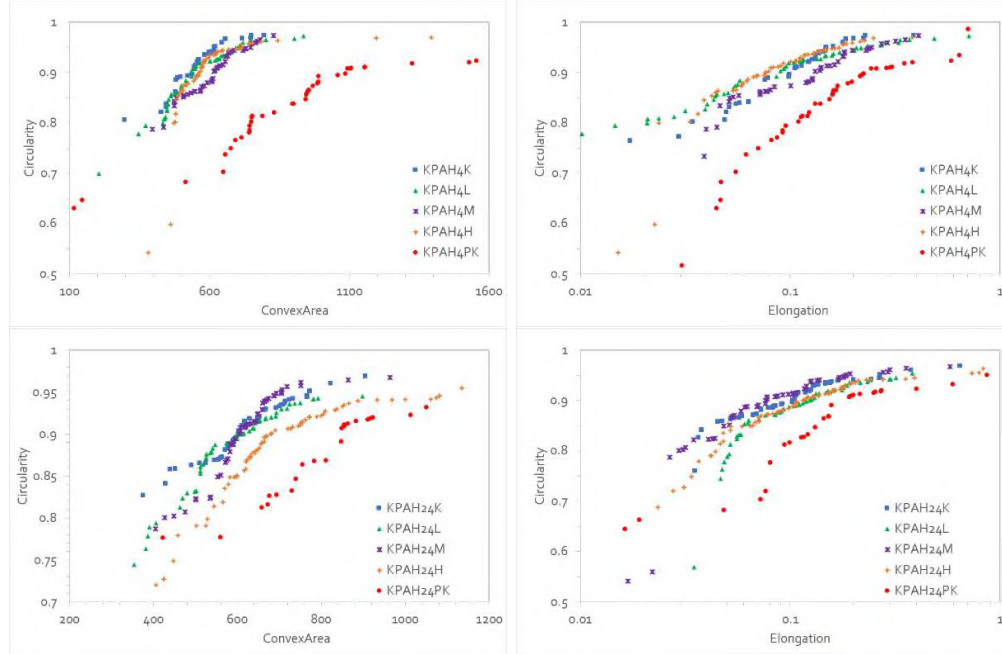


Figure 37 Correlations of Percentile distribution of Circularity vs. Convex Areas (left) and of Circularity vs. Elongations in Image Collages for 4h (upper row) and for 24 h (lower row)

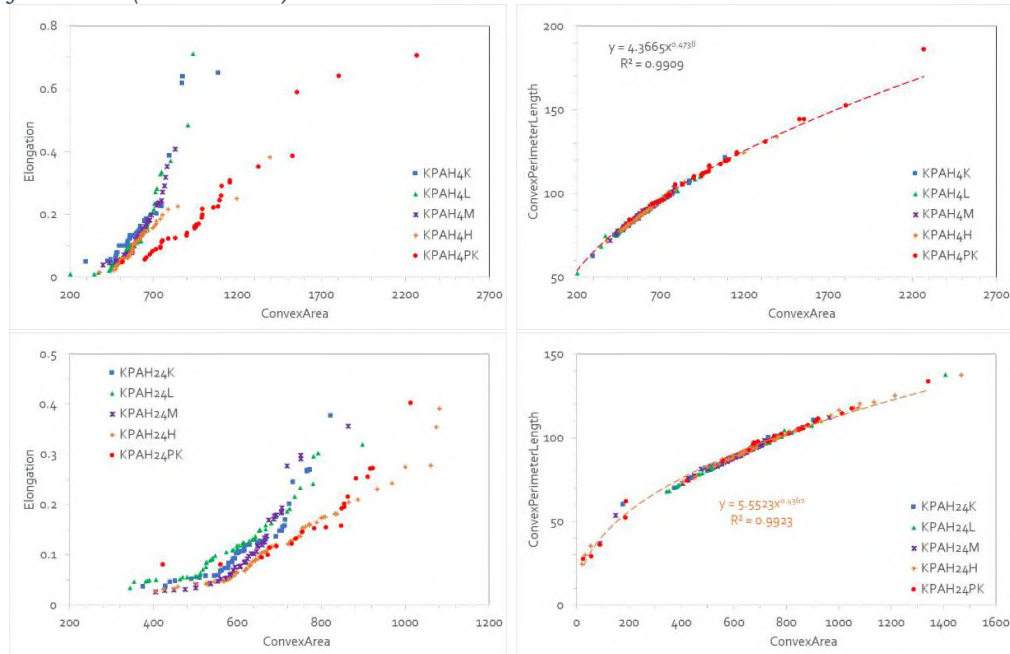


Figure 38 Correlations of Percentile distribution of Elongations vs. Convex Areas (left) and of Convex Perimeter Lengths vs. Convex Areas in Image Collages for 4h (upper row) and for 24 h (lower row)

Textural analysis: Haralick features

Image-based texture rich information in images offer possibilities to identify similarities or differences among biological samples⁹ and to decipher the toxicity profile and other biological effects.

Statistical and structural approaches categorize texture analysis methods into those that describe texture by its random, statistical variations or the arrangement of deterministic structural elements¹⁰. Statistical methods use measures like autocorrelation functions and gray-level co-occurrence matrices to extract quantitative features from pixel data, while structural methods rely on finding repeating patterns, primitives, and their relationships to describe texture. Grey Level Co-occurrence Matrices (GLCM)^{11,12,13,14,15} are one of the earliest techniques used for image texture analysis introduced by Haralick.

The texture parameters have been used in several areas of the biomedical research^{16,17,18,19}. Correlation between pixels means that there is a predictable and linear relationship between the two neighbouring pixels within the window. Entropy is a notoriously difficult term to understand; the concept comes from thermodynamics. The term is used to mean irremediable chaos or disorder. High values of ASM occur when the window is orderly. The square root of the ASM is sometimes used as a texture measure, and is called Energy. Energy is, in this context, the opposite of entropy. A high Correlation texture means high predictability of pixel relationships.

Textural analysis

Average values of GLCM parameters with standard deviations are presented in table 1. The values of entropy, angular second moment, and inverse difference moment did not significantly change after the exposures. There was no significant correlation between GLCM and fractal parameters.

Table 2 Haralick features ASM: Angular Second Moment (Homogeneity), Contrast, Correlation, IDM: Inverse Difference Moment and Entropy

	K		L		M		H		PK	
	Av.	st. dev.	Av.	st. dev.	Av.	st. dev.	Av.	st. dev.	Av.	st. dev.
4h										
ASM	0.707	0.187	0.611	0.234	0.518	0.200	0.465	0.147	0.571	0.188
Contrast	12.33	6.37	13.67	7.79	16.48	7.43	13.39	3.41	13.04	3.40
Correlation	0.026	0.015	0.016	0.006	0.016	0.006	0.016	0.009	0.019	0.007
IDM	0.904	0.065	0.890	0.076	0.865	0.076	0.890	0.036	0.895	0.044
Entropy	0.706	0.357	0.893	0.440	1.043	0.310	1.194	0.293	1.004	0.346
24h										
ASM	0.928	0.024	0.862	0.193	0.640	0.266	0.891	0.059	0.561	0.249
Contrast	4.02	0.74	5.08	2.92	14.15	9.21	4.90	2.39	16.58	9.76
Correlation	0.029	0.006	0.039	0.010	0.025	0.012	0.028	0.012	0.024	0.012
IDM	0.981	0.010	0.969	0.032	0.851	0.119	0.974	0.024	0.841	0.115
Entropy	0.259	0.065	0.385	0.385	0.811	0.510	0.348	0.142	0.966	0.440

Entropy and Contrast at different levels of exposure are going up for 4h and for 24 h (except for the H-exposure level at 24h) (Figure 39, left and Figure 40, left). Angular Second Moment, ASM (Homogeneity), and Inverse Difference

Moment, IDM, are going down for 4h and for 24 h (except for the H-exposure level at 24h) (Figure 39, right and Figure 40, right). is also going down for 4h and for 24 h (except for the H-exposure level at 24h).

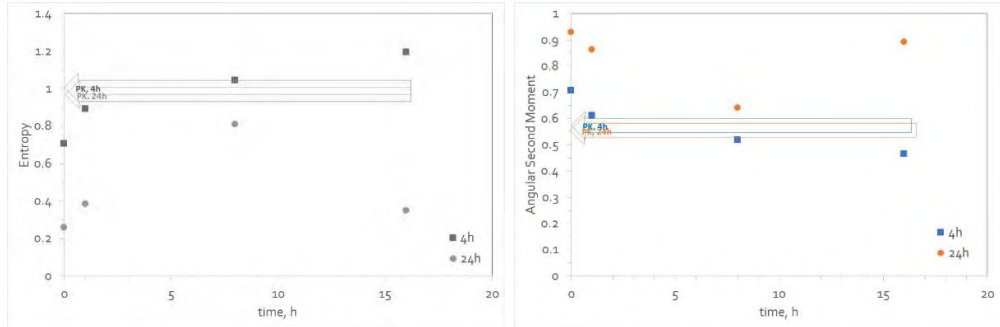


Figure 39 Haralick features Entropy (left) and ASM: Angular Second Moment (Homogeneity) (right) for different time of exposure.

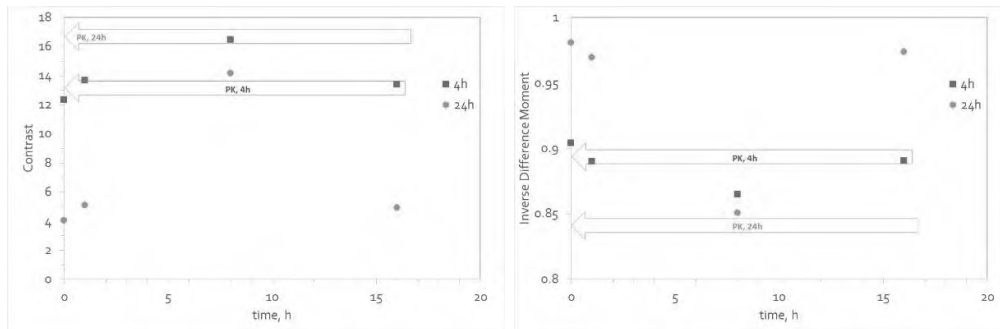


Figure 40 Haralick features Contrast (left) and IDM: Inverse Difference Moment (right) for different time of exposure.

Fractal analysis

For each Comet assay structure, values of fractal dimension, lacunarity, circularity, as well as parameters of grey level co-occurrence matrix (GLCM) texture, were determined.

Fractal dimension is the most important indicator of structural complexity of the analysed object. In biomedical research it has so far been successfully applied to measure structural properties of various tissues and organs, such as liver, lungs, brain, and kidney^{20,21,22}. In cell biology fractal dimension of nuclear structure was a sensitive parameter in evaluation of early stages of programmed cell death^{23,24}. Fractal dimension measures surface complexity, with higher values showing details at increasingly smaller scales²⁵.

Monofractals: Characterizing space

The goal is to define a numerical index for the two-dimensional images, using a fractal dimension, D , that provides a statistical measure on how details in the pattern change with the scale at which it is measured. To help with this study, the binary black and white contrast images were used to determine the fractal dimensions. The fractal dimension (D) is determined by the equation:

$$D = \lim_{s \rightarrow 0} \left[\frac{\log N(\varepsilon)}{\log(\frac{1}{\varepsilon})} \right]$$

where ε is the length of the box which creates a mesh covering the image, and $N(s)$ is the minimal number of boxes required to cover the entire pattern. This method counts the number of boxes (N) of length s required to cover the object being measured. In this method, a grid of boxes is superimposed over the binary image, and the number of nonempty boxes (N) for different scale (ε), that at least partially cover the object is determined. Fractal dimension (D) was calculated based on the slope of the regression line from the logarithmic relationship between the box count and the scale:

Range Fractal indices of separate subImages

Control K

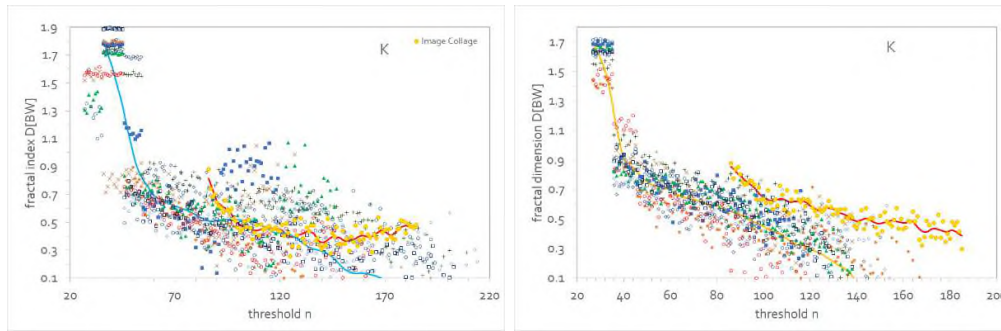


Figure 41 Fractal indices $D[BW]$ of separate subimages for 4h (left) and 24 h (right); Image collage: ● yellow, red line

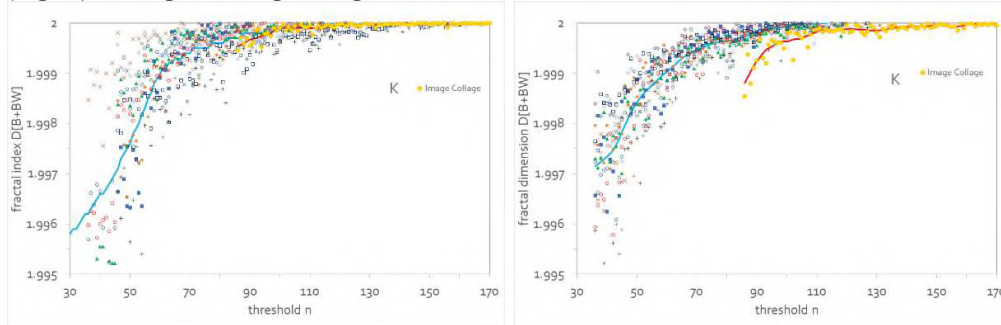


Figure 42 Fractal indices $D[B+BW]$ of separate subimages for 4h (left) and 24 h (right); Image collage: ● yellow, red line

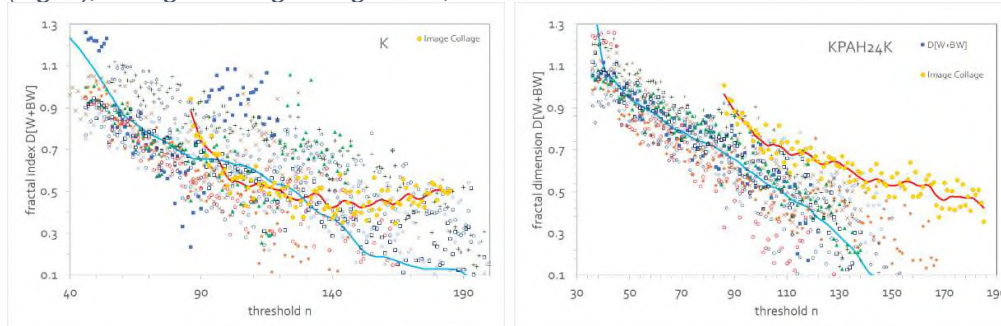


Figure 43 Fractal indices $D[W+BW]$ of separate subimages for 4h (left) and 24 h (right); Image collage: ● yellow, red line

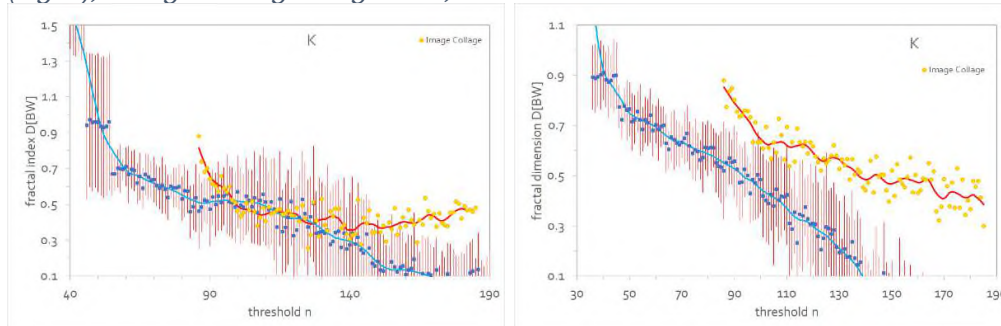


Figure 44 Average Fractal indices $D[BW]$ (blue) for 4h (left) and 24 h (right) (standard deviations (red)); Image collage: ● yellow, red line

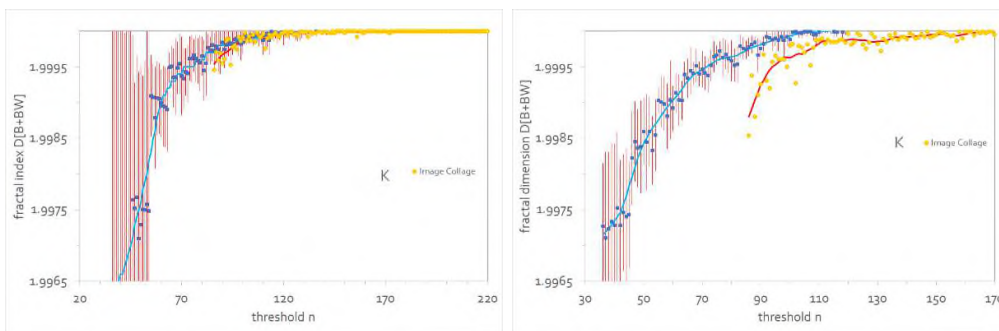


Figure 45 Average Fractal indices $D[B+BW]$ (blue) for 4h (left) and 24 h (right) (standard deviations (red); Image collage: ● yellow, red line)

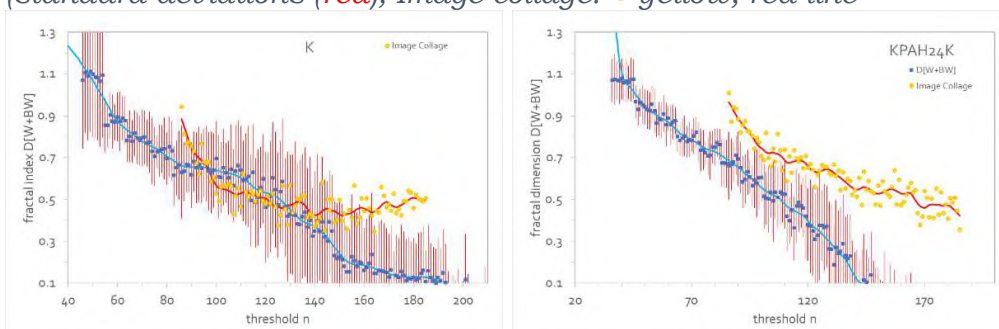


Figure 46 Average Fractal indices $D[W+BW]$ (blue) for 4h (left) and 24 h (right) (standard deviations (red); Image collage: ● yellow, red line)

Exposure L

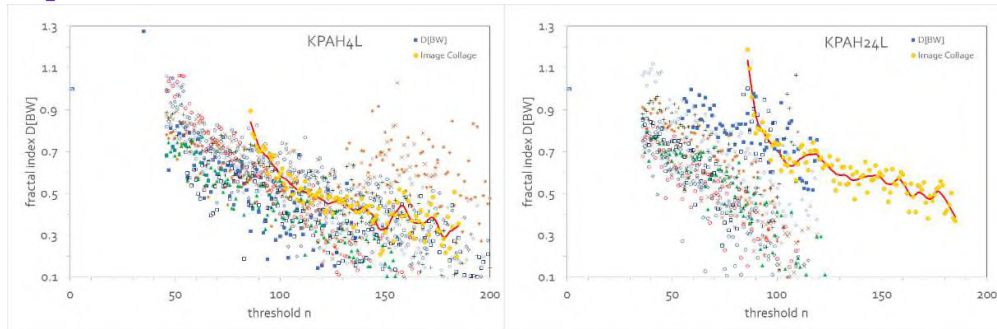


Figure 47 Fractal indices $D[BW]$ of separate subimages for 4h (left) and 24 h (right); Image collage: ● yellow, red line

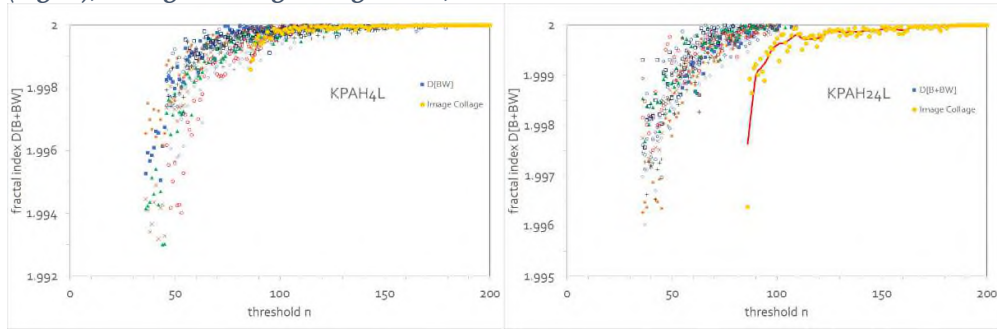


Figure 48 Fractal indices $D[B+BW]$ of separate subimages for 4h (left) and 24 h (right); Image collage: ● yellow, red line

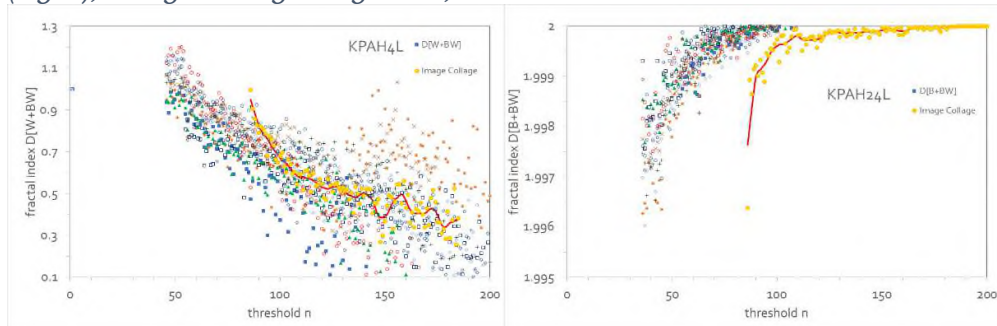


Figure 49 Fractal indices $D[W+BW]$ of separate subimages for 4h (left) and 24 h (right); Image collage: ● yellow, red line

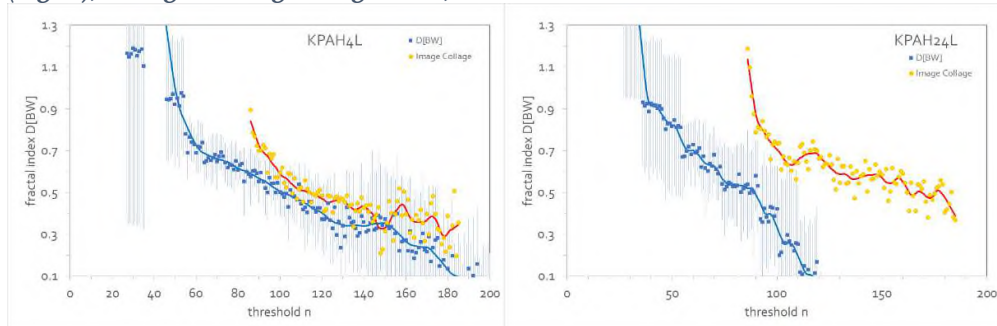


Figure 50 Average Fractal indices $D[BW]$ (blue) for 4h (left) and 24 h (right) (standard deviations (red); Image collage: ● yellow, red line

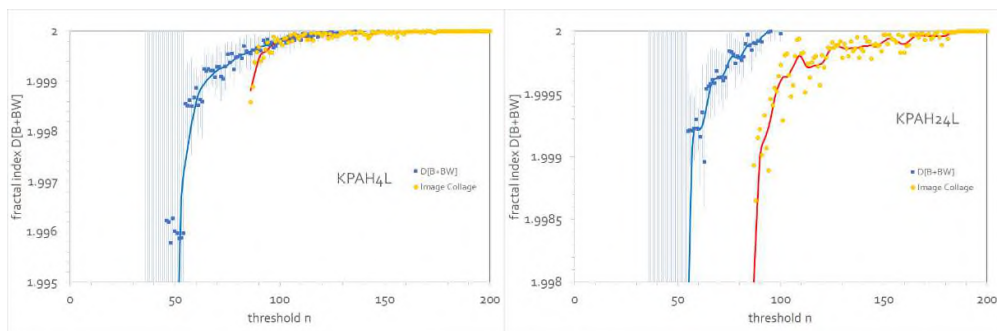


Figure 51 Average Fractal indices $D[B+BW]$ (blue) for 4h (left) and 24 h (right) (standard deviations (red); Image collage: ● yellow, red line)

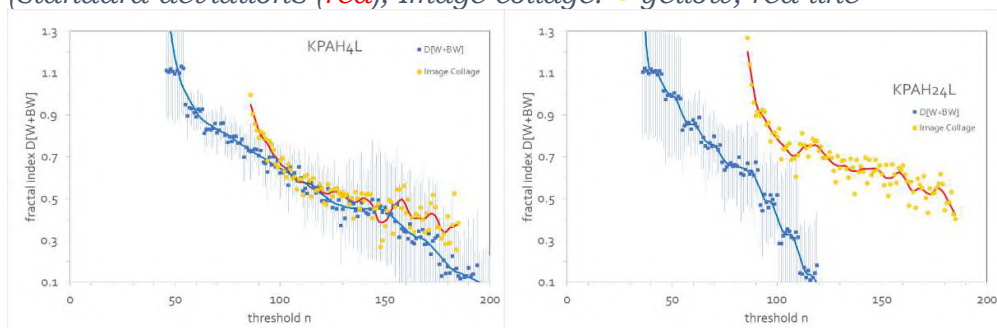


Figure 52 Average Fractal indices $D[W+BW]$ (blue) for 4h (left) and 24 h (right) (standard deviations (red); Image collage: ● yellow, red line)

Exposure M

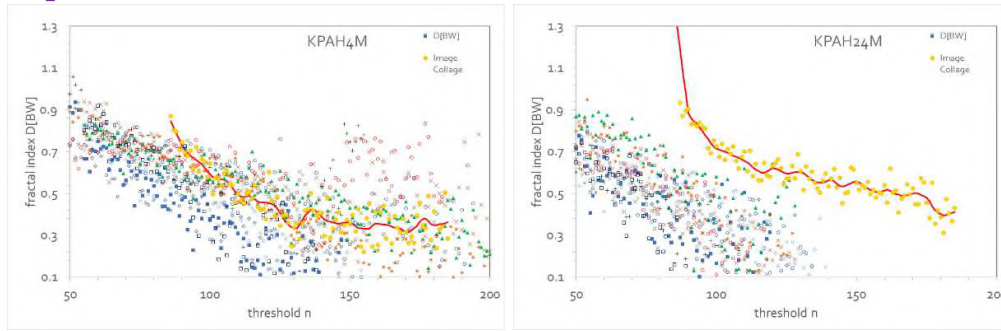


Figure 53 Fractal indices $D[BW]$ of separate subimages for 4h (left) and 24 h (right); Image collage: ● yellow, red line

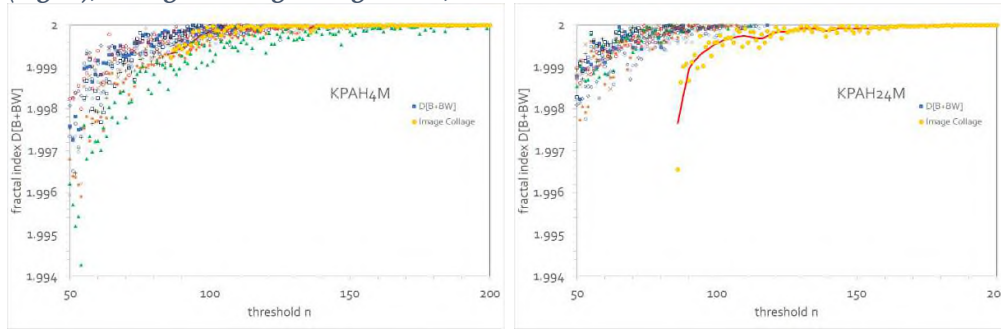


Figure 54 Fractal indices $D[B+BW]$ of separate subimages for 4h (left) and 24 h (right); Image collage: ● yellow, red line

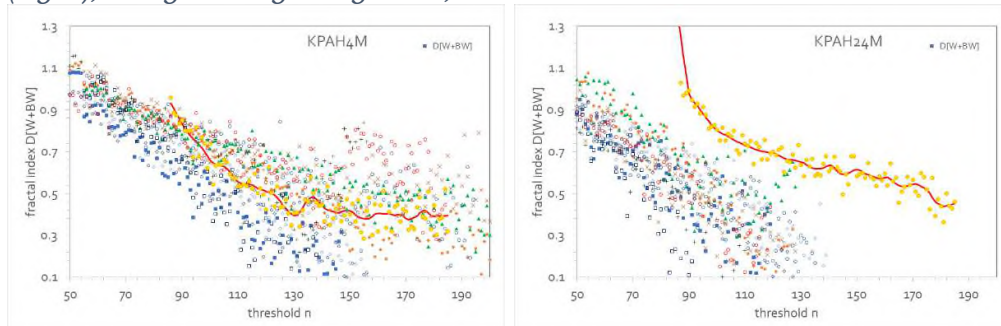


Figure 55 Fractal indices $D[W+BW]$ of separate subimages for 4h (left) and 24 h (right); Image collage: ● yellow, red line

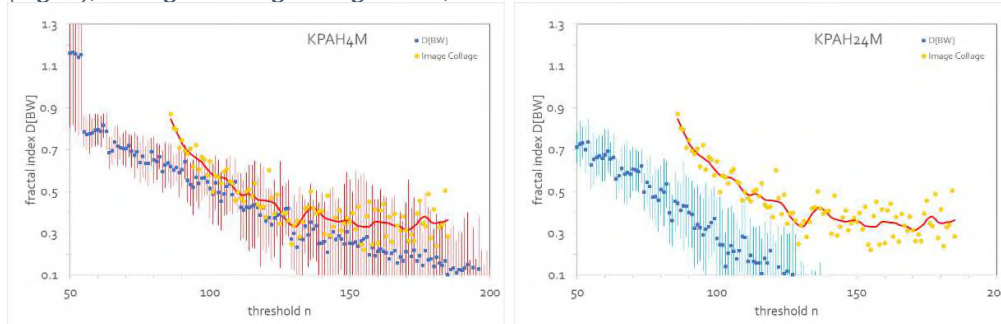


Figure 56 Average Fractal indices $D[BW]$ (blue) for 4h (left) and 24 h (right) (standard deviations (red)); Image collage: ● yellow, red line

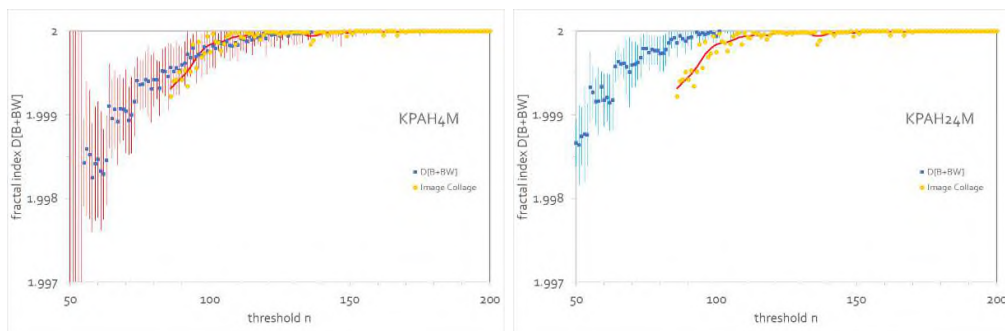


Figure 57 Average Fractal indices $D[B+BW]$ (blue) for 4h (left) and 24 h (right) (standard deviations (red); Image collage: ● yellow, red line

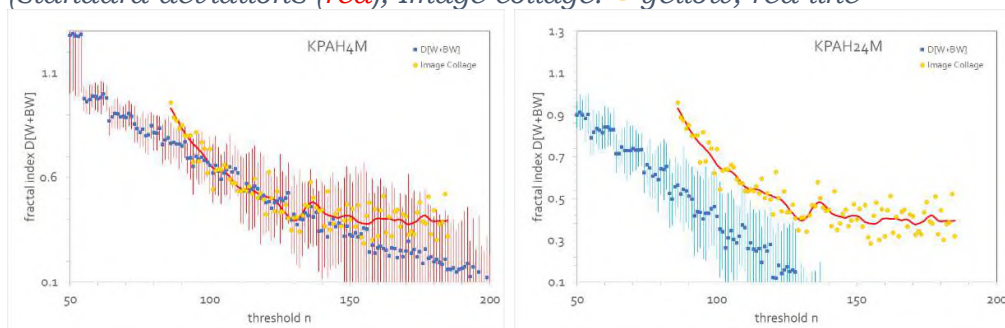


Figure 58 Average Fractal indices $D[W+BW]$ (blue) for 4h (left) and 24 h (right) (standard deviations (red); Image collage: ● yellow, red line

Exposure H

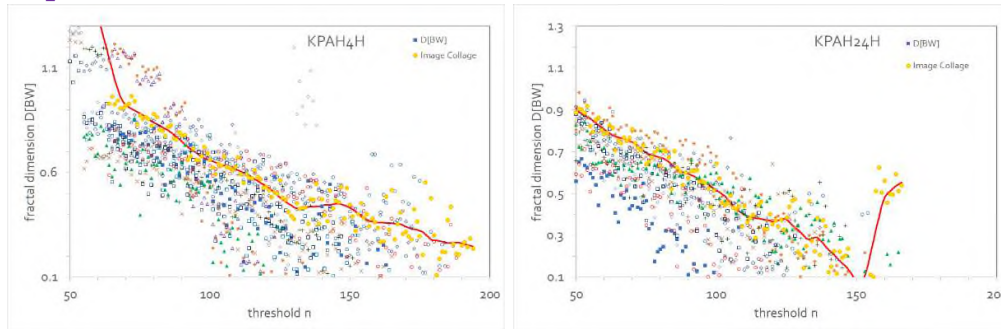


Figure 59 Fractal indices $D[BW]$ of separate subimages for 4h (left) and 24 h (right); Image collage: ● yellow, red line

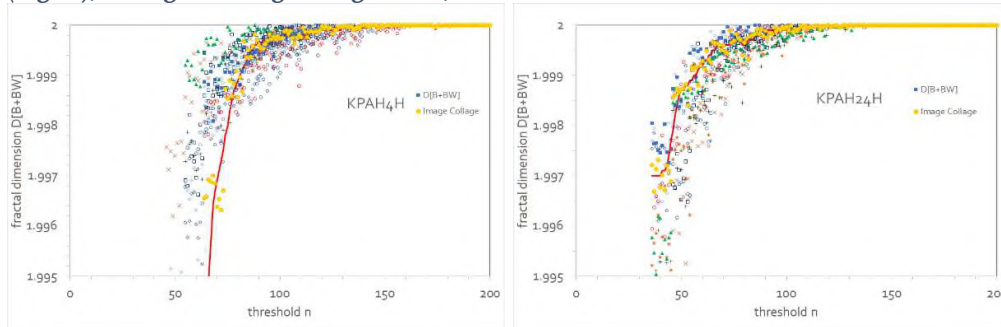


Figure 60 Fractal indices $D[B+BW]$ of separate subimages for 4h (left) and 24 h (right); Image collage: ● yellow, red line

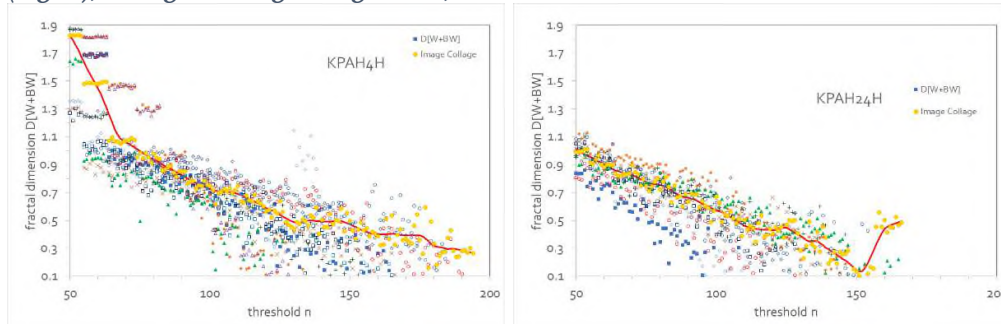


Figure 61 Fractal indices $D[W+BW]$ of separate subimages for 4h (left) and 24 h (right); Image collage: ● yellow, red line

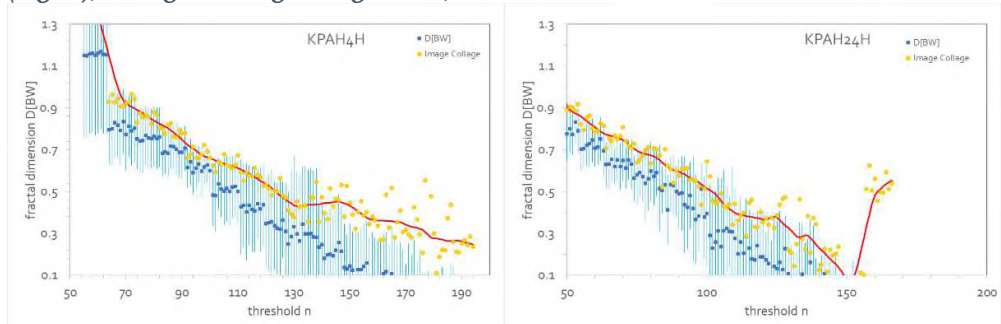


Figure 62 Average Fractal indices $D[BW]$ (blue) for 4h (left) and 24 h (right) (standard deviations (red)); Image collage: ● yellow, red line

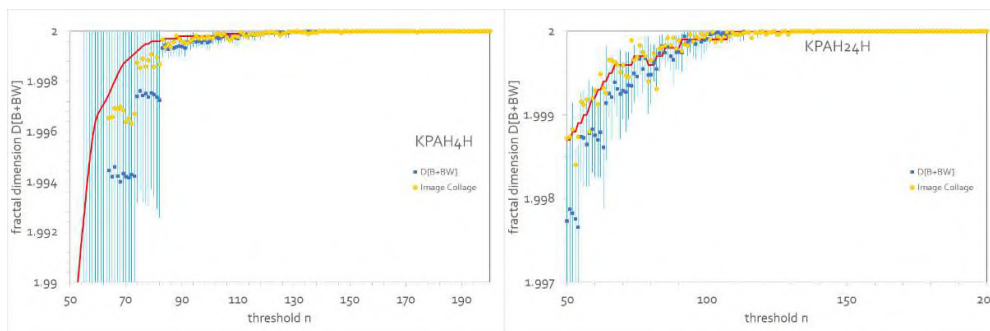


Figure 63 Average Fractal indices $D[B+BW]$ (blue) for 4h (left) and 24 h (right) (standard deviations (red); Image collage: ● yellow, red line

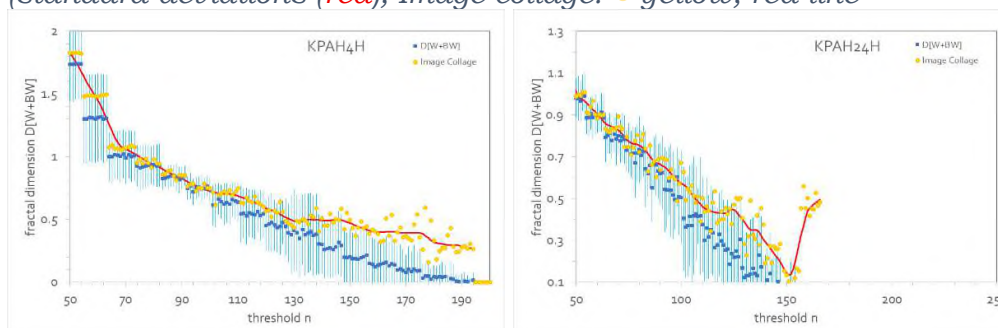


Figure 64 Average Fractal indices $D[W+BW]$ (blue) for 4h (left) and 24 h (right) (standard deviations (red); Image collage: ● yellow, red line

Positive control PK

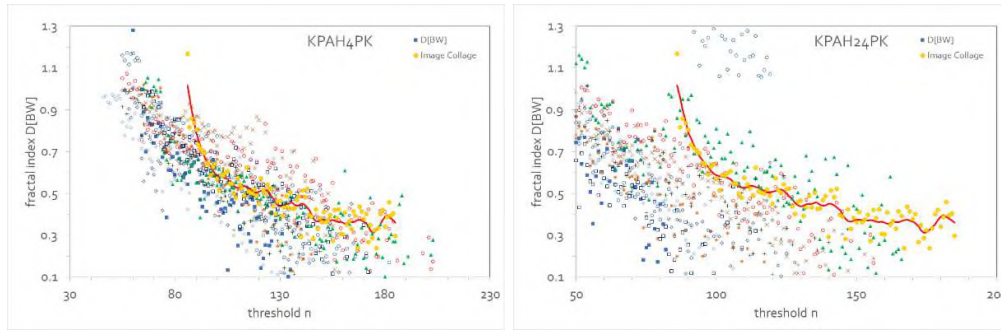


Figure 65 Fractal indices $D[BW]$ of separate subimages for 4h (left) and 24 h (right); Image collage: ● yellow, red line

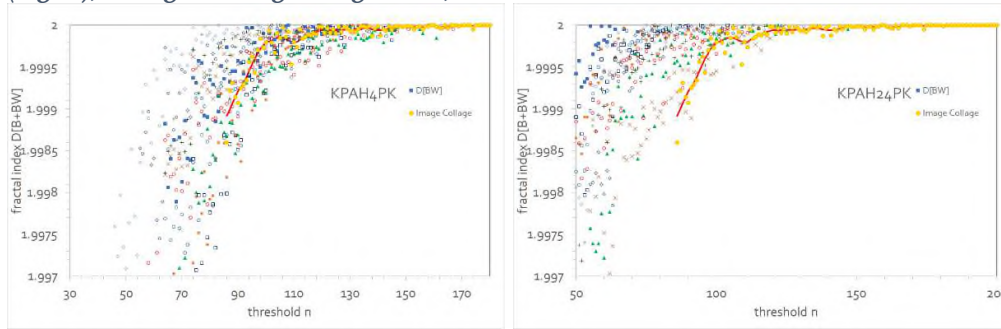


Figure 66 Fractal indices $D[B+BW]$ of separate subimages for 4h (left) and 24 h (right); Image collage: ● yellow, red line

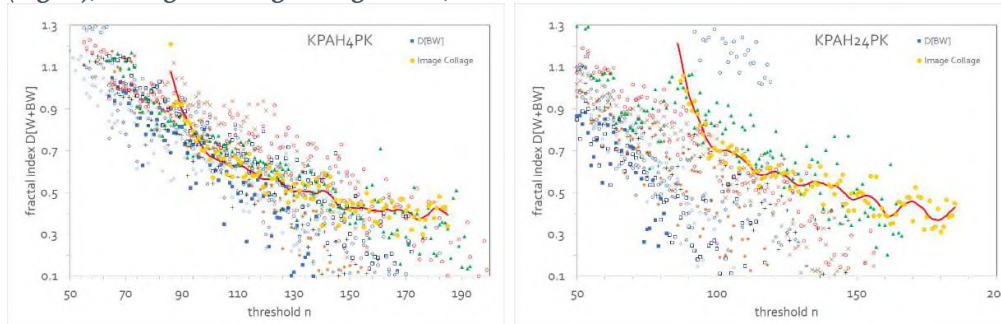


Figure 67 Fractal indices $D[W+BW]$ of separate subimages for 4h (left) and 24 h (right); Image collage: ● yellow, red line

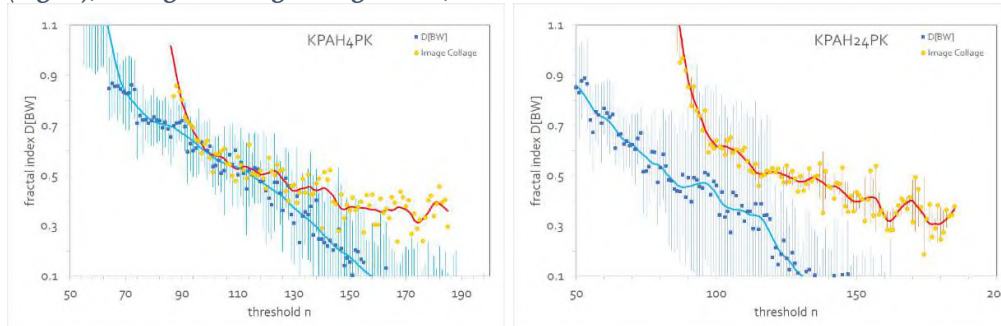


Figure 68 Average Fractal indices $D[BW]$ (blue) for 4h (left) and 24 h (right) (standard deviations (red)); Image collage: ● yellow, red line

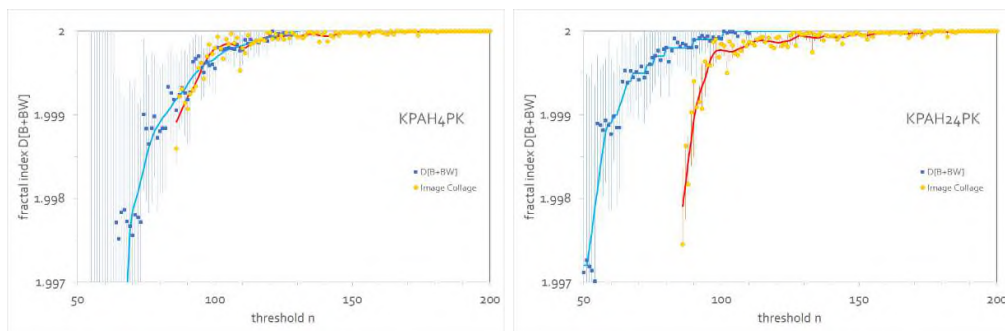


Figure 69 Average Fractal indices $D[B+BW]$ (blue) for 4h (left) and 24 h (right) (standard deviations (red); Image collage: ● yellow, red line

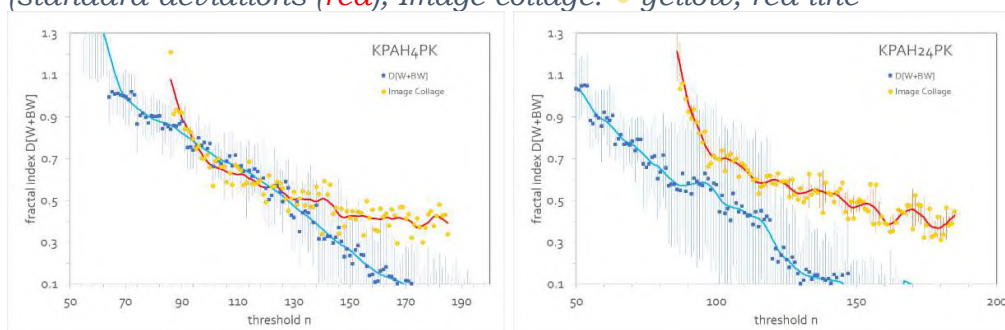


Figure 70 Average Fractal indices $D[W+BW]$ (blue) for 4h (left) and 24 h (right) (standard deviations (red); Image collage: ● yellow, red line

Range Fractal indices of Image Collage for 4h and 24 h

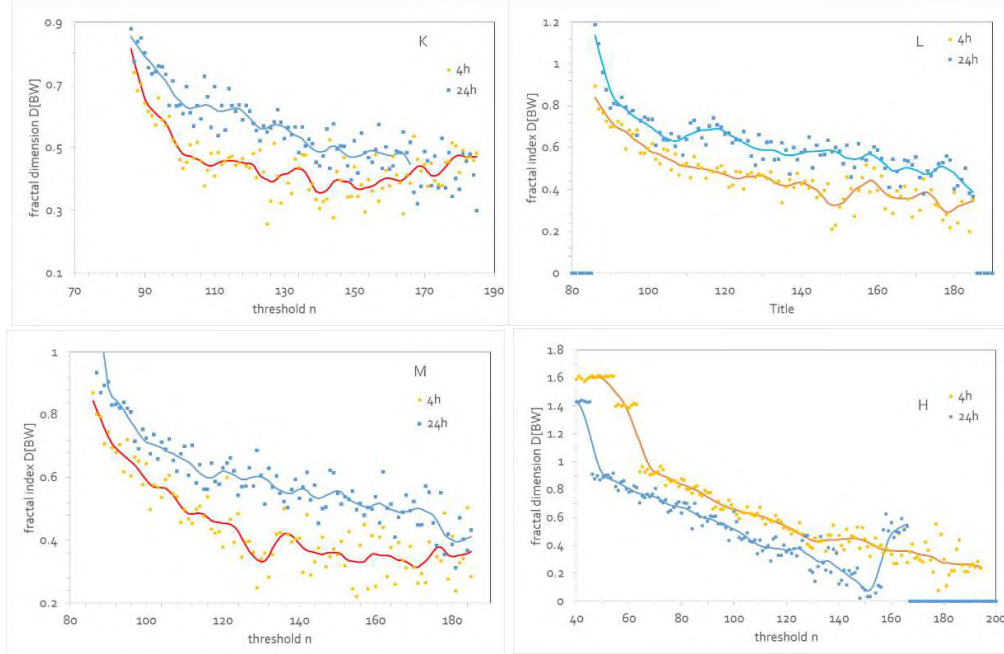


Figure 71 Fractal indices $D[BW]$ of Image Collage for 4h (yellow) and 24 h (blue)

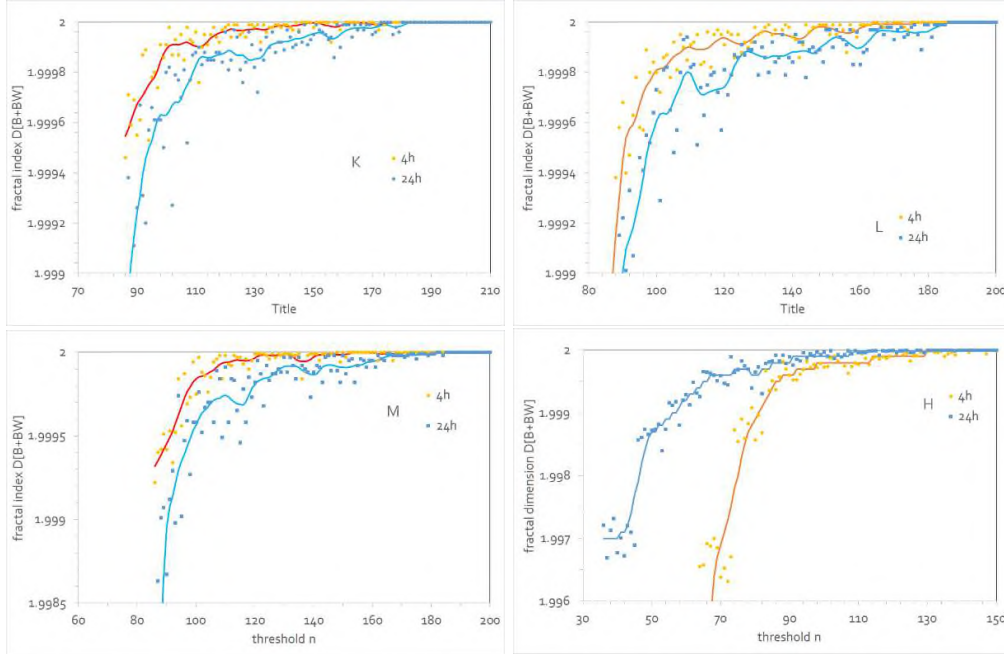


Figure 72 Fractal indices $D[B+BW]$ of Image Collage for 4h (yellow) and 24 h (blue)

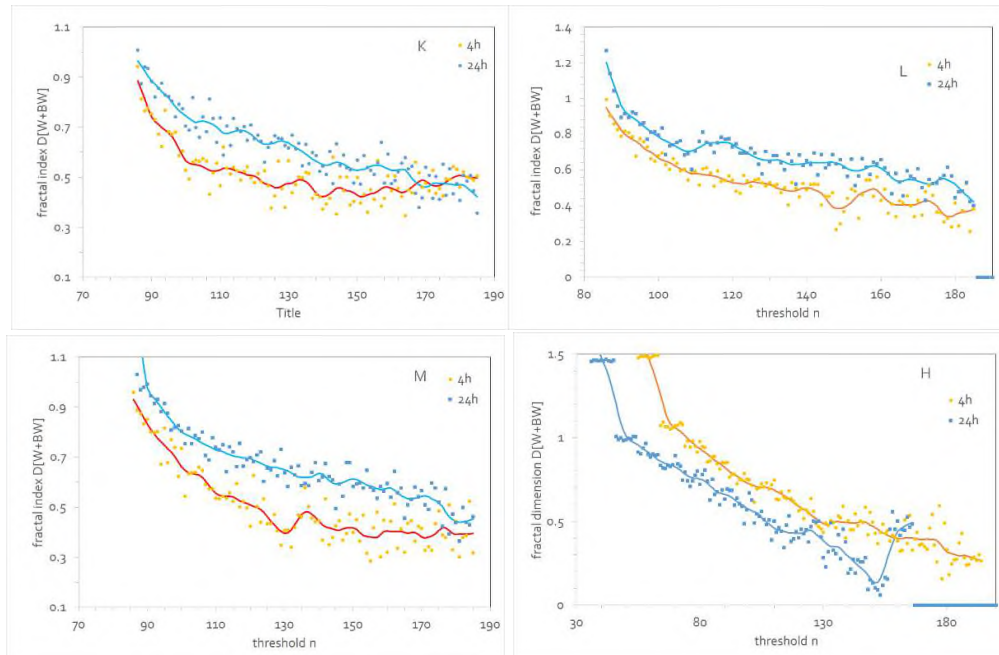


Figure 73 Fractal indices $D[W+BW]$ of Image Collage for 4h (yellow) and 24 h (blue)

Fractal indices for different exposures, L , M , H

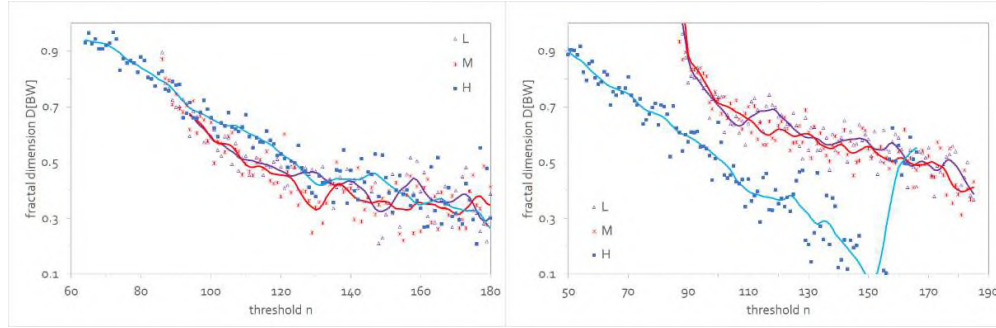


Figure 74 Fractal indices $D[BW]$ of Image Collage for 4h (left) and 24 h(right); Δ L , \times M , \bullet H

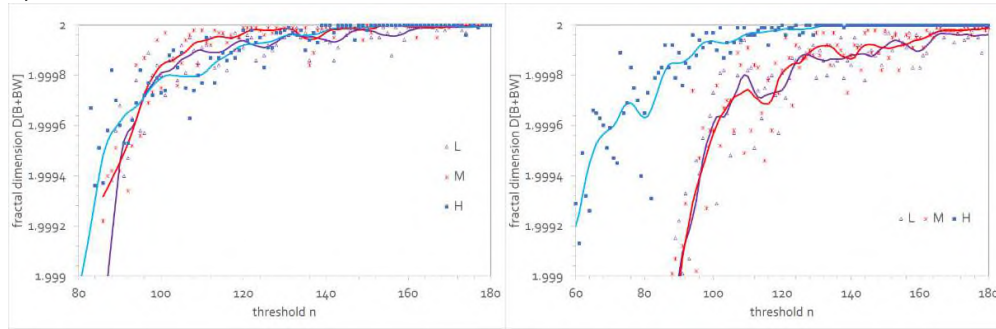


Figure 75 Fractal indices $D[B+BW]$ of Image Collage for 4h (left) and 24 h(right); Δ L , \times M , \bullet H

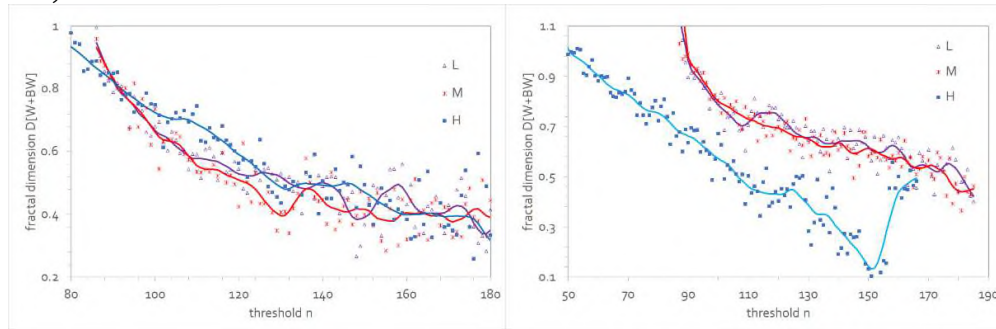


Figure 76 Fractal indices $D[W+BW]$ of Image Collage for 4h (left) and 24 h(right); Δ L , \times M , \bullet H

Fractal indices for control (K), positive control (PK) and different exposures, L, M, H

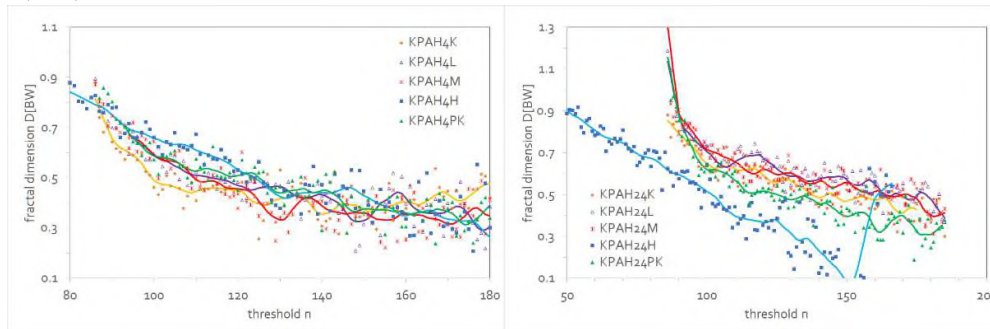


Figure 77 Fractal indices $D[BW]$ of Image Collage for 4h (left) and 24 h(right); ♦ control K, △ L, ✕ M, ● H, ▲ positive control PK

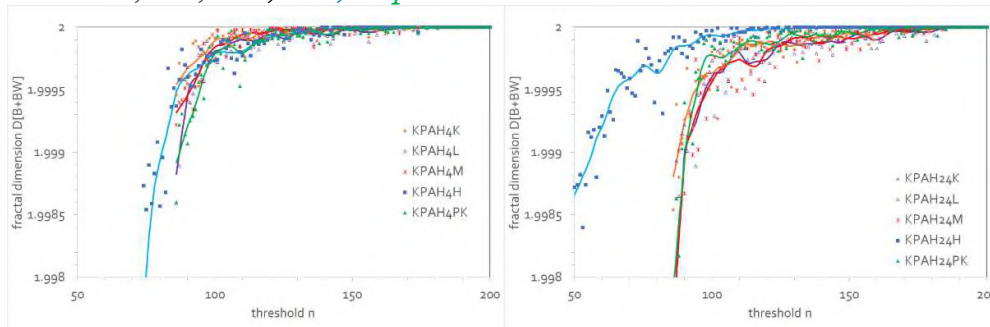


Figure 78 Fractal indices $D[B+BW]$ of Image Collage for 4h (left) and 24 h(right); ♦ control K, △ L, ✕ M, ● H, ▲ positive control PK

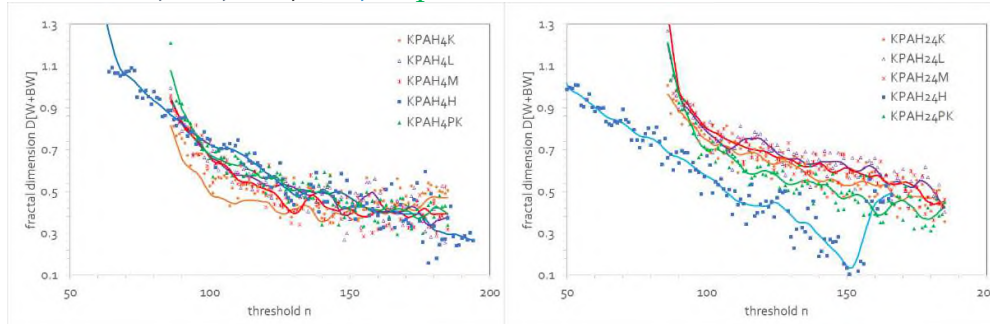


Figure 79 Fractal indices $D[W+BW]$ of Image Collage for 4h (left) and 24 h(right); ♦ control K, △ L, ✕ M, ● H, ▲ positive control PK

Multi-scale Focal fractal dimension

Multi-scale Focal fractal dimension is a local measurement of fractal dimension²⁶ that analyses the roughness and complexity of a surface or signal within a specific, moving window of a defined size using the variogram method²⁷. Unlike a standard fractal dimension that provides a single global value, Focal fractal dimension generates a raster map of fractal values, allowing for the identification of areas with varying degrees of self-similarity and irregularity across the landscape or data. For example, a landscape with smooth features like flat lakes would have low focal fractal dimension values, while rough, complex areas would have higher values.

Focal fractal dimension uses a moving window of a specified size (e.g., 11x11 pixels) to analyse part of the data. Within each window, a fractal dimension is calculated using a standard method, such as the box-counting method. The calculated fractal dimension for that window is then assigned to the central part of that window and store the results in a raster map. This process is repeated for every part in the dataset, creating a raster map of local fractal dimensions. compare locations with their neighbours. and are attributed Focal fractal dimension values for the white parts is =3, and for black parts is =2.

Focal fractal dimension can identify variations in image complexity, serving as a diagnostic feature for different structures²⁸. The Focal fractal dimension analysis is a convenient method to measure contrasting morphologic changes within stimulated Comet assays.

There are several ways to assess structural Comet assay organization. One method today commonly used for evaluation of Comet assay structural features is fractal analysis. Fractal dimension and lacunarity²⁹ of binarized tissue micrographs as fractal analysis parameters are important indicators of complexity and heterogeneity^{28,30,31}.

Treated Comet assays show the greatest *Focal fractal dimension* values and complexity, and represent the greatest degree of stress.

Multi-scale focal fractal dimension cannot be performed at separate images, as they are too small, so, this analysis was performed on the image collage. However, the window seems too small (11 times 11 pixels) for images as large as 3066 times 1022 pixel and strong noise is produced in the Multi-scale focal fractal dimension spectra!

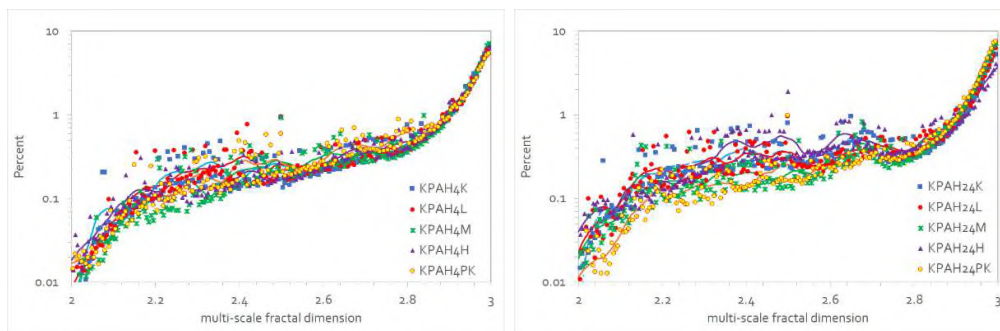


Figure 80 Multi-scale focal fractal dimension spectra in Image Collages for 4h (left) and 24 h (right)

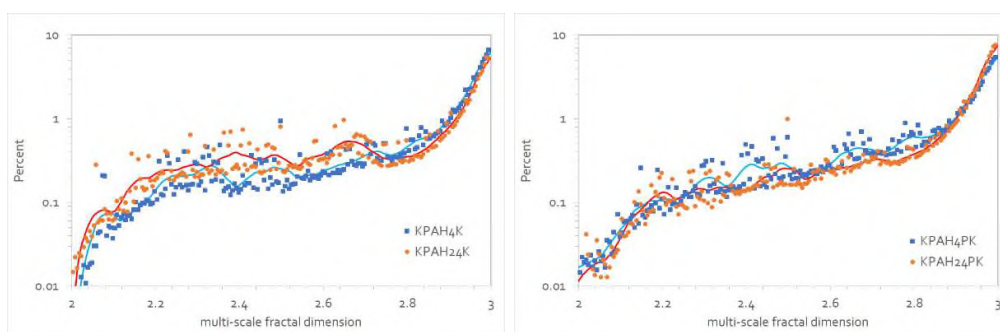


Figure 81 Multi-scale focal fractal dimension spectra in Image Collages for control K and positive control PK at 4h (blue, ■) and 24 h (brown, ●)

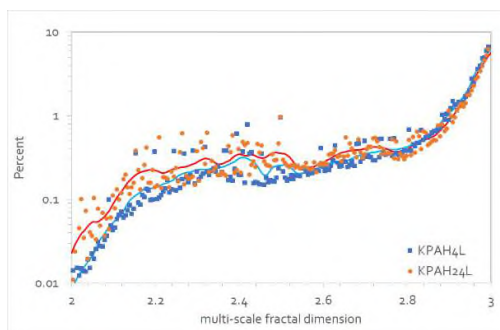


Figure 82 Multi-scale focal fractal dimension spectra in Image Collages for exposure L at 4h (blue, ■) and 24 h (brown, ●)

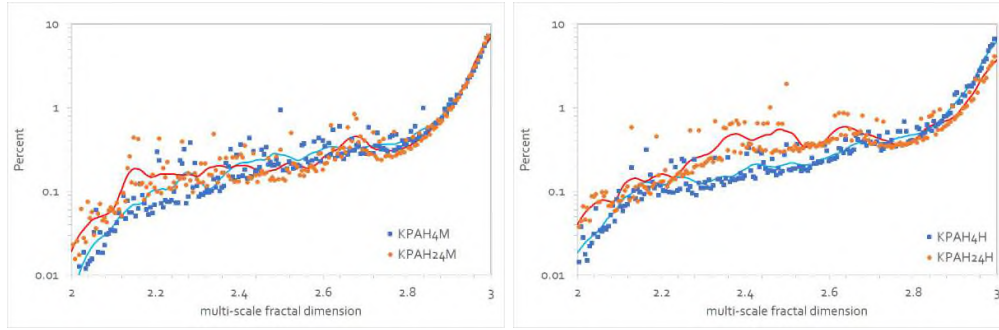


Figure 83 Figure 84 Multi-scale focal fractal dimension spectra in Image Collages for exposure M (left) and exposure H (right) at 4h (blue, ■) and 24 h (brown, ●)

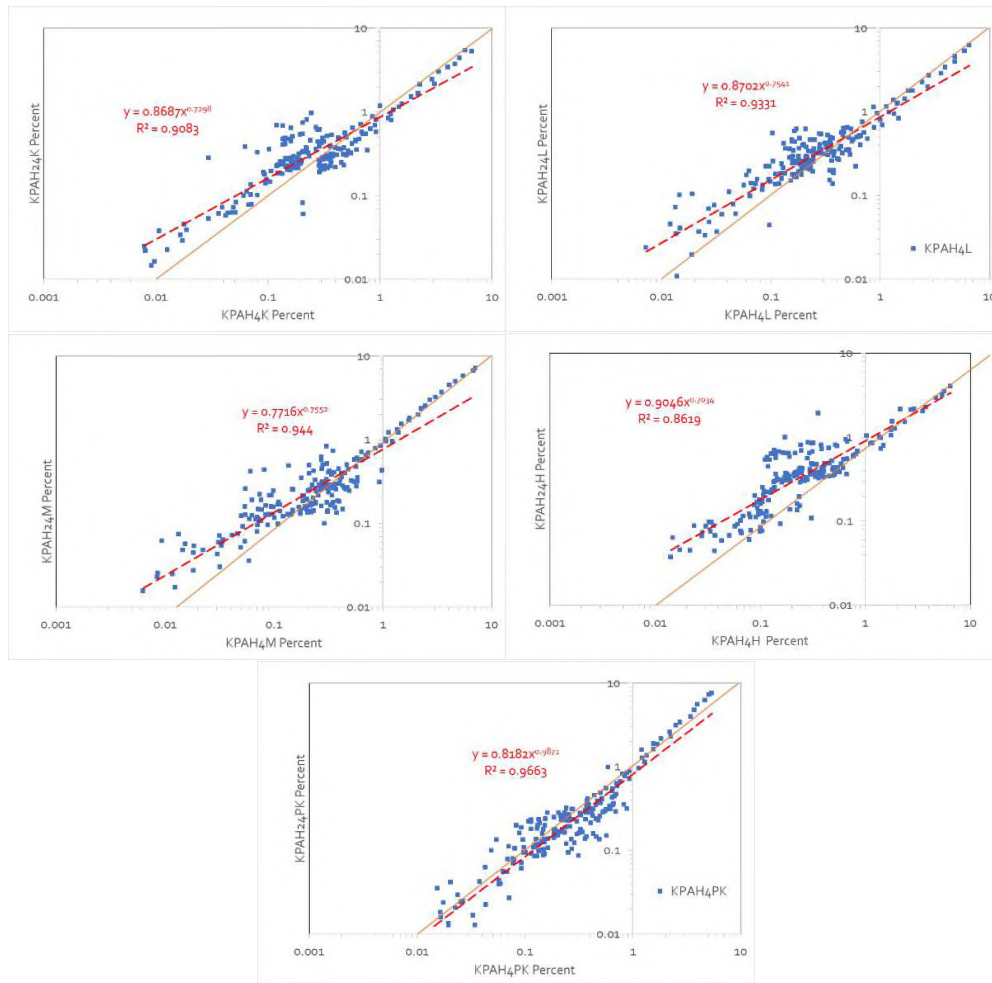


Figure 85 Correlation of percents at defined Focal multiscale fractal dimension (window 11x11) of Image Collages at 4h (x-ordinate) and 24h (y-ordinate)

Correlation of percents at defined Focal multiscale fractal dimension (window 11x11) of Image Collages at 4h (x-ordinate) and 24h (y-ordinate) are almost perfect for Focal multiscale fractal dimensions percents above 1% (i.e., for above Focal multiscale fractal dimension ≈ 2.8) (Figure 85). However, the

exposure M is an outlier maybe due to apparent piecewise double power above 1% (i.e., for above Focal multiscale fractal dimension \approx 2.9).

Table 3 Power function ($y=\text{intercept } x^{\text{power}}$) fit of Correlation of percents at defined Focal multiscale fractal dimension (window 11x11) of Image Collages at 4h (x-ordinate) and 24h (y-ordinate).

	intercept	power	R ²
K	0.8687	0.7298	0.9083
L	0.8702	0.7541	0.9331
M	0.7716	0.7552	0.944
H	0.9046	0.7034	0.8619
PK	0.8182	0.9871	0.9663

Table 4 50% Focal multiscale fractal dimension and percent at 50% Focal fractal dimension Focal multiscale fractal dimension (window 11x11) of Image Collages at 4h and 24h.

	4h		24h	
	50% Focal fractal dimension	Mean percent at 50% Focal fractal dimension, %	50% Focal fractal dimension	Mean percent at 50% Focal fractal dimension
K	2.605	0.2209	2.79	0.2971
L	2.61	0.2292	2.49	0.2939
M	2.21	0.2326	2.32	0.2142
H	2.575	0.2250	2.565	0.3578
PK	2.59	0.2445	2.36	0.2101

Multifractals: A theory of measures

While monofractals are mostly concerned with spaces, multifractals deal with measures. Even if the idea behind multifractals is also to study the complexity and reveal the scaling properties of a mathematical object, those two concepts are distinct. Multifractal analysis is a mathematical tool used to measure morphologies characterized by multiple scaling laws^{32,33,34}. Originating from research related to energy dissipation during developed turbulence, multifractal analysis has since been applied to a wide range of complex systems. The comprehensive theoretical foundation of multifractal analysis has been elaborated in many works³⁵. This paper focuses on several key calculations for almost implementing this analysis using computational tools. Multifractal analysis offers a sophisticated method to examine the complex morphology of Comet assays, which traditionally have been analysed using monofractal techniques.

We employed the box-counting method with non-overlapping boxes to determine the two most commonly applied spectra: the spectra of generalized dimensions $DQ(Q)$, and the singularity spectra $f(a)$ vs. a . The multifractal spectrum is the curve $f(a)$ against a . It gives, roughly speaking, the “fractal dimension” of sets where the measure scales locally with the same exponent. Singularity spectrum $f(a)$ calculates the fractal dimension of the subset of pixels in an image described by a particular exponent, referring to the relative dominance of various fractal exponents involved in the structure³⁶. The spectrum is formed by calculating the parameter α , also known as the Hölder exponent, which represents the local regularity or irregularity of a point on an object at a specific scale³⁷. This allows the formation of subsets of points with similar α , which will act as objects with monofractal properties stemmed from the original multifractal object.

Fractal analysis was performed³⁸ using ImageJ software (National Institutes of Health, USA) and the plugin FracLac^{39,40,41,42} (Version 2.5, Release 1e).

During the fractal analysis, comet assays were automatically converted into binary format (black and white image) by using the following protocol/setting: ImageJ> Plugins> Fractal analysis> FracLac_2.5 Release 1e> Standard box count> Auto Threshold to Binary. For each comet assays image, values of fractal dimension were determined.

In this study, values of the parameter Q ranged from -10.0 to 10.0 with a step of 0.1 , providing each multifractal spectrum with 201 points. Each point in the spectra was treated as an individual variable. Each image was associated with two spectra: $DQ(Q)$ vs. Q , and $f(a)$ vs. a .

Dimensional Ordering is a general trend important in interpreting results in multifractal analysis a rule for the generalize dimension, where the Capacity Dimension \geq the Information Dimension \geq the Correlation Dimension or:

$$D(Q=0) \geq D(Q=1) \geq D(Q=2)$$

Several metrics were used for the characterization of multifractal spectra [49]. These metrics analyse the differences between the mean spectrum of the control group (healthy retinas) and those of the DR and Glaucoma groups, to

determine the multifractal behaviour within each of the three retinal groups. The multifractal spectra are represented by the values of α and $f(\alpha)$ and are calculated based on the magnifying parameter q . The multifractal spectra are characterized not only by α_0 and $f(\alpha_0)$, but also by several other parameters derived from α and $f(\alpha)$ to describe the spectrum more comprehensively. These parameters include α_{max} , the dominant singularity; the asymmetry, A :

$$\text{Asymetry, } A = \frac{\alpha_{max} - \alpha_{left}}{\alpha_{right} - \alpha_{max}}$$

Other parameters descriptor parameters of the multifractal spectrum have been defined. These include $\Delta\alpha_{left} = \alpha_{max} - \alpha_{left}$, and $\Delta\alpha_{right} = \alpha_{right} - \alpha_{max}$, which represent the width of the left and right sides of the multifractal spectrum, respectively; $\Delta f(\alpha_{left}) = f(\alpha_{max}) - f(\alpha_{left})$, and $\Delta f(\alpha_{right}) = f(\alpha_{right}) - f(\alpha_{max})$, representing the height of the left and right side of the spectrum;

$$L_{left} = \sqrt{(\alpha_{max} - \alpha_{left})^2 + (f(\alpha_{max}) - f(\alpha_{left}))^2}$$

$$L_{rig} = \sqrt{(\alpha_{right} - \alpha_{max})^2 + (f(\alpha_{right}) - f(\alpha_{max}))^2}$$

which correspond to the length of the left (multifractal) and right (monofractal) sides of the multifractal spectrum, respectively.

We extracted more variables per spectrum: maximum and minimum value of a parameter, the range or span (the difference between the maximum and minimum values): DQ_{min} , DQ_{max} , DQ_{span} ; $f(\alpha)_{min}$, $f(\alpha)_{max}$, $f(\alpha)_{span}$.

The calculation of the multifractal spectrum of comet assays images is ineffective due to artifacts that can lead to an unreliable result. The search for methodological approaches to using multifractal image analysis in relation to the assessment of the comet assays is relevant.

Multifractals

Consistent with the previous findings, some of the test results showed no statistically significant differences between the groups for any variable. Thus, the morphology of Comet assays in the comet assay images within this sample did not change significantly with treatment.

The generalized dimension DQ for the three treatments are presented in Figure 86 and compared to control (K) and positive control (PK) in Figure 87.

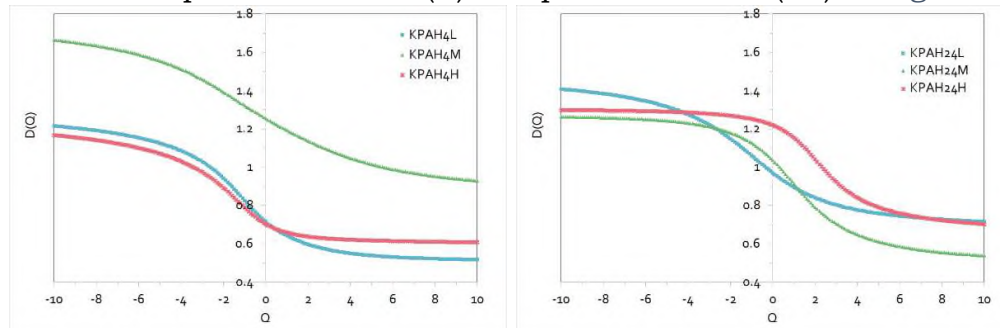


Figure 86 The generalized dimension DQ spectra for the three treatments at 4h (left) and 24h (right): ■ L, ▲ M and x H exposures.

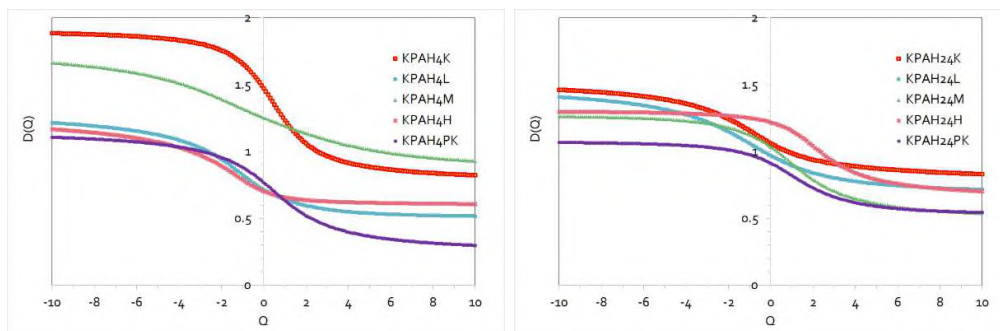


Figure 87 The generalized dimension DQ spectra for the three treatments at 4h (left) and 24h (right): ■ L, ▲ M and x H exposures, and control (K) □ and positive control (PK) ●.

Values of the generalized dimension DQ showed a “step-like” go down across almost the entire Q - range of the spectra with increasing exposure, except for M-exposure (Figure 87). The biggest differences between the groups were observed for Q negative range (between exposure M and exposures {L, H, and positive control (PK)}) (Figure 87, left). For 24h time, these differences are not so obvious (Figure 87, left).

The multifractal spectra for the three treatments are presented in Figure 88 and compared to control (K) and positive control (PK) in Figure 89.

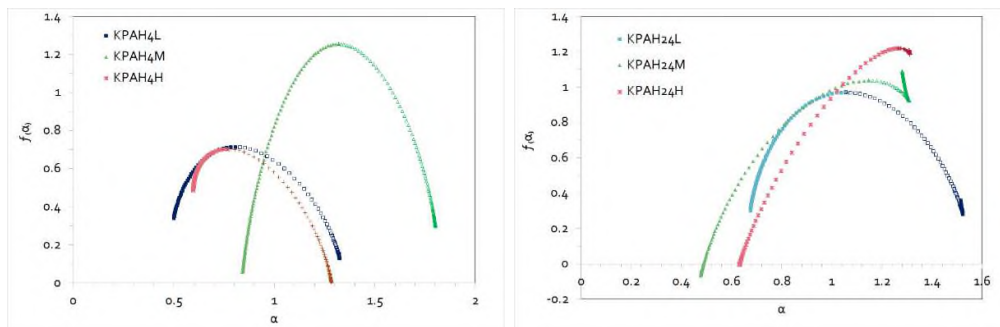


Figure 88 The multifractal spectra for the three treatments at 4h (left) and 24h (right): ■ L, ▲ M and x H exposures.

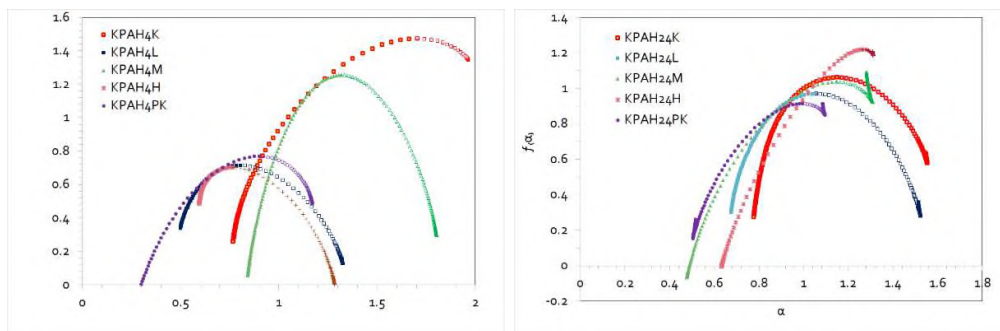


Figure 89 The multifractal spectra for the three treatments at 4h (left) and 24h (right): ■ L, ▲ M and x H exposures, and control (K) □ and positive control (PK) ●.

The generalized dimension DQ spectra for separate images

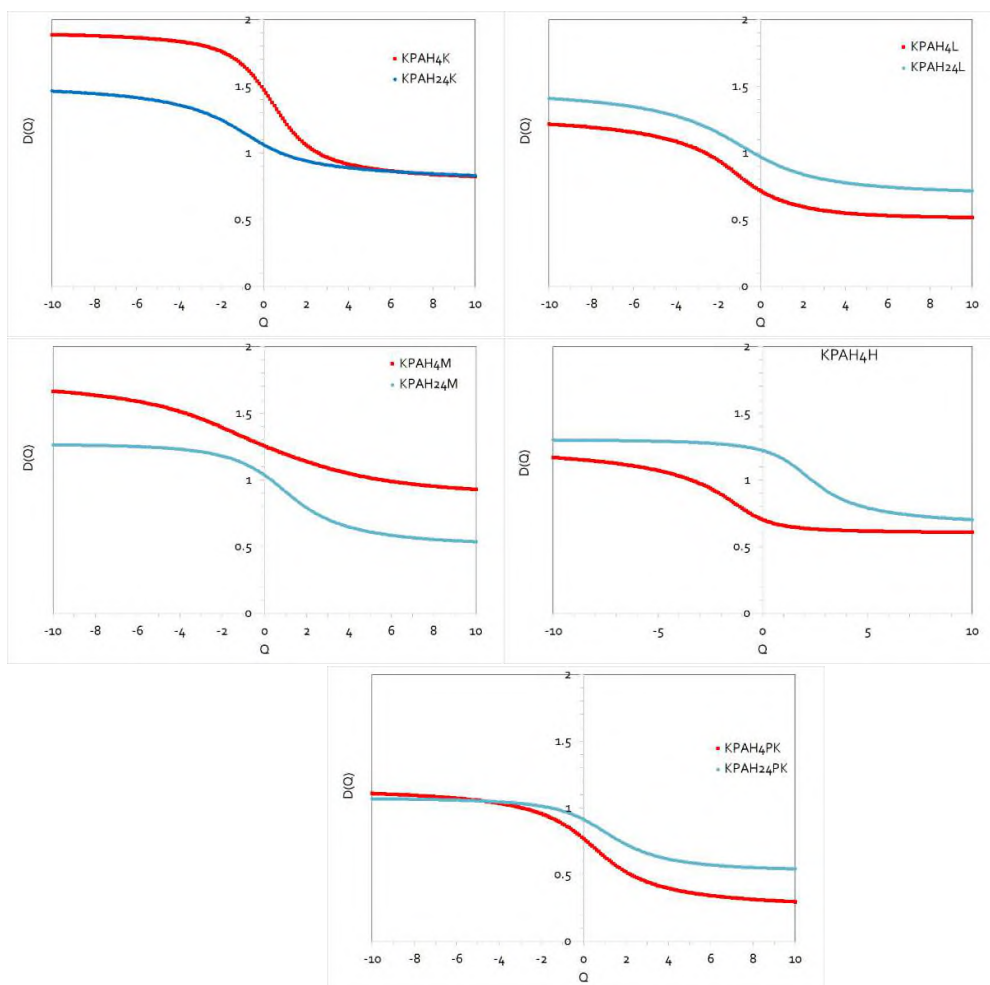


Figure 90 The generalized dimension DQ spectra for the three treatments, and control (K) and positive control (PK), at 4h (red, ■) and 24h (blue, ●).

The multifractal spectra for separate images

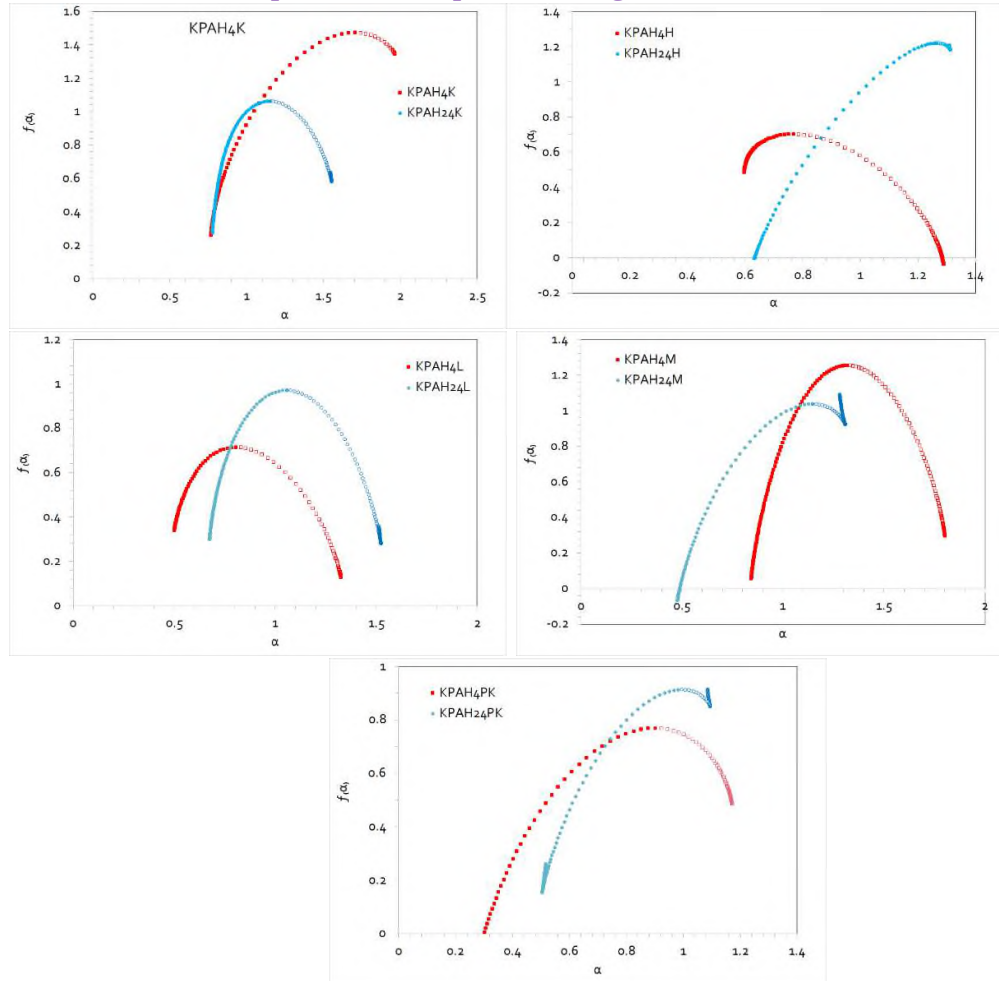


Figure 91 The multifractal spectra for the three treatments, and control (K) and positive control (PK), at 4h (red, ■) and 24h (blue, ●).

A decrease in the height of the spectrum ($(a)_{max}$ and $f(a)_{max}$, Table 6) indicates a “simplification” of the pattern of comet assays.

A decrease in the width of the singularity spectrum (i.e., $(a)_{span}$ and $f(a)_{span}$, Table 7) indicates a more uniform distribution of the structural parts of comet assays.

It is important to note that one challenge of this analysis is the practical interpretation of the singularity spectrum results. The practical significance of these spectra can often be “elusive” and abstract, making it difficult to intuitively understand individual spectra and relate the results to the actual images. Also, the limitation is the pixel-based image. Despite using relatively high-resolution images (3066 × 1022 pixels) in this study, various fine structures of the analysed Comet assays’ morphology can still be obscured, which may prove problematic with larger mathematical distortions.

The spectrum value parameters are presented in Table 5-Table 7 for each treatment and time group. These parameters were extracted for each collage of individual images.

Table 5 The generalized dimension DQ spectra parameters

	DQ_{max}	DQ_{min}	DQ_{span}	Capacity Dimension, D0	Information Dimension, D1	Correlation Dimension, D2
4h						
K	1.8858	0.8186	1.0672	1.4732	1.2342	1.0600
L	1.2151	0.5157	0.6994	0.7146	0.6421	0.5969
M	1.6645	0.9214	0.7430	1.2541	1.1925	1.1372
H	1.1678	0.6066	0.5612	0.7030	0.6583	0.6377
PK	1.1079	0.2927	0.8152	0.7691	0.6335	0.5209
24h						
K	1.4628	0.8264	0.6364	1.0617	0.9876	0.9404
L	1.4080	0.7126	0.6955	0.9703	0.8949	0.8404
M	1.2630	0.5328	0.7302	1.0363	0.9124	0.7915
H	1.2987	0.6964	0.6023	1.2194	1.1539	1.0414
PK	1.0689	0.5422	0.5267	0.9133	0.8203	0.7277

Table 6 The multifractal spectra parameters

	(a)_{left}	f(a)_{left}	(a)_{max}	f(a)_{max}	(a)_{right}	f(a)_{right}
4h						
K	0.7707	0.2912	1.7033	1.4732	1.9636	1.3469
L	0.4996	0.3401	0.8059	0.7146	1.3211	0.1554
M	0.8420	0.0558	1.3189	1.2541	1.8012	0.2969
H	0.5955	0.4852	0.7669	0.7030	1.2881	-0.0349
PK	0.2351	-0.3346	0.9000	0.7692	1.1701	0.4853
24h						
K	0.7758	0.2747	1.1499	1.0617	1.5457	0.6337
L	0.6748	0.3004	1.0555	0.9703	1.5131	0.3577
M	0.4811	-0.0314	1.1462	1.0363	1.2803	1.0900
H	0.6324	-0.0009	1.2645	1.2194	1.3076	1.2093
PK	0.5165	0.2616	0.9942	0.9133	1.0844	0.9141

Table 7 Derived multifractal spectra parameters

	(a) span	f(a) span	(a) Asymetry	f(a) Asymetry	L₁ (multifractal) side	L₂ (monofractal) sides
4h						
K	0.9326	1.0557	3.5833	-9.3541	1.5057	0.2893
L	0.3064	-0.1847	0.5947	-0.6697	0.4838	0.7603
M	0.4769	0.2411	0.9888	-1.2519	1.2897	1.0719
H	0.1714	-0.5201	0.3289	-0.2952	0.2772	0.9034
PK	0.6649	0.8199	2.4620	-3.8886	1.2885	0.3918
24h						
K	0.3741	0.3590	0.9450	-1.8388	0.8714	0.5830
L	0.3807	0.0573	0.8321	-1.0936	0.7705	0.7646
M	0.6651	1.1213	4.9590	19.9055	1.2579	0.1445
H	0.6321	1.2102	14.6418	-120.5203	1.3743	0.0443
PK	0.4778	0.6525	5.2980	785.3386	0.8080	0.0902

The length of the left (multifractal) and right (monofractal) sides of the multifractal spectrum

The length of the left (multifractal) and right (monofractal) sides of the multifractal spectrum are decreasing and decreasing at 24h, respectively (Figure 92).

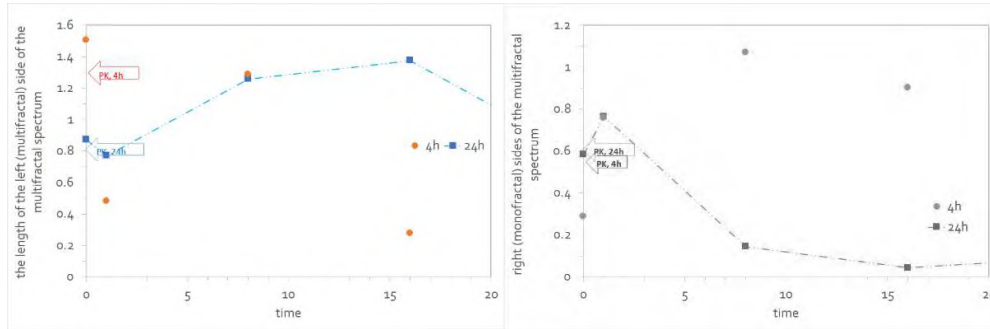


Figure 92 The length of the left (multifractal) and right (monofractal) sides of the multifractal spectrum, respectively at 4h (left, ●) and 24h (right, ■); and positive control (PK) is indicated by shape arrows.

Euler number in COMET ASSAY images

At its core, the Euler number is a topological invariant that describes an object's connectivity properties independent of geometric transformations such as scaling or rotation. In image processing, it is typically applied to binary images, where pixels are categorized as either foreground (object) or background. The morphological Euler number^{43,44,45,46} is a topological invariant calculated as the number of connected components (objects) in an image minus the number of holes in those objects. For a 2D image, the formula is $E = \beta_0 - \beta_1$, where β_0 is the number of components and β_1 is the number of holes.

The Euler typical of a binary image on a continuous 2D plane is calculated. The Euler number method consists in thresholding the image at t different levels, computing the Euler number for each binarization⁴⁷. The Euler number is plotted against the threshold level. The graph typically shows a decrease as the threshold increases. For grayscale images, repeating the Euler typical calculation for multiple threshold levels will result in complexity, related to the number of pixels in the image and the number of threshold values

Component Count β_0 : This is the number of separate, distinct objects or "blobs" in the image. A connected component in a binary image refers to the connected region composed of adjacent pixels with the same pixel value (usually object pixel) in a binary image.

Holes Count β_1 : This is the total number of "holes" within the components. A hole in the image refers to the regions with non-object pixels surrounded by a connected component.

Morphological Euler Number, E: The final result, which indicates the overall topology of the image, is found by subtracting the number of holes from the number of components.

A single, solid disk would have one component $\beta_0=1$ and zero holes $\beta_1=0$, resulting in an Euler number of $E=1-0=1$. A figure-eight shape would have one component ($\beta_0=1$) but two holes, so its Euler number would be $E=1-2=-1$. Noise can introduce small artifacts or holes in the image, which can change the number of objects or holes counted, thus changing the Euler number. Preprocessing steps like denoising and morphological operations are often needed to obtain correct Euler number measurements.

Comparison with Betti Numbers

Betti numbers generalize the Euler characteristic by separately counting the number of connected components, holes, and higher-dimensional voids. In 2D images, the Euler number relates to Betti numbers as:

$$\text{Euler Number} = \beta_0 - \beta_1$$

where β_0 represents the number of connected components and β_1 the number of holes. While Betti numbers provide more granular information, the Euler number condenses this into a single scalar, helping with simpler analysis but potentially losing detail.

Complementary Features

Other features like perimeter, area, convexity, and fractal dimension often go along with Euler number in image analysis pipelines. The Euler number uniquely captures connectivity and topology, parts often invisible to geometric descriptors.

Results

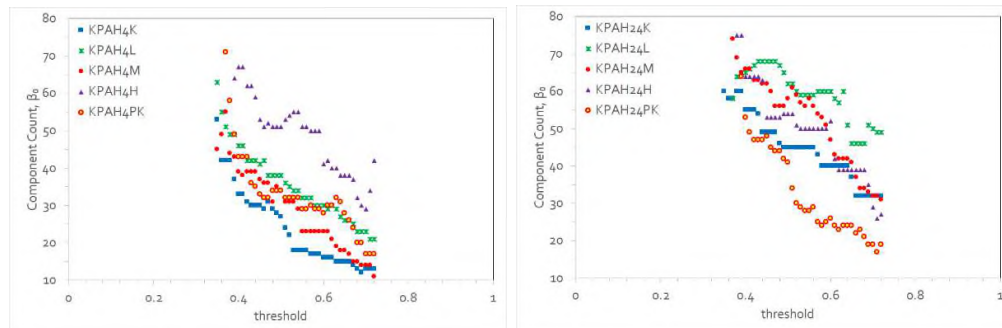


Figure 93 Component Count β_0 at 4h (left) and 24h (right)

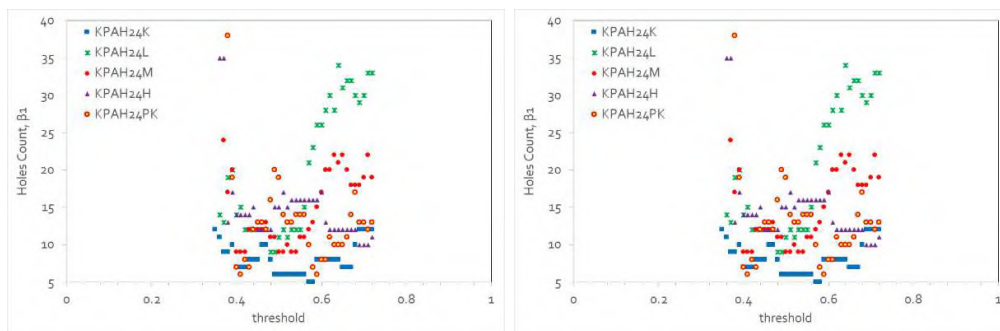


Figure 94 Holes Count β_1 at 4h (left) and 24h (right)

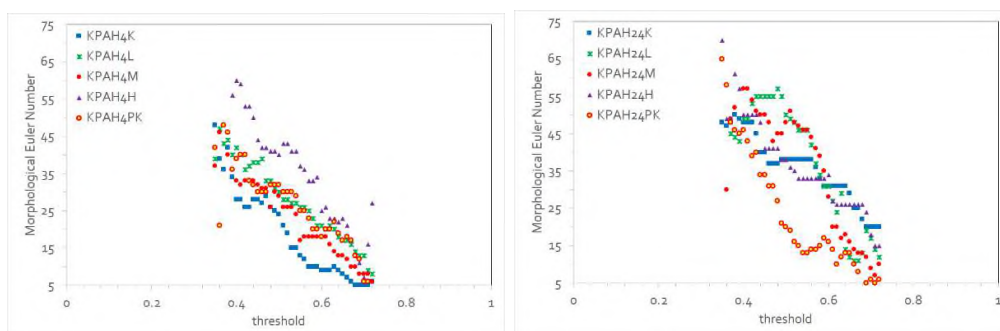


Figure 95 Morphological Euler Number, E at 4h (left) and 24h (right)

Table 8 Component Count, Holes Count and Morphological Euler Number @ threshold=0.5 in COMET ASSAY images

	exposure	Component Count @ t=0.5	Holes Count@ t=0.5	Morphological Euler Number @ t=0.5
KPAH4K	0	27	3	24
KPAH4L	1	38	8	30
KPAH4M	8	34	5	29
KPAH4H	16	51	9	40
KPAH4PK	24	34	2	32
KPAH24K	0	45	6	38
KPAH24L	1	62	11	50
KPAH24M	8	58	9	48
KPAH24H	16	54	15	38
KPAH24PK	24	41	19	20

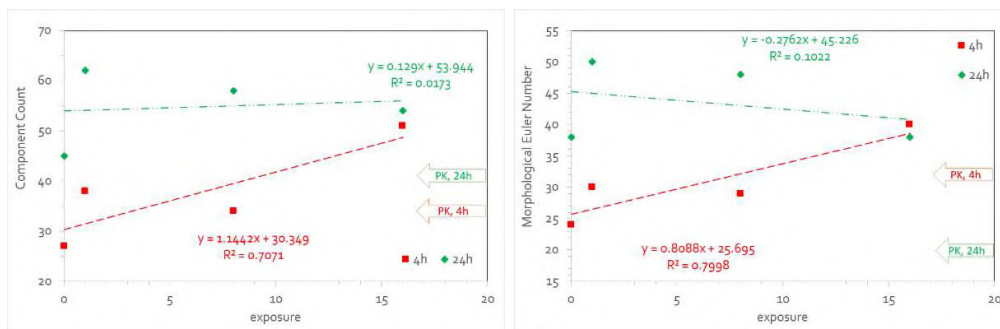


Figure 96 Correlations of Component Count @ $t=0.5$ versus exposure (left) Morphological Euler Number @ threshold=0.5 versus exposure (right) at 4h (red, ■) and 24h (green, ◆)

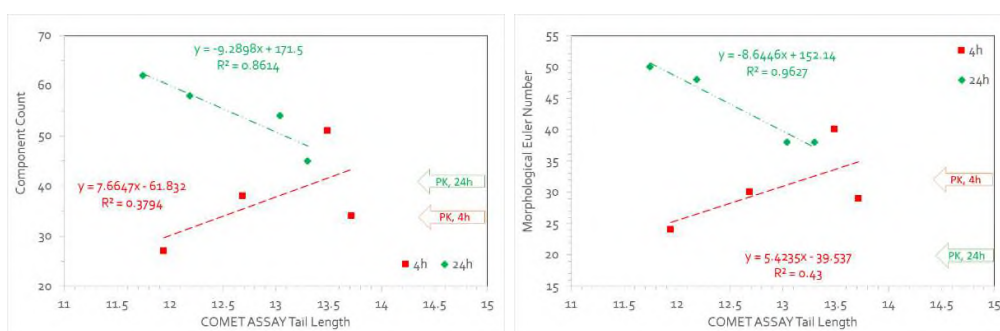


Figure 97 Correlations of Component Count @ $t=0.5$ versus **Comet Assay Tail Length** (left) and Morphological Euler Number @ threshold=0.5 versus **Comet Assay Tail Length** (right) at 4h (red, ■) and 24h (green, ◆)

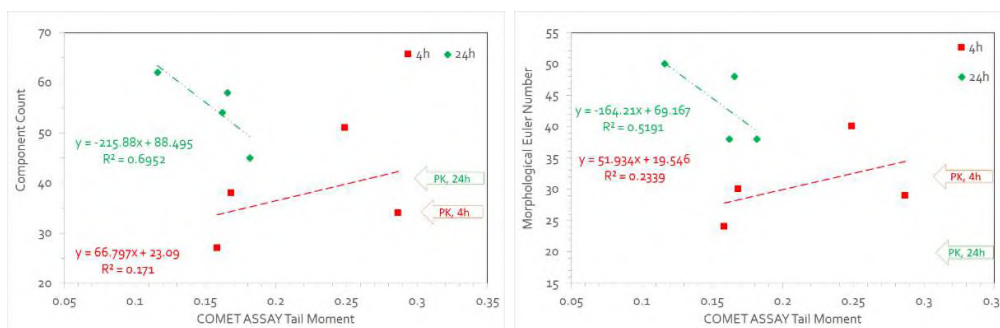


Figure 98 Correlations of Component Count @ $t=0.5$ versus **Comet Assay Tail Moment** (left) and Morphological Euler Number @ threshold=0.5 versus **Comet Assay Tail Moment** (left) and Morphological Euler Number @ threshold=0.5 versus (right) at 4h (red, ■) and 24h (green, ◆)

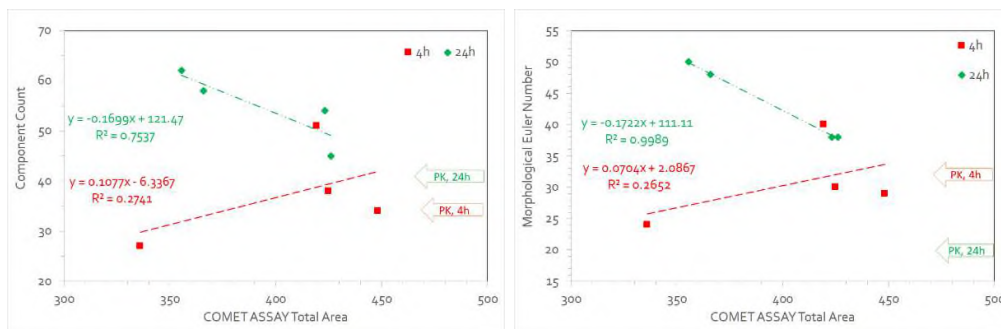


Figure 99 Correlations of Component Count @ t=0.5 versus **Comet Assay Tail Moment** (left) and Morphological Euler Number @ threshold=0.5 versus **Comet Assay Tail Moment** (left) and Morphological Euler Number @ threshold=0.5 versus (right) at 4h (red, (right) at 4h (red, ■) and 24h (green, ◆)

Minkowski scalars

Minkowski functionals from integral geometry^{48,49,50} are powerful and versatile shape descriptors for random spatial structures in real space^{51,52}. They provide a localized and comprehensive shape analysis by characterizing, among others, symmetries and preferred directions. Minkowski scalars⁵³, also known as intrinsic volumes, are rotation-invariant shape descriptors that measure the size and complexity of objects robustly and sensitively⁵⁴, such as area, perimeter, and the Euler characteristic for 2D shapes, and volume, surface area, mean width, and Euler characteristic for 3D shapes. These scalars come from integral geometry and are based on integrals of the shape's boundary and curvature. Due to their rotation invariance, they don't capture shape orientation, but are the foundation for Minkowski tensors, which provide more detailed information about shape anisotropy. The Minkowski scalars can be interpreted as area, perimeter, and the Euler characteristic. The latter is a topological measure; for compact bodies it is given by the number of parts minus the number of cavities. All three can be written as integrals, weighted by the area, the contour length, and the local curvature of the contour.

Isoperimetric Ratio

A classical index of anisotropy is the isoperimetric ratio Q , (or perimetric complexity) which measures compactness of a shape. The perimetric complexity (PC) is a measure of the complexity of binary images⁵⁵. The general concept goes back to the early days of vision research where this measure, originally called dispersion, was used to describe the perceptual complexity of visual shapes. In 2D, it is the dimensionless quotient of the area and perimeter of an object. For a simple closed plane curve, the isoperimetric ratio is $(4\pi A/P^2)$, where P is the length (perimeter) of the curve, and A is the area enclosed by the curve, it encloses. This ratio is a dimensionless value that measures how "round" or far from circular a shape is. This ratio is invariant under scaling: changing the size of the shape does not change this ratio:

$$Q = 4\pi \frac{Area}{(2 \times Perimeter)^2}$$

The isoperimetric ratio of a planar domain, as a stable means of evaluating geometric compactness, measures the minimum perimeter needed to inscribe a shape with prescribed area varying from 0 to the area of the domain. For example, a perfect circle is the most compact shape achievable, with isoperimetric ratio 1, while a shape with fractalesque boundary will have a ratio close to 0.

Examples:

Circle:

For a circle with radius r , the perimeter L is $2\pi r$ and the area A is πr^2 . The ratio

$$\frac{L^2}{A} = \frac{(2\pi r)^2}{\pi r^2} = \frac{4\pi^2 r^2}{\pi r^2} = 4\pi$$

Square:

A square with side length s has a perimeter L of $4s$ and an area A of s^2 . The ratio

$$\frac{L^2}{A} = \frac{(4s)^2}{s^2} = \frac{16s^2}{s^2} = 16$$

L^2/A is $(4s)^2 / s^2 = 16s^2 / s^2 = 16$. Since $16 > 4\pi$, *the square is less circular than the circle.*

While the isoperimetric ratio is a scalar measurement of compactness, the isoperimetric profile provides an entire plot that measures the multiscale compactness of a domain. It is more suited to noisy data that often exhibit fractal boundaries. Since boundaries can have fractal shapes, their lengths are not well-defined and can change at different length scales. The isoperimetric ratio is unstable to boundary perturbation and will change its compactness score depending on the resolution of the input shape. Different shape information is in the isoperimetric ratio and the Minkowski anisotropy indices, as seen in the comparison of $\beta_1^{0.2}$ and Q for various geometric shapes. Elliptical objects can have a relaxed, spherical shape, or be of prolate or oblate morphology. In addition, imperfect denoising and incomplete deblending may leave merging objects of unidentifiable class and amorphous shape. An anisotropy analysis for each object yields the following graph of the isoperimetric ratio Q versus q_2 , which easily allows us to separate the objects from the elliptical objects. The latter are difficult to distinguish from the amorphous merging structure. Converting q_2 to the ratio of eigenvalues $\beta_1^{0.2}$ yields a similar graph, in which the objects are even better separable because their distances with respect to each other are larger due to the (non-linear) transformation between q_2 and $\beta_1^{0.2}$. The general trend of “boxiness” versus “sphericity” as also in this plot is recovered for these objects.

We measure these structures by calculating global and local topological and morphological measures, namely Minkowski functionals (MFs) and by deriving suitable texture measures based on MF.

Gray-scale images are analysed by constructing isocontour with the (interpolated) marching squares algorithm and analysing the shapes with Minkowski Tensors. Gray-scale images are analysed with the interpolated marching squares algorithm, an enhancement of the regular marching squares algorithm.

Minkowski Structure Metrics qs

Often, the orientation and size of an object is not of interest, and a metric invariant under rotation, scaling, and translation is required. This is provided by the Minkowski Structure Metrics (MSMs) qs. The qs are invariant under rigid motions (rotation and translation) and scaling. qs detects s-fold-symmetric part in the normal density. For example, q2 is sensitive to the quadrupole part, and is sensitive to rod-like shapes. q3 can be used to find shapes with predominantly threefold symmetry, such as equilateral triangles. q6 is used to detect regular hexagons, and to find hexagonal order in 2D. The qs are fingerprints of a particular class of shapes, and blind to rotation and scaling. The qs may thus be used to classify shapes, ignoring their size and distinguished orientation. Each of Minkowski tensor of rank s distinguishes s spatial directions. These are the directions in which “most” of the interface is oriented. For example, distinguished directions for q2 is \mid , for q3 is \top , for q4 is \perp , for q5 is \ast and for q6 is \times .

Table 9 Minkowski Structure Metrics qs and Isoperimetric ratio $Q=4\pi A/P^2$; $\beta_1^{02}=(1 - q_2)/(1 + q_2)$ (anisotropy index)

	K		L		M		H		PK	
4h	Av.	st. dev.	Av.	st. dev.	Av.	st. dev.	Av.	st. dev.	Av.	st. dev.
q6	0.069	0.040	0.098	0.060	0.087	0.046	0.095	0.052	0.086	0.019
q5	0.033	0.020	0.027	0.016	0.029	0.017	0.026	0.015	0.016	0.013
q4	0.267	0.151	0.202	0.128	0.297	0.134	0.214	0.061	0.404	0.157
q3	0.033	0.020	0.027	0.016	0.029	0.017	0.026	0.015	0.016	0.013
q2	0.069	0.040	0.098	0.060	0.087	0.046	0.095	0.052	0.086	0.019
Area	4677.0	2358.4	6440.6	3217.6	5810.6	2966.9	6481.2	3992.4	9213.8	4183.8
Perimeter	1181.3	1726.6	1693.5	1913.0	1287.5	1605.5	1210.1	903.1	2775.8	1746.8
Area fraction	0.020	0.010	0.026	0.014	0.023	0.012	0.033	0.028	0.037	0.017
Q	0.135	0.076	0.133	0.074	0.112	0.061	0.093	0.061	0.028	0.022
β_1^{02}	0.873	0.069	0.840	0.093	0.842	0.079	0.830	0.086	0.842	0.032

Table 10 Minkowski Structure Metrics q_s and Isoperimetric ratio $Q=4\pi A/P^2$; $\beta_1^{02} = (1 - q_2)/(1 + q_2)$ (anisotropy index)

	K		L		M		H		PK	
24h	Av.	st. dev.	Av.	st. dev.	Av.	st. dev.	Av.	st. dev.	Av.	st. dev.
q6	0.081	0.019	0.066	0.024	0.070	0.022	0.067	0.017	0.120	0.053
q5	0.019	0.009	0.030	0.015	0.031	0.012	0.015	0.013	0.020	0.011
q4	0.337	0.088	0.302	0.120	0.212	0.104	0.535	0.163	0.276	0.077
q3	0.019	0.009	0.030	0.015	0.031	0.011	0.015	0.013	0.020	0.011
q2	0.081	0.019	0.066	0.024	0.070	0.022	0.067	0.017	0.120	0.053
Area	5965.8	1486.4	5028.3	1283.0	5914.7	1900.4	12622.1	8090.4	7239.1	3734.5
Perimeter	1083.4	378.0	919.3	358.1	1005.5	470.2	2458.8	1021.0	1417.1	1025.1
Area fraction	0.024	0.006	0.020	0.005	0.024	0.008	0.051	0.032	0.029	0.015
Q	0.076	0.035	0.091	0.038	0.091	0.038	0.021	0.019	0.079	0.059
β_1^{02}	0.851	0.033	0.877	0.042	0.871	0.039	0.881	0.035	0.790	0.082

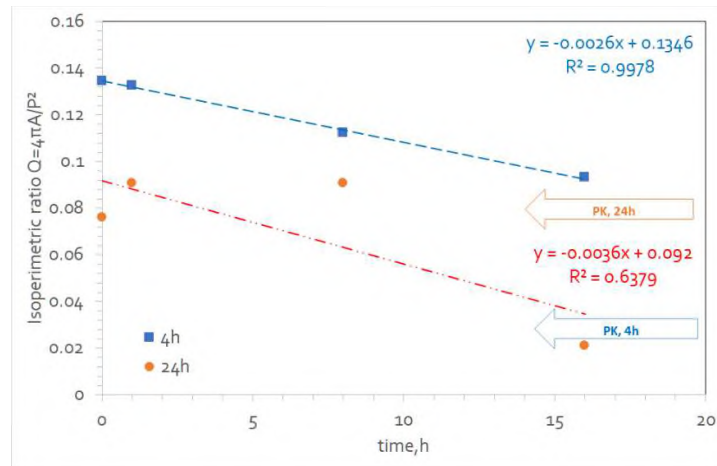


Figure 100 Isoperimetric ratio $Q=4\pi A/P^2$ in Image Collages

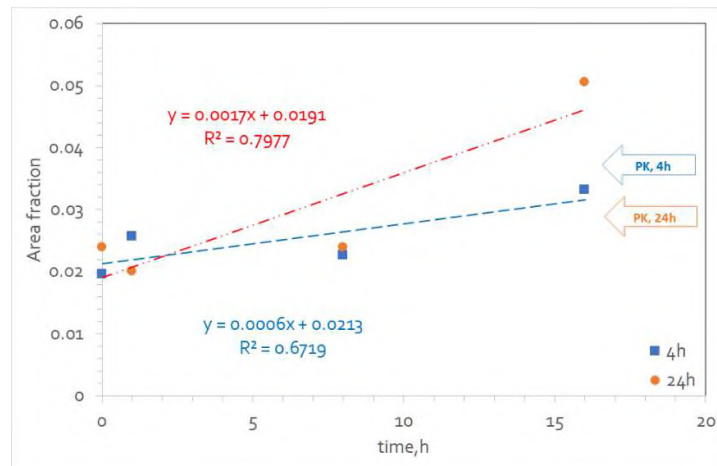


Figure 101 Area fraction in Image Collages)

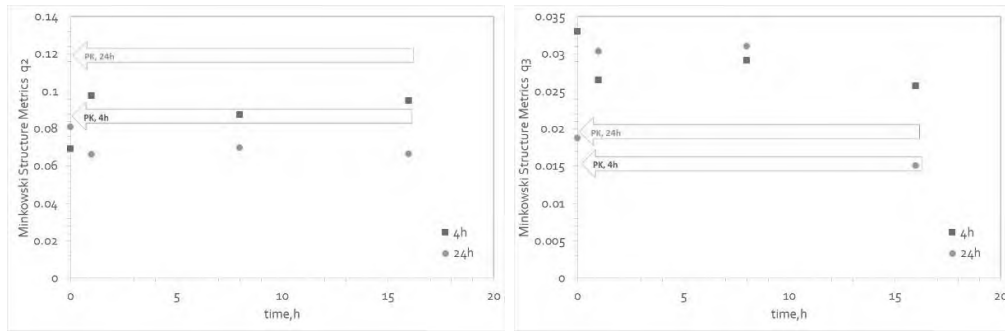


Figure 102 Minkowski Structure Metrics q_2 (left) and q_3 (right) in Image Collages

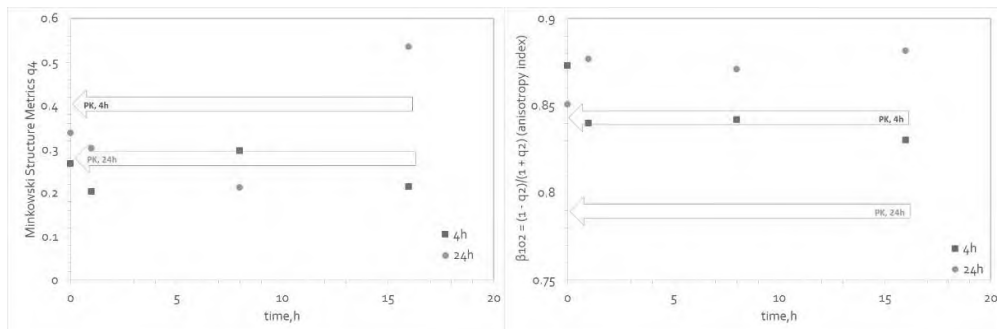


Figure 103 Minkowski Structure Metrics q_4 (left) and $\beta_1^{02} = (1 - q_2)/(1 + q_2)$ (anisotropy index) (right) in Image Collages

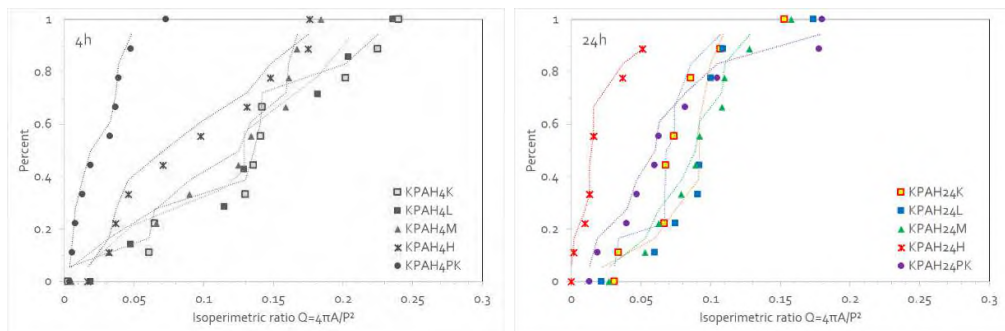


Figure 104 Percentile distribution of Isoperimetric ratio $Q=4\pi A/P^2$ in Image Collages for 4h (left) and 24 h (right)

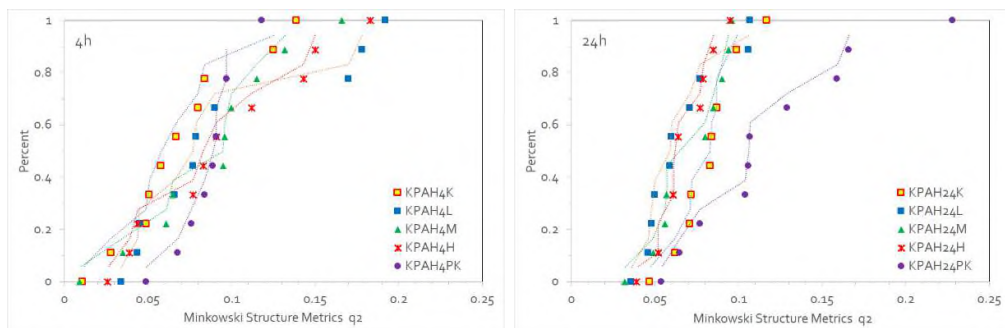


Figure 105 Percentile distribution of Minkowski Structure Metrics q_2 in Image Collages for 4h (left) and 24 h (right)

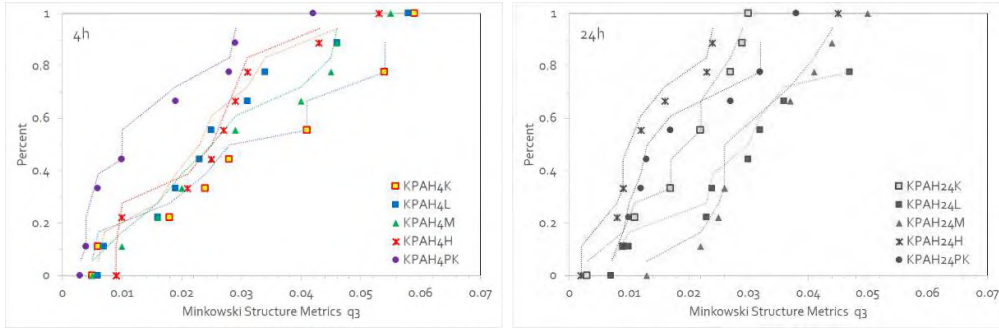


Figure 106 Percentile distribution of Minkowski Structure Metrics q_3 in Image Collages for 4h (left) and 24 h (right)

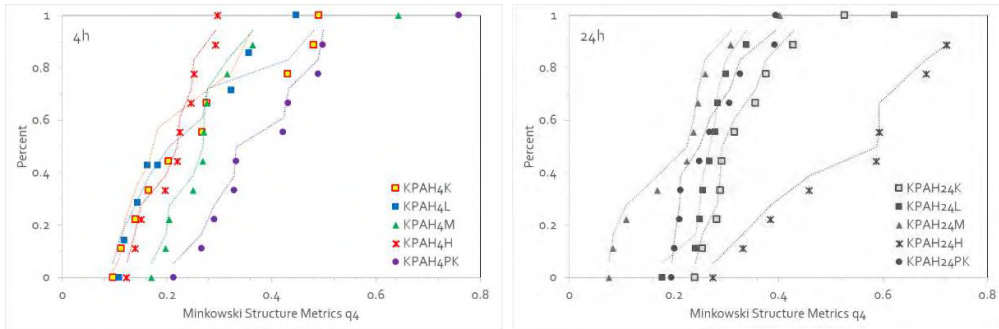


Figure 107 Percentile distribution of Minkowski Structure Metrics q_4 in Image Collages for 4h (left) and 24 h (right)

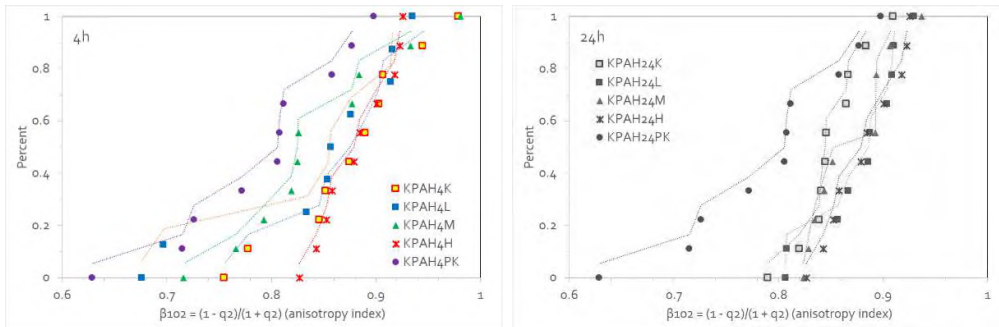


Figure 108 Percentile distribution of $\beta_1^{02} = (1 - q_2)/(1 + q_2)$ (anisotropy index) in Image Collages for 4h (left) and 24 h (right)

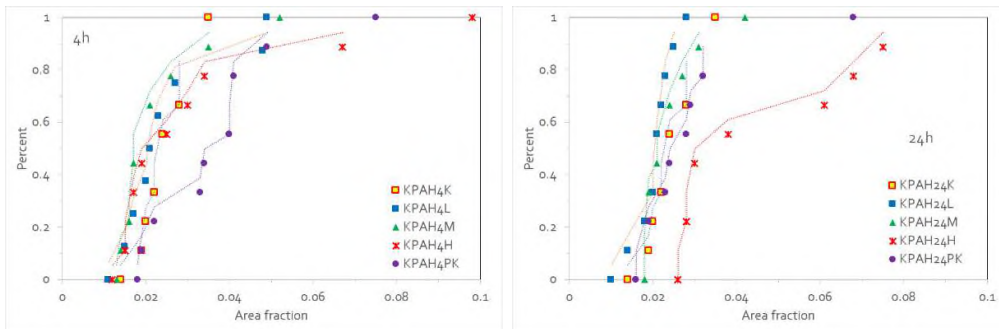


Figure 109 Percentile distribution of Area fraction in Image Collages for 4h (left) and 24 h (right)

Shannon Entropy of Collage Images

Shannon entropy returns the local randomness of a signal, commonly used to measure textures in an image, entropy is related to the complexity in a neighbourhood that can be used to characterize the texture of the input image. Shannon entropy, H , is computed from a sum of the form: $H = -\sum p_i \log[p_i]$, summing over the number of gray levels (256 for 8-bit images) and where p_i is the probability of each pixel value (gray level) occurring. The summation is over all possible pixel values. Using the base-2 logarithm provides the information in bits per pixel. Image entropy^{56,57,58} is a statistical measure of randomness or disorder in an image, reflecting its information content. It is calculated using the image's histogram (where the histogram shows the different grey level probabilities in the image) and the Shannon entropy formula, where a high entropy indicates a complex image with many pixel values, and a low entropy suggests a simple, uniform image with predictable patterns. Image entropy is calculated for images that were thresholded in the range 1-100. Entropy provides a way to measure the information in an image. When dealing with a distribution of different colors of pixels, the concepts of information, uncertainty, and evenness are roughly equivalent.

Things get more complicated when we consider the spatial structure of pixels in an image. Entropy takes different values in different locations.

Notice that the entropy of an image is rather different from the entropy feature extracted from the GLCM (Gray-Level Co-occurrence Matrix) of an image.

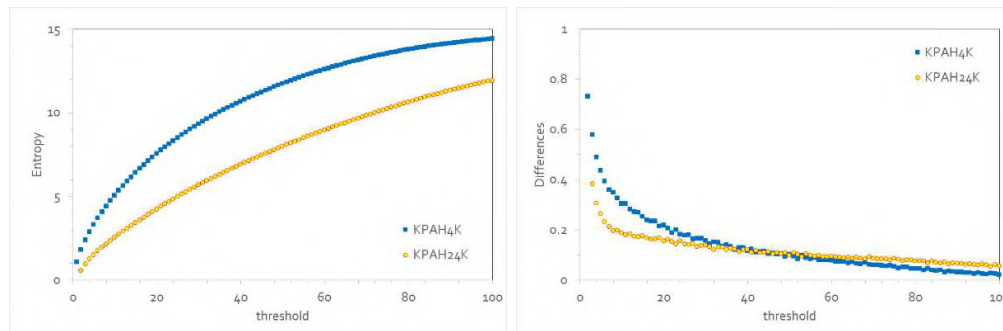


Figure 110 Comparison of Shanonn Entropy (left) and its derivation by threshold differences (right) at 4 h and 24 h.

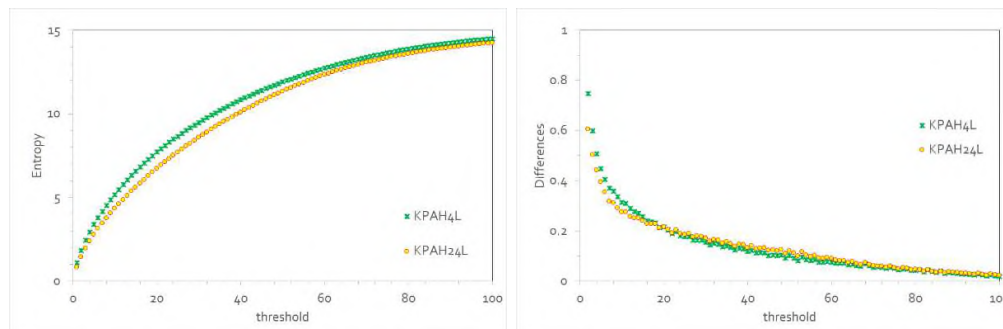


Figure 111 Comparison of Shanonn Entropy (left) and its derivation by threshold differences (right) at 4 h and 24 h.

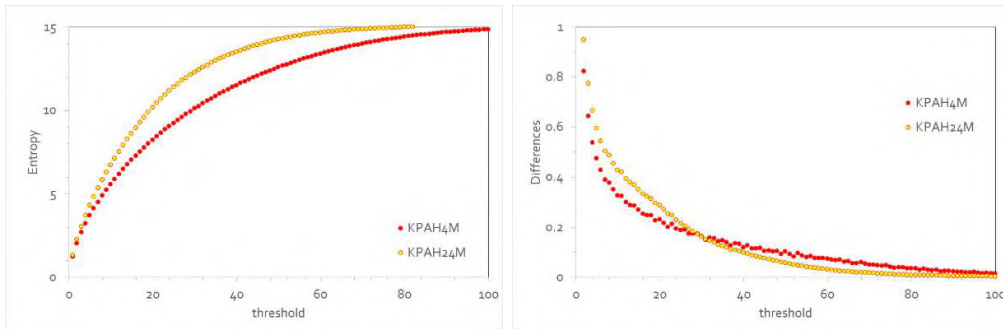


Figure 112 Comparison of Shanonn Entropy (left) and its derivation by threshold differences (right) at 4 h and 24 h.

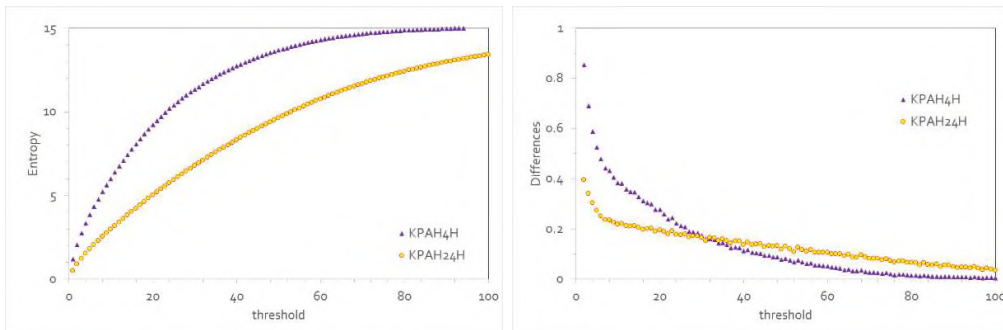


Figure 113 Comparison of Shanonn Entropy (left) and its derivation by threshold differences (right) at 4 h and 24 h.

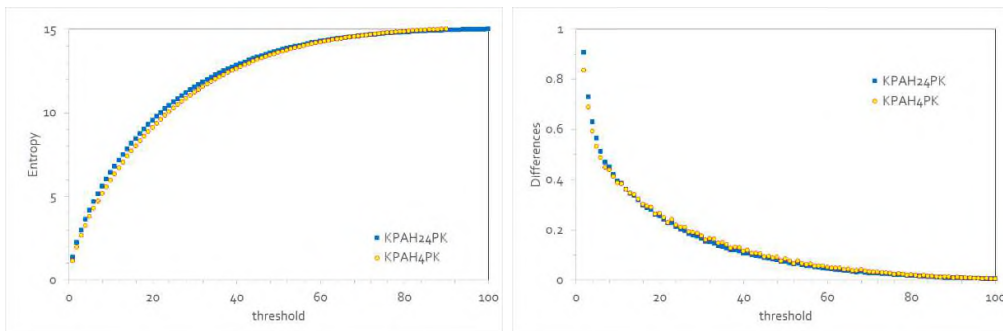


Figure 114 Comparison of Shanonn Entropy (left) and its derivation by threshold differences (right) at 4 h and 24 h.

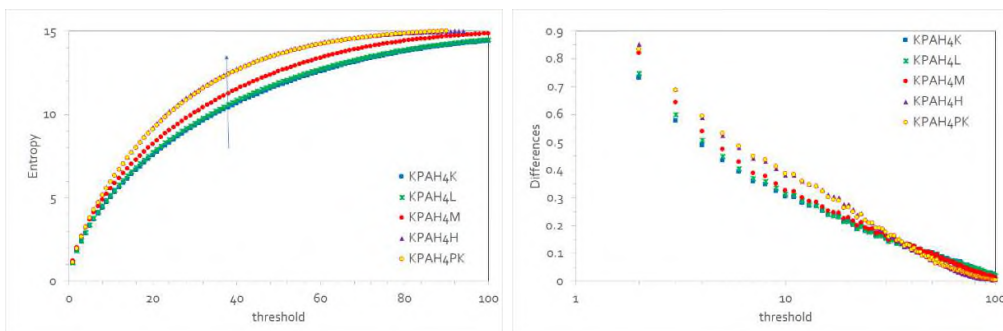


Figure 115 Comparison of Shanonn Entropy (left) and its derivation by threshold differences (right) at 4 h and 24 h.

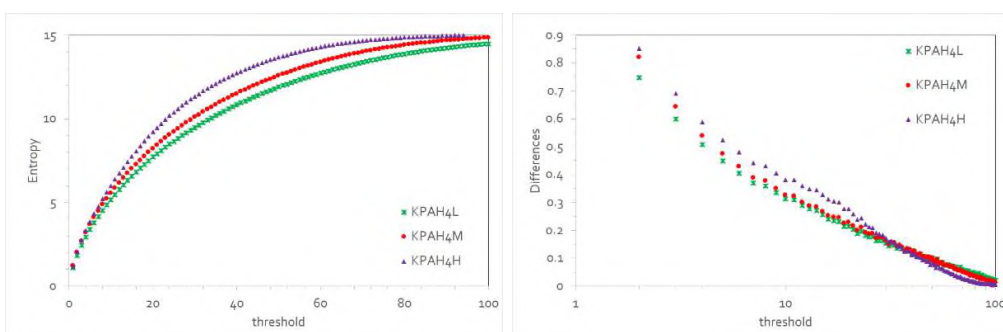


Figure 116 Comparison of Shanonn Entropy (left) and its derivation by threshold differences (right) at 4 h and 24 h.

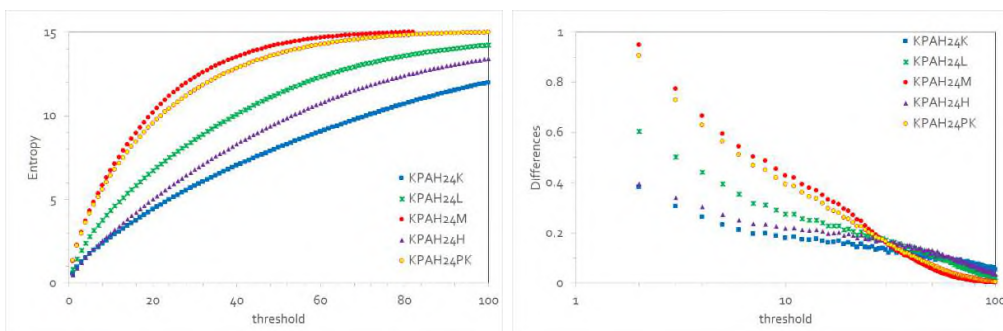


Figure 117 Comparison of Shanonn Entropy (left) and its derivation by threshold differences (right) at 4 h and 24 h.

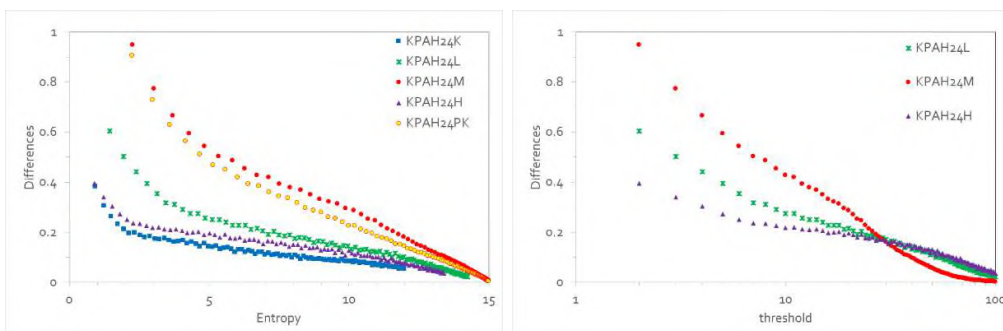


Figure 118 Comparison of Shanonn Entropy (left) and its derivation by threshold differences (right) at 4 h and 24 h.

Most significant threshold have been determined by the 2D cross-correlation analysis (synchronous part) at threshold 35-40 (for 4h) and 25-30 (for 24 h) (Figure 119, Figure 120).

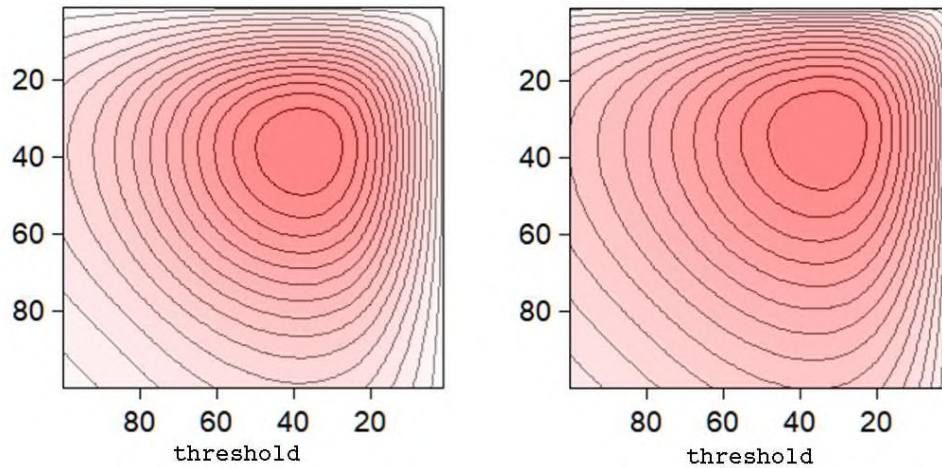


Figure 119 2D cross-correlation analysis (synchronous part) of Shannon entropy for 4h (left) and 24 h (right)

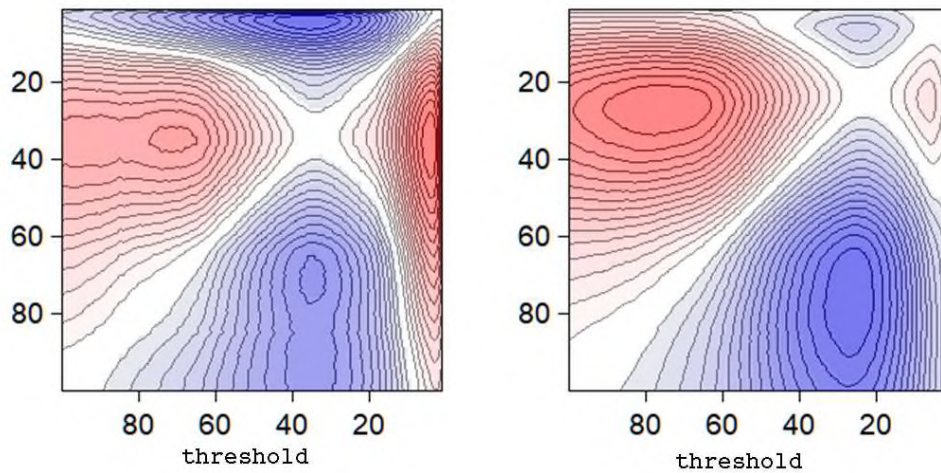


Figure 120 2D cross-correlation analysis (asynchronous part) of Shannon entropy for 4h (left) and 24 h (right)

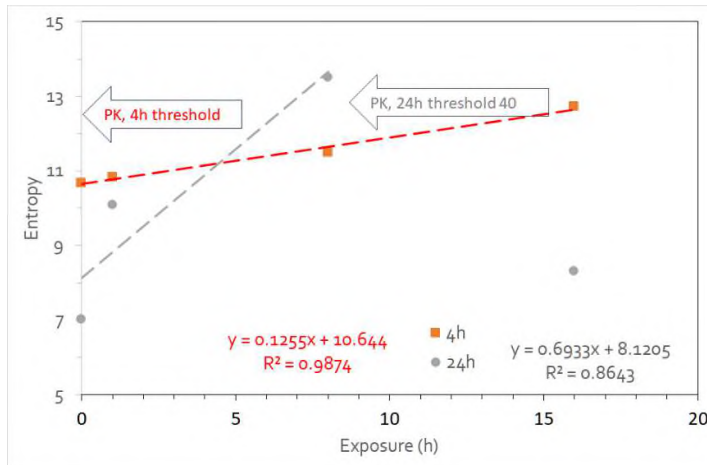


Figure 121 Shannon entropy for 4h (red) and 24 h (gray)

Shannon Entropy of red, green and blue (RGB) colors in Collage Images
Red, green, and blue are the three additive primary colors of light, forming the RGB color model.

Also, Energy is calculated using normalized image histogram, as the sum of squared elements in the normalized histogram.

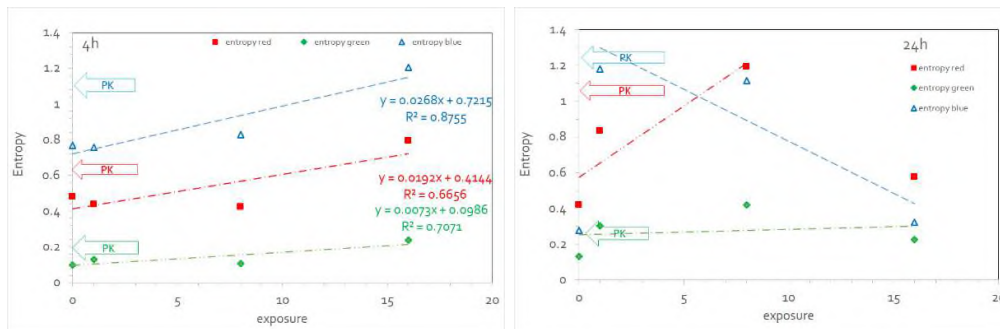


Figure 122 Entropy of red, green and blue (RGB) colors in Collage Images for 4h (left) and 24 h (right)

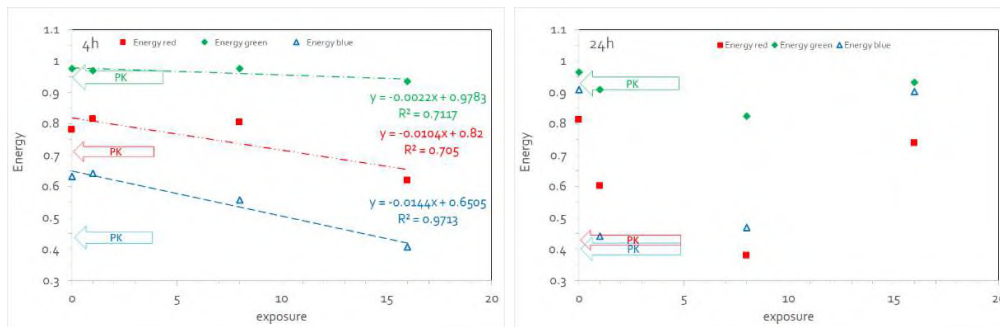


Figure 123 Energy of red, green and blue (RGB) colors in Collage Images for 4h (left) and 24 h (right)

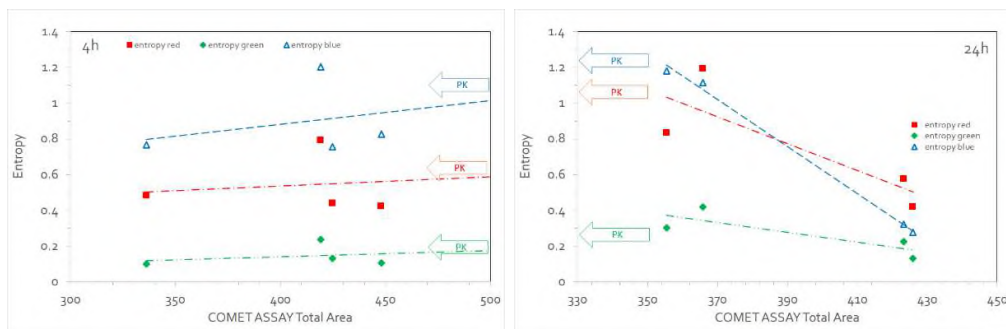


Figure 124 Entropy of red, green and blue (RGB) colors in Collage Images vs. COMET assays Total Area for 4h (left) and 24 h (right)

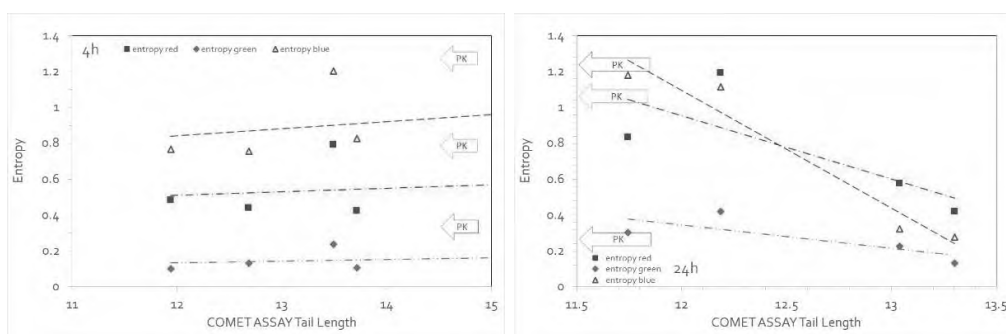


Figure 125 Energy of red, green and blue (RGB) colors in Collage Images vs. COMET assays Tsil Length for 4h (left) and 24 h (right)

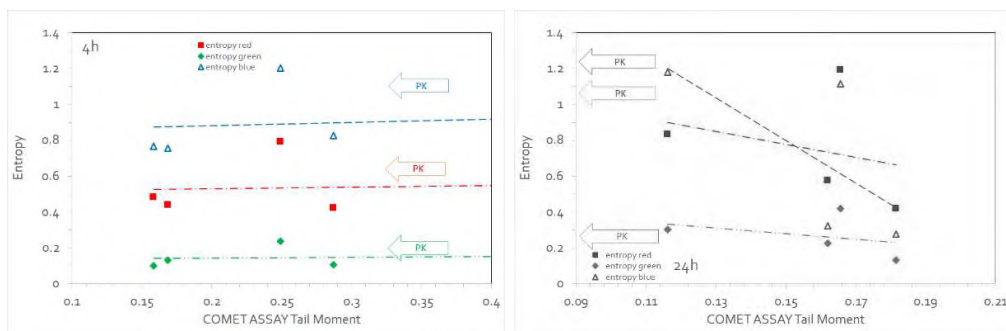


Figure 126 Entropy of red, green and blue (RGB) colors in Collage Images vs. COMET assays Total Moment for 4h (left) and 24 h (right)

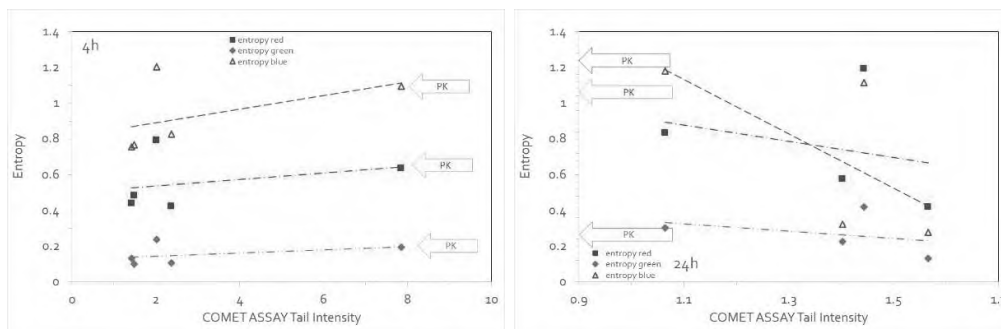


Figure 127 Entropy of red, green and blue (RGB) colors in Collage Images vs. COMET assays Tail Intensity for 4h (left) and 24 h (right)

Image Vectorscope Plot

Image Vectorscope Plot is typically used to measure and adjust the hue and saturation of colors of an image, where indicates chroma and hue are given on the polar coordinate basis. Pixels are plotted in the YCbCr space, derived from the RGB space using the following transformation:

$$\begin{pmatrix} Y \\ C_b \\ C_r \end{pmatrix} = \begin{pmatrix} K_r & K_g & K_b \\ -\frac{K_r}{2(1-K_b)} & -\frac{K_g}{2(1-K_b)} & \frac{1}{2} \\ \frac{1}{2} & -\frac{K_g}{2(1-K_r)} & -\frac{K_b}{2(1-K_r)} \end{pmatrix} \times \begin{pmatrix} r \\ g \\ b \end{pmatrix}$$

Coefficients K_r , K_g , K_b equal the component of the XYZ representation of red, green and blue color primaries, respectively. In the vectorscope plots C_b and C_r are chrominance components. An Image Vectorscope Plot is a visual tool that represents an image's color information (chrominance) as a circular graph, showing color saturation and hue. Each dot in the vectorscope gives information about chrominance. This information has two components: the hue (the specific color) and the saturation (the strength of the hue) of the colors in the image. Pixels with no color are at the center, while saturation increases as they move toward the outer edge. Hue is shown by the color's position around the circle, with target points for primary and secondary colors like red, green, blue, cyan, magenta, and yellow. The chroma strength of the signal is indicated by its distance from the center of the vectorscope. The closer the trace is to the outer edge of the vectorscope, the greater the chrominance or the more vivid the color. The hue of the image is indicated by its rotational position around the circle. The red, blue, and green form a triangle. Between each primary color is the colors formed by mixing those primaries. So, the color between red and blue is magenta. The color between blue and green is cyan, and the color between red and green is yellow. These secondary colors form another triangle.

Image Vectorscope Plots and 3D chromaticity of these plots

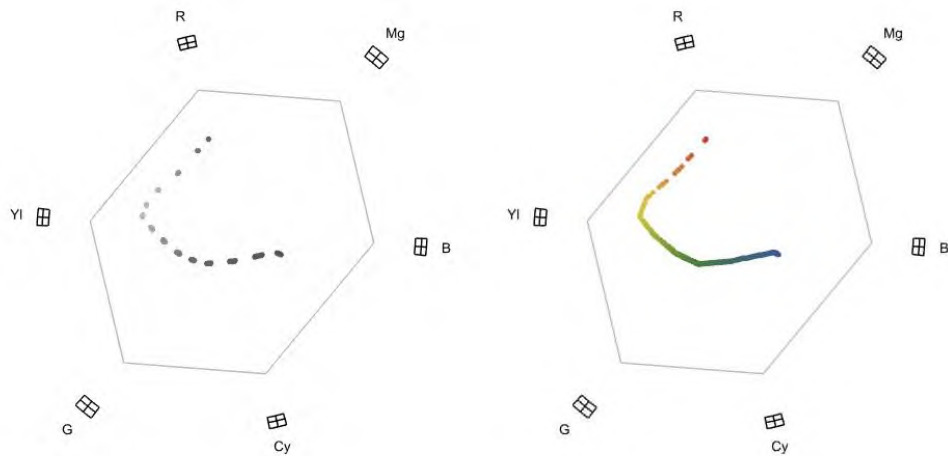


Figure 128 Image Vectorscope Plot of control, K, at 4h (left) and 24h (right)

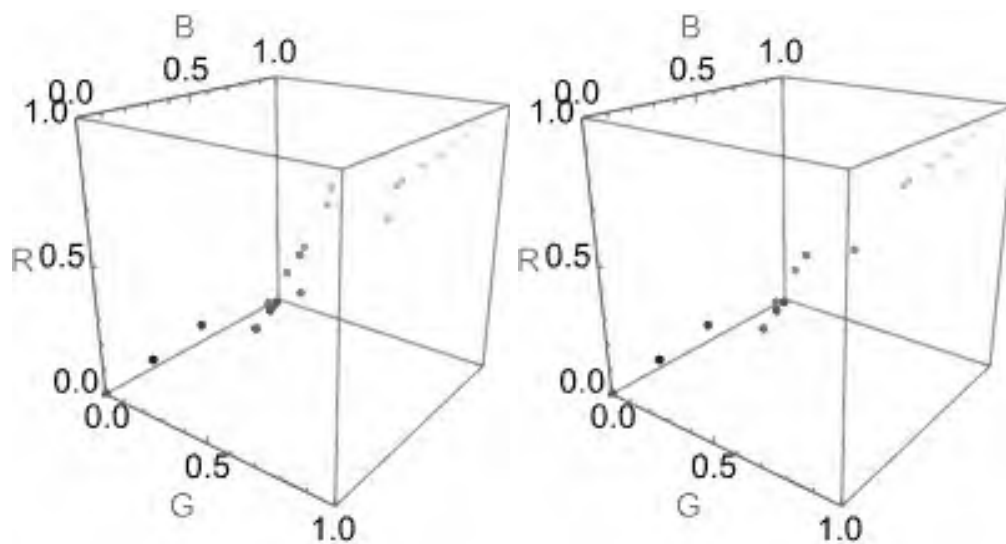


Figure 129 RGB 3D- color distribution of Image Vectorscope Plot of control, K, at 4h (left) and 24h (right)

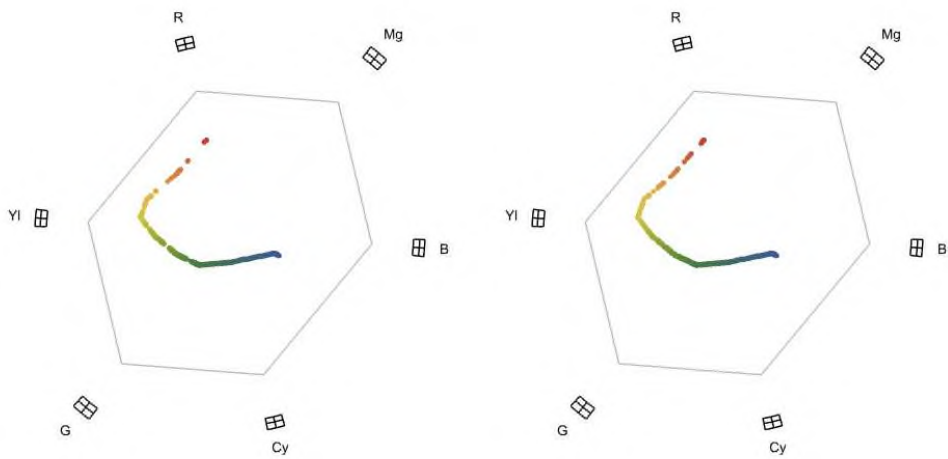


Figure 130 Image Vectorscope Plot of L-exposure, at 4h (left) and 24h (right)

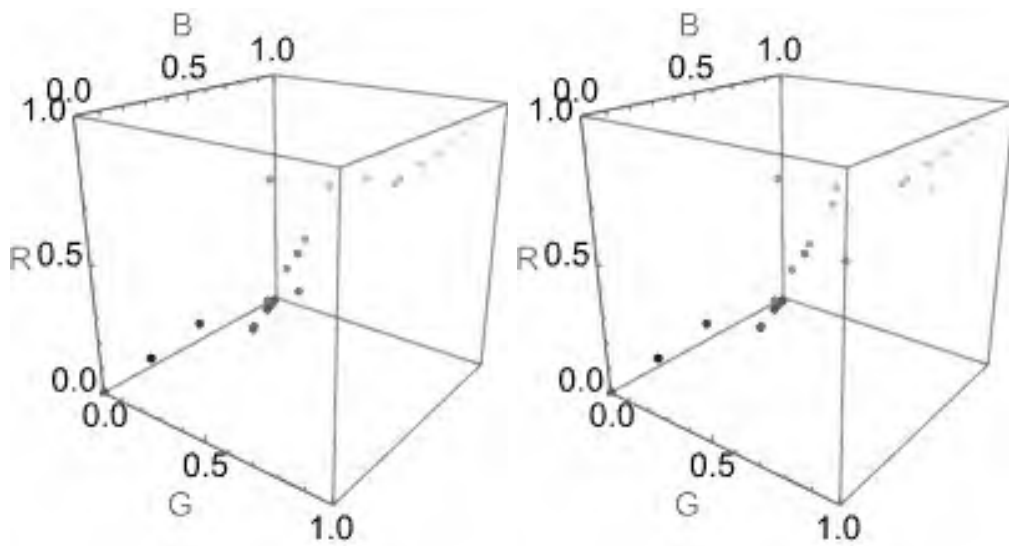


Figure 131 RGB 3D- color distribution of Image Vectorscope Plot of L-exposure, at 4h (left) and 24h (right)

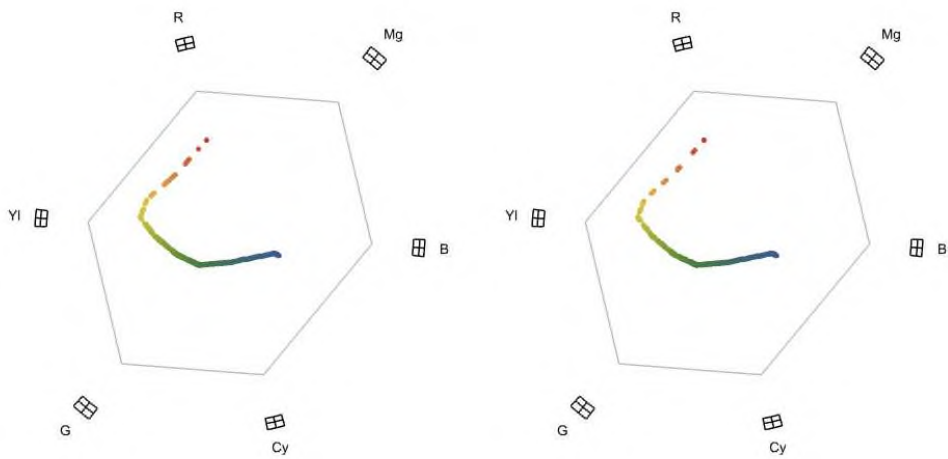


Figure 132 Image Vectorscope Plot of M-exposure, at 4h (left) and 24h (right)

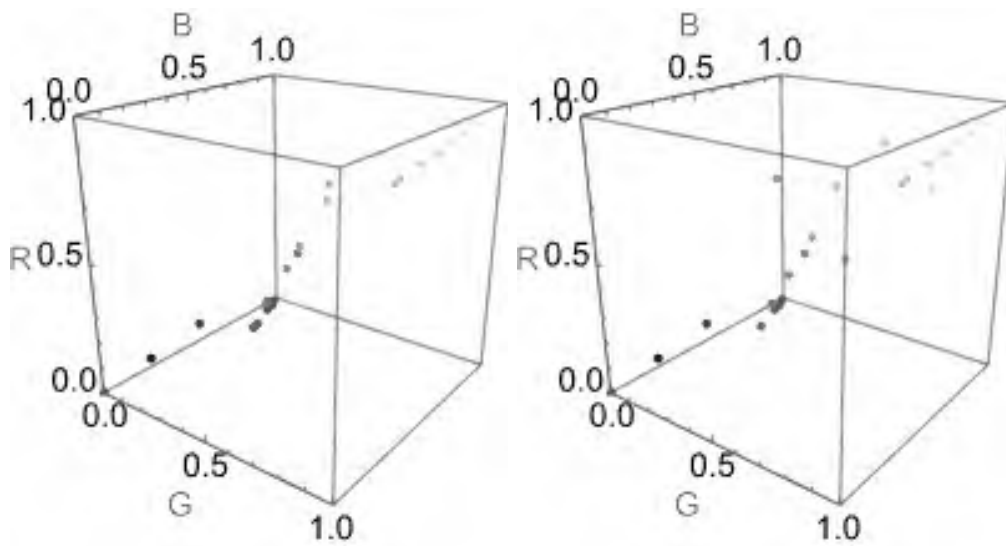


Figure 133 RGB 3D- color distribution of Image Vectorscope Plot of M-exposure, at 4h (left) and 24h (right)

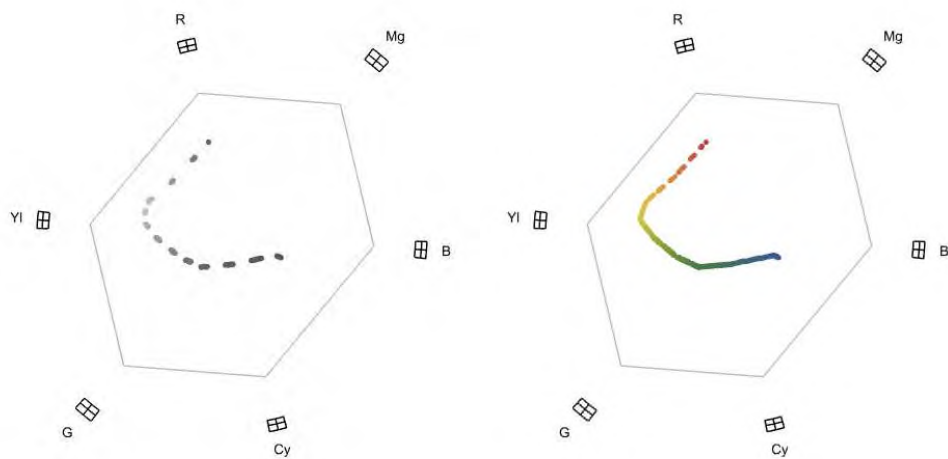


Figure 134 Image Vectorscope Plot of H-exposure, at 4h (left) and 24h (right)

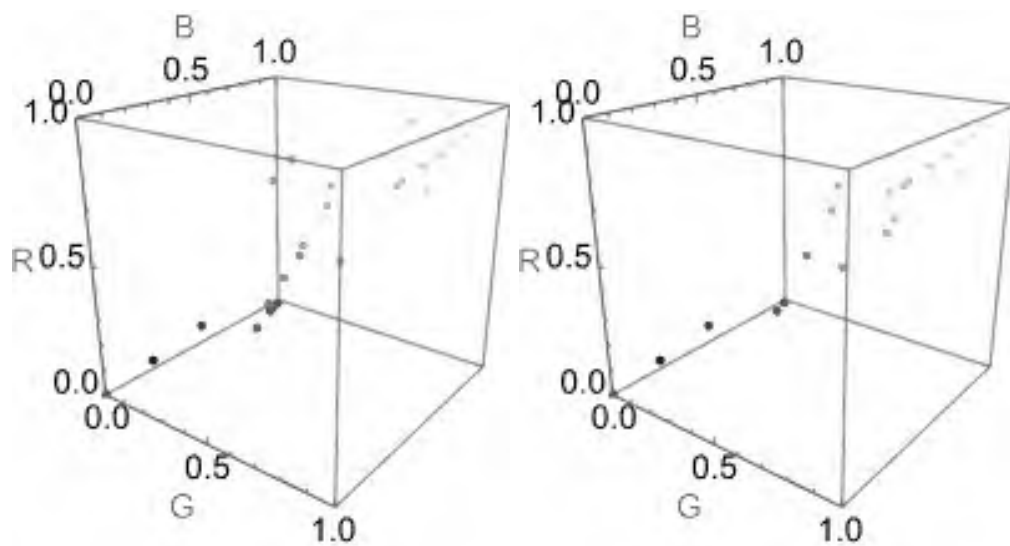


Figure 135 RGB 3D- color distribution of Image Vectorscope Plot of H-exposure, at 4h (left) and 24h (right)

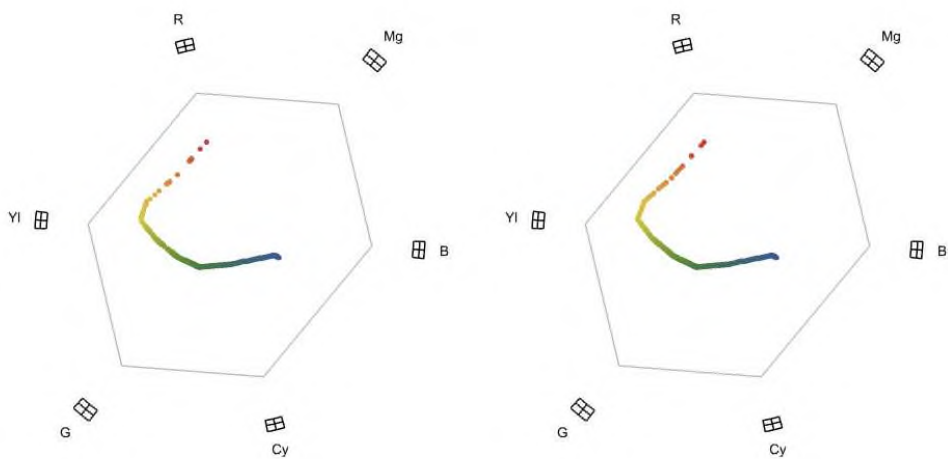


Figure 136 Image Vectorscope Plot of positive control, PK, at 4h (left) and 24h (right)

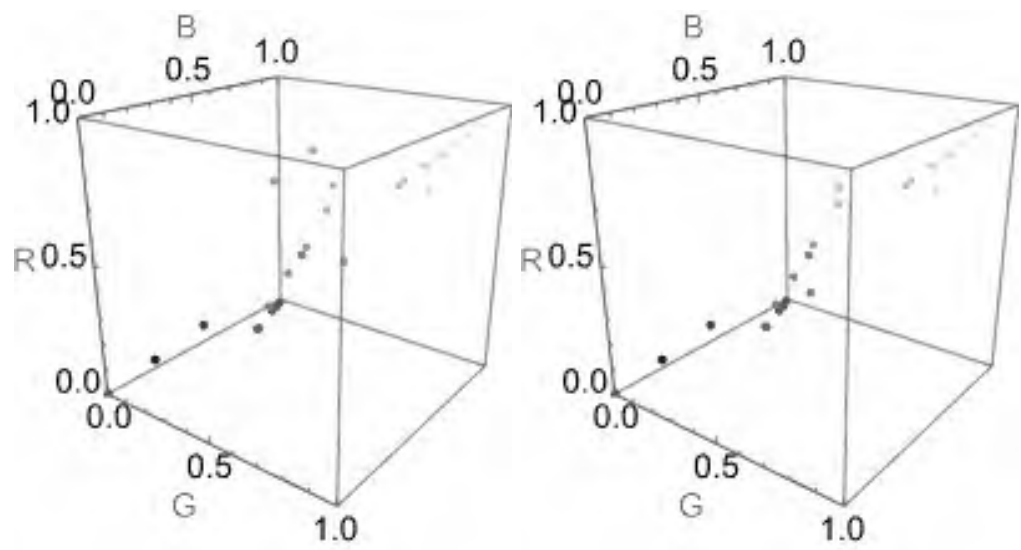


Figure 137 RGB 3D- color distribution of Image Vectorscope Plot of positive control, PK, at 4h (left) and 24h (right)

Graph morphology of comet assays images

In Comet assays imaging, the goal of image analysis is the extraction of Comet assays parts. This problem is a difficult problem due to the large variations of the features present in such structures. Comet assays analysis is an application dependent task and no general scheme can be given. So, we propose a set of graph-based tools to address Comet assays extraction problems. These tools are a framework. Within this framework, many operations can be performed, combined or derived to produce a specific scheme for a problem. One strong specificity of the proposed framework is to use graphs as a discrete modelling of images at different levels (pixels) and different component relationships (distance graph, etc.). Working on graphs of arbitrary structure, our framework leads to a set of flexible tools for Comet assays analysis. To provide a fast scheme, we also propose a discrete image simplification inspired by an approach based on the generalized Voronoi diagram⁵⁹. This simplification scheme can be viewed as an image/graph reduction. The main purpose of this paper is not to solve a particular class of cytology or histology problems but to show how, with our graph-based tools, we can provide particular schemes to address several classes of problems in microscopic images segmentation.

Comet assays images are inherently complex and have a variety of visual information. Several image analysis techniques are being explored in this direction. However, graph-based methods^{60,61} are gaining most popularity, as these methods can describe tissue architecture and provide adequate numeric information for subsequent computer-based analysis. Graphs can represent spatial arrangements and neighbourhood relationships of different tissue components, which are essential features observed visually by pathologists during investigation of specimens. The representation of some system by means of graphs allow us to analyse complex structures, both their structure and properties, and summarize them in terms of some numerical parameters called (topological) indices. These indices, which are graph invariants, carry information on the structure of the graph.

The ideas of “classical morphology” can be extended to spaces of images modelled by graphs^{62,63,64,65}. Mathematical morphology^{66,67} is a tool for extracting image components that are useful for representation and description. A set-theoretic method of image analysis is developed that quantitatively describes geometrical structures. Morphology can provide boundaries of objects, their skeletons, and many preprocessing and postprocessing techniques, especially in edge thinning and pruning. Generally, most morphological operations are based on simple expanding and shrinking operations. A graph includes a collection of points, called vertices, and a binary relation between them: two vertices either are related or they are not related. A graph $G = (V, E)$ includes V , a nonempty set of vertices (or nodes), and E , a set of edges. Each edge has two vertices associated with it, known as its endpoints, and is said to connect these endpoints. The primary objective of

graph theory is to identify topological indices that closely correlate with the properties. The structuring graph is used to construct a neighbourhood function on the vertices by relating individual vertices to each other whenever they belong to a local instantiation of the structuring graph. Graph morphology provides nice tools for the study of heterogeneous media at a macroscopic level, based on information on their microstructure. The basic idea underlying classical morphology is to extract information from an image by probing it at any position with some small geometric shape called a structuring element. Graph morphology provides a collection of morphological tools for the investigation of populations of objects for which neighbourhood relations are of interest. Recently, graph theory has been used to characterize the spatial arrangement of the cells by constructing a graph with cell/nuclei as the node^{68, 69, 70, 71, 72, 73, 74, 75}.

The high discriminative capability of the graph node centrality feature played a significant role in the experiments⁷⁶.

Morphological Graph

A graph $G=(V,E)$ includes V , a nonempty set of vertices (or vertices), and E , a set of edges⁷⁷. Each edge has two vertices associated with it, known as its endpoints, and is said to connect these endpoints. In chemical graph theory, a chemical graph represents individual atoms as vertices and the bonds between them as edges. The main goal of chemical graph theory is to identify topological indices that closely correlate with the properties of chemical compounds^{78,79}.

A graph is a structure that represents components and their interactions. Morphological Graph gives a graph object that represents the connectivity of the morphological branch points and endpoints of the objects in image after applying morphological thinning^{80,81}. The morphological branch points and endpoints of the image skeleton are converted into vertices. The graph shows an undirected edge between two vertices if the corresponding branch points or endpoints are 8-connected by the skeleton.

The image is transformed into a graph using Morphological Graph function.

Results

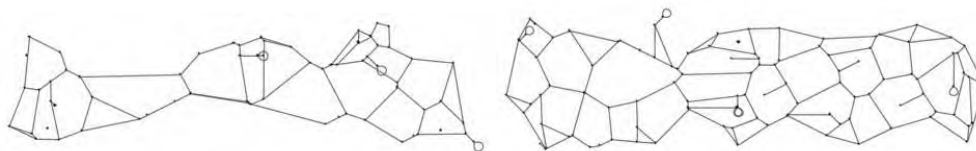


Figure 138 Morphological Graph of control (K): 4h (left) and 24h (right)

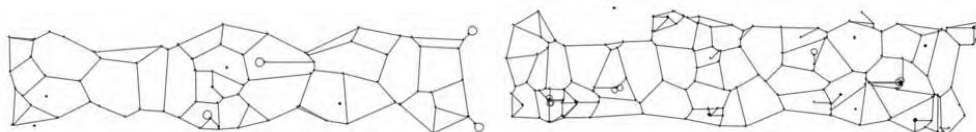


Figure 139 Morphological Graph of low exposure (L): 4h (left) and 24h (right)

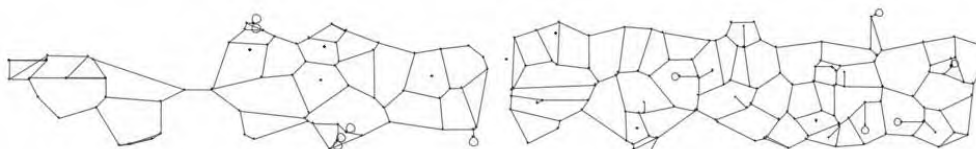


Figure 140 Morphological Graph of middle exposure (M): 4h (left) and 24h (right)

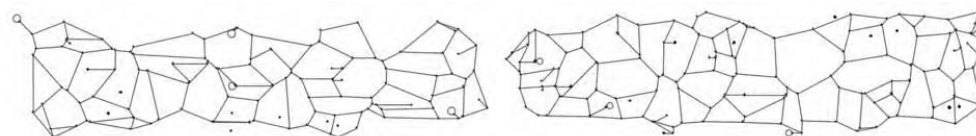


Figure 141 Morphological Graph of high exposure (H): 4h (left) and 24h (right)

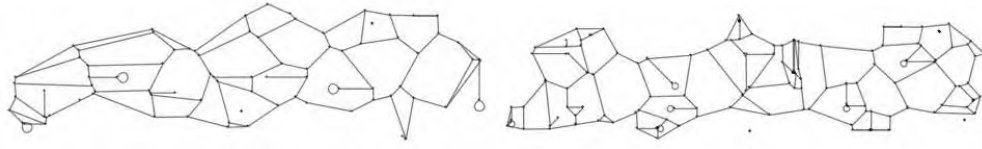


Figure 142 Morphological Graph of positive control (PK): 4h (left) and 24h (right)

Correlation dimension

Correlation dimension is the measure of dimensionality of the space occupied by a set of random points^{82,83}. Correlation dimension is estimated by modelling as a power law the fraction of vertices within a distance. Correlation dimension is estimated as the slope of the correlation integral *versus* the range of radius of similarity. The correlation dimension was initially applied to describe geometric object. Correlation dimension can be a typical measure to distinguish between deterministic chaos and random noise.

The correlation integral as:

$$C(\epsilon) \equiv \lim_{N \rightarrow \infty} \frac{1}{N^2} H(\epsilon - |x_i - x_j|)$$

where H is the Heaviside step function. When the below limit exists, the correlation dimension is then defined as:

$$D_2 \equiv d_{corr} \equiv \lim_{\epsilon, \epsilon' \rightarrow 0} \frac{\ln \left[\frac{C(\epsilon')}{C(\epsilon)} \right]}{\ln \left(\frac{\epsilon'}{\epsilon} \right)}$$

where D_2 is the correlation dimension.

Results

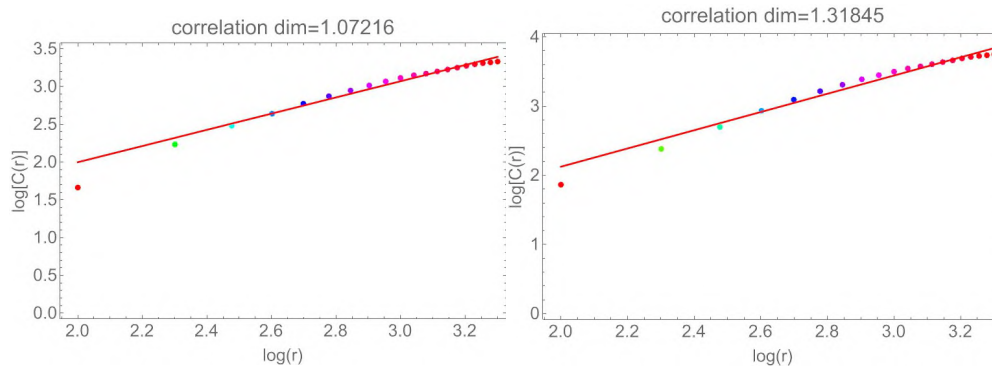


Figure 143 Correlation dimension fit of control (K): 4h (left) and 24h (right)

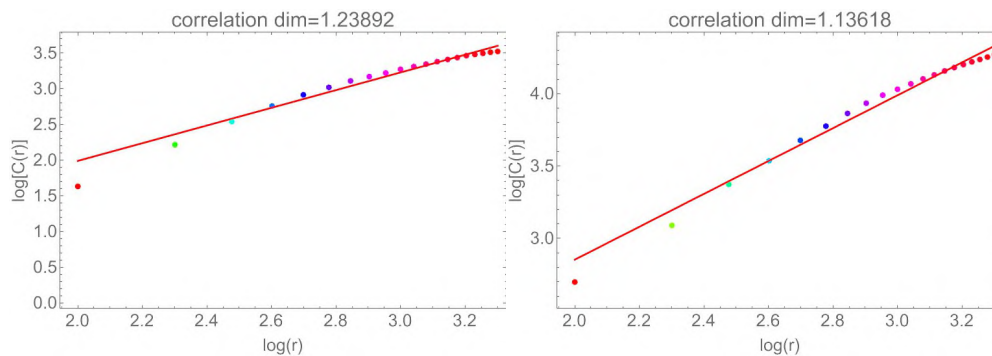


Figure 144 Correlation dimension fit of low exposure (L): 4h (left) and 24h (right)

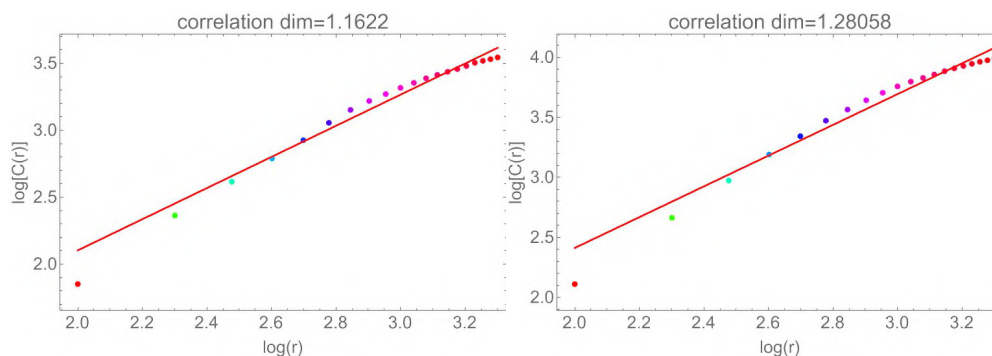


Figure 145 Correlation dimension fit of middle exposure (M): 4h (left) and 24h (right)

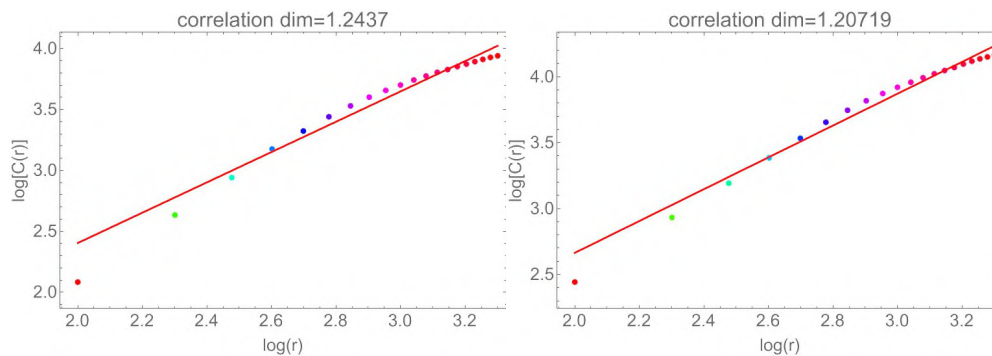


Figure 146 Correlation dimension fit of high exposure (H): 4h (left) and 24h (right)

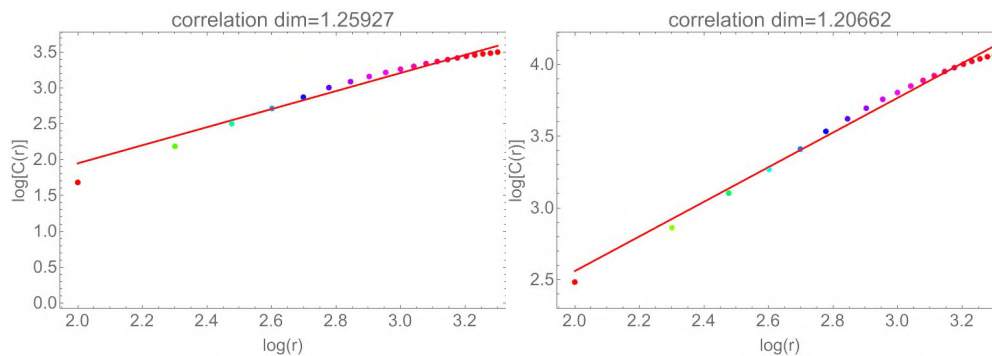


Figure 147 Correlation dimension fit of positive control (PK): 4h (left) and 24h (right)

Table 11 Correlation dimension fit parameters

	exposure	correlation dimension	Standard Error	t-Statistic	P-Value
KPAH4K	0	1.0722	0.0339	31.6083	1.53E-16
KPAH4L	1	1.2389	0.0489	25.3100	6.19E-15
KPAH4M	8	1.1622	0.0409	28.4395	8.91E-16
KPAH4H	16	1.2437	0.0526	23.6392	1.92E-14
KPAH4PK		1.2593	0.0526	23.9240	1.57E-14
KPAH24K	0	1.3184	0.0523	25.1897	6.70E-15
KPAH24L	1	1.1362	0.0376	30.2147	3.24E-16
KPAH24M	8	1.2806	0.0524	24.4451	1.10E-14
KPAH24H	16	1.2072	0.0408	29.6092	4.55E-16
KPAH24PK		1.2066	0.0313	38.5782	5.39E-18

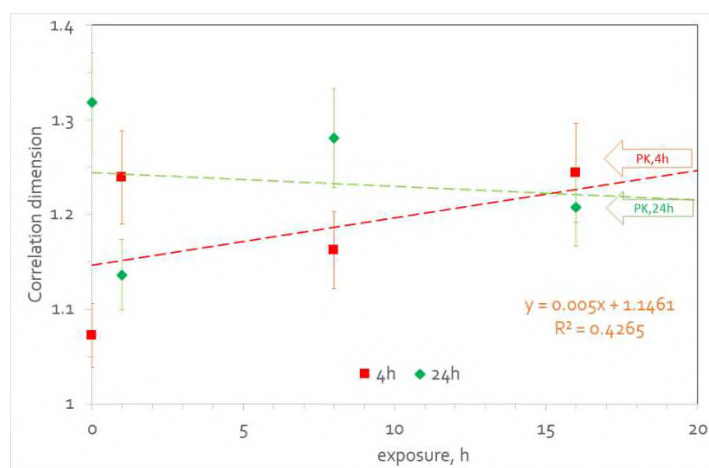


Figure 148 Correlation dimensions at different exposures at 4h (red) and 24h (green).

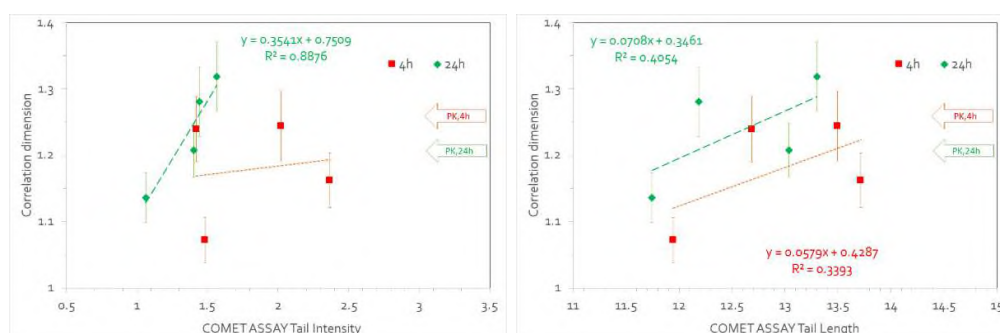


Figure 149 Correlations of Correlation dimensions and **Comet Assay Tail Intensity** (left) and Correlation dimensions and **Comet Assay Tail Length** (right) at 4h (red, ■) and 24h (green, ◆)

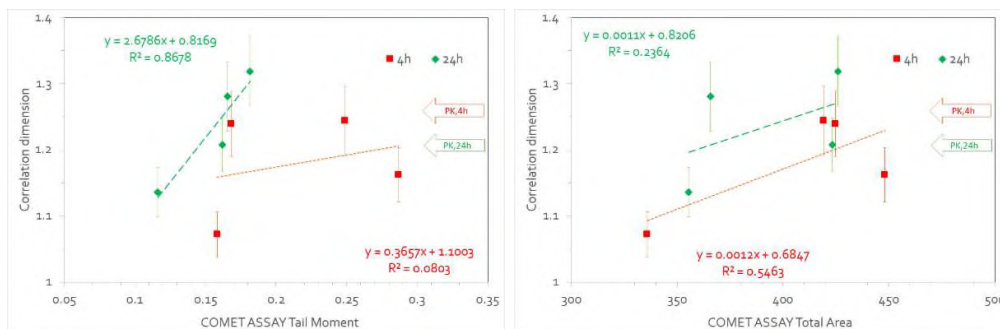


Figure 150 Correlations of Correlation dimensions and **Comet Assay Tail Moment** (left) and Correlation dimensions and **Comet Assay Total Area** (right) at 4h (red, ■) and 24h (green, ◆)

Graph Density

Graph Density is the ratio of edges divided by the number of edges of a complete graph with the same number of vertices. Graph Density measures the density of the Adjacency Matrix. The graph density is between 0 and 1. The graph density is the ratio of its edges to the maximum number of edges and shows how sparse or dense a graph is according to the number of connections per node set and is defined as:

$$\text{Graph density} = \frac{2 \times \text{Edge count}}{(\text{Vertex count})^2 - \text{Vertex count}}$$

Results

Table 12 Graph Density and Shannon entropy (or base 2 entropy) of Graph Distance Matrix

	exposure	Graph Density	Shannon entropy (or base 2 entropy) of Graph Distance Matrix
KPAH4K	0	0.0333	6.2479
KPAH4L	1	0.0295	6.4919
KPAH4M	8	0.0283	6.5078
KPAH4H	16	0.0173	7.1699
KPAH4PK		0.0306	6.4263
KPAH24K	0	0.0231	6.8580
KPAH24L	1	0.0113	7.7944
KPAH24M	8	0.0172	7.2479
KPAH24H	16	0.0128	7.6073
KPAH24PK		0.0148	7.4009

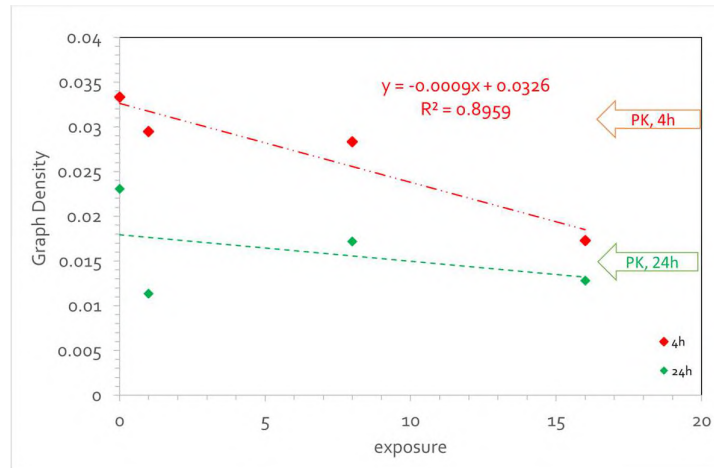


Figure 151 Graph Density at 4h (red) and 24h (green).

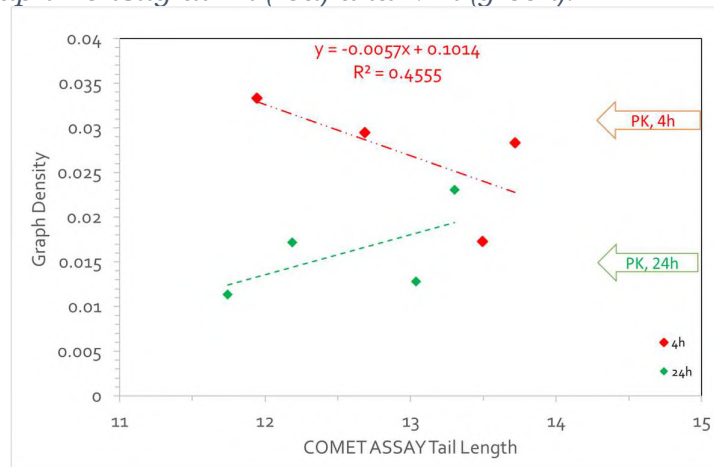


Figure 152 Correlation of Graph Density vs. COMET ASSAY Tail Length at 4h (red) and 24h (green).

Graph Distance Matrix, entropy

Graph Distance Matrix⁸⁴ returns an ordinary matrix of distances between all pairs of vertices in the graph H . The distance matrix of a graph is defined similarly as the adjacency matrix: the entry in the i -th row, j -th column is the distance (length of a shortest path) between the i -th and j -th vertex. The entries of the distance matrix d_{ij} give the shortest distance from vertex v_i to vertex v_j . The diagonal entries d_{ii} of the distance matrix are always zero. The entry d_{ij} is infinity (∞) if there is no path from vertex v_i to vertex v_j .

Distance matrix is a square matrix (two-dimensional array) containing the distances, taken pairwise, between the elements of a set. A graph's distance matrix is created^{85,86,87}, where each entry (i,j) represents the shortest path length (geodesic distance) between vertex i and vertex j .

The entropy is then estimated for this matrix, which measures the uncertainty or “disorder” of the distances it has.

Graph Distance Matrix entropy^{88,89,90} is a measure of disorder within a graph's distance matrix. It is calculated by estimating the entropy⁹¹ for the graph's distance matrix, which has the shortest path distances between all pairs of vertices in a graph.

The Shannon entropy is given by $H = -\sum p_i \log(p_i)$, where the probability of each group p_i (e.g., the number of vertices in a group divided by the number of vertices).

Analysing these entropy values can reveal patterns (Table 3, Figure 23). A lower entropy value can show a more ordered structure in the graph's distances. A higher entropy value can show more disorder or uncertainty. The entropy measures based on distances were proposed for interpreting the molecular branching of molecular graphs⁹².

Results

Graph Distance Matrix

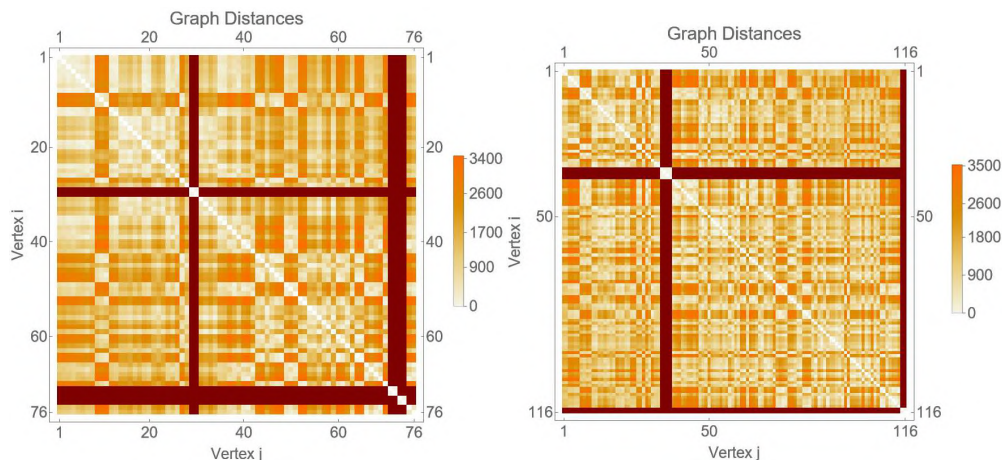


Figure 153 Graph Distance Matrix of control (K): 4h (left) and 24h (right)

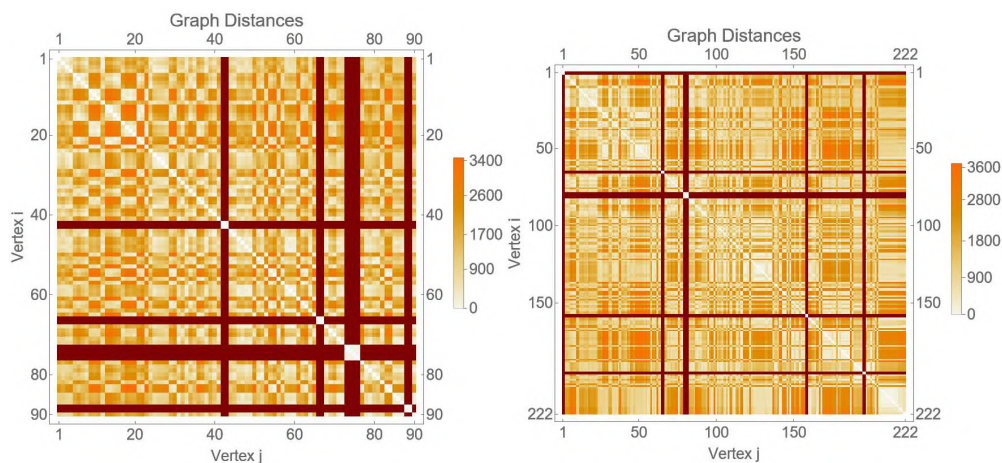


Figure 154 Graph Distance Matrix of low exposure (L): 4h (left) and 24h (right)

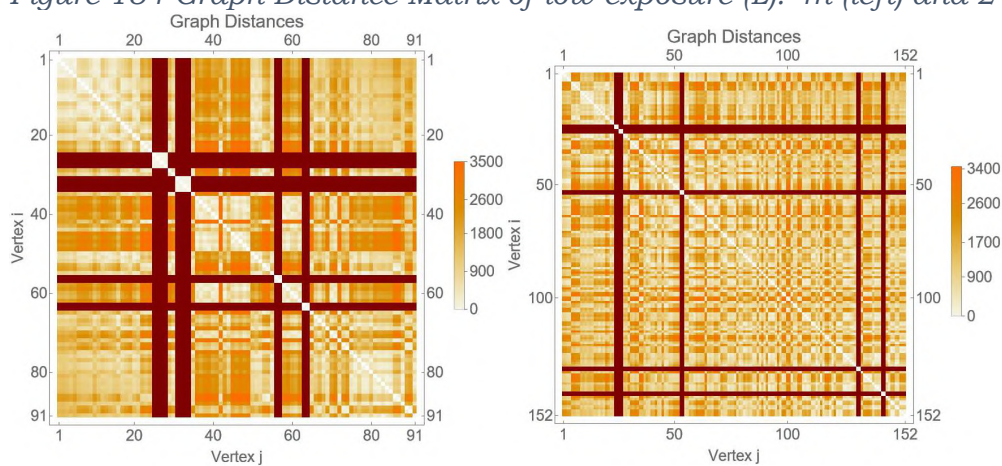


Figure 155 Graph Distance Matrix of middle exposure (M): 4h (left) and 24h (right)

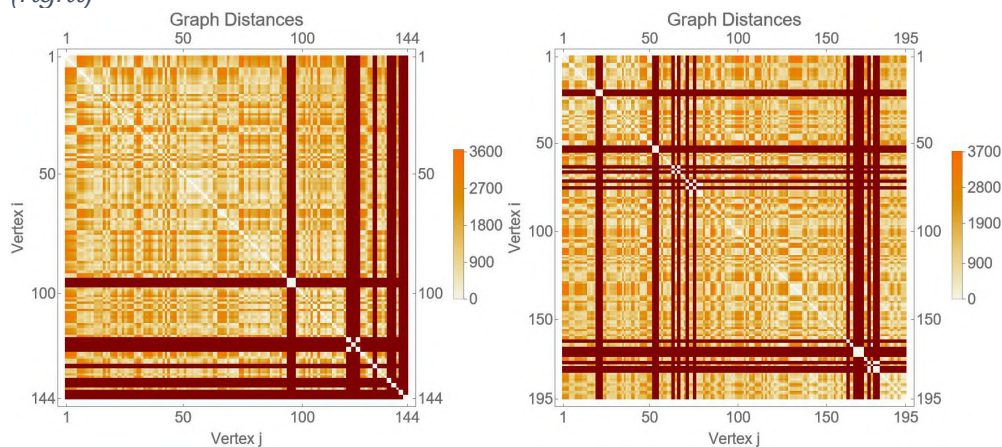


Figure 156 Graph Distance Matrix of high exposure (H): 4h (left) and 24h (right)

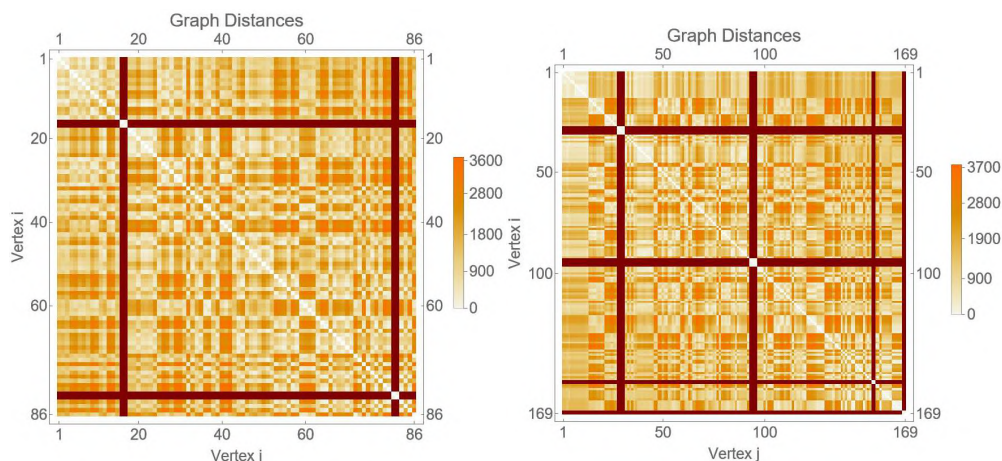


Figure 157 Graph Distance Matrix of positive control (PK): 4h (left) and 24h (right)

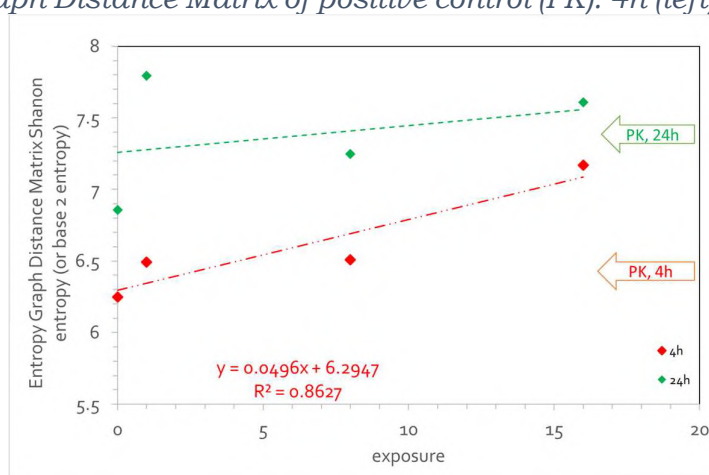


Figure 158 Entropy Graph Distance Matrix or Shannon entropy (or base 2 entropy) at 4h (red) and 24h (green).

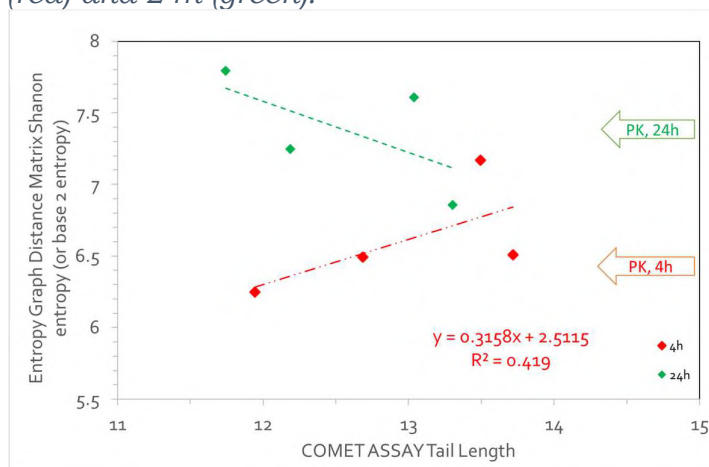


Figure 159 Correlation of Entropy Graph Distance Matrix or Shannon entropy (or base 2 entropy) vs. COMET ASSAY Tail Length at 4h (red) and 24h (green).

Topological measures, Centrality measures

Evaluating the graph is using various topological measures. These measures include degree, betweenness centrality, closeness centrality, and clustering coefficient. The degree of a node refers to the number of edges linked to a node in the graph. Vertices with high degrees are known as graph *hubs*, suggesting their many connections to other vertices. Hubs play important roles in information flow and are important in the graph. Betweenness centrality measures how often a node serves as a bridge or lies on the shortest paths connecting pairs of other vertices in the graph. Vertices with high betweenness centrality act as bottlenecks, controlling the flow of information between different parts of the graph. Closeness centrality can provide insights into the information flow and accessibility of vertices within the graph. The clustering coefficient determines how closely connected a vertices neighbour are to one another. A high clustering coefficient indicates that the neighbours of a vertex are well connected.

Centrality metrics have become a popular concept in graph science. Over the years, centrality has been used to assign importance and identify influential elements in various settings. Centrality metrics refer to indicators identifying the varying importance of vertices in complex graphs⁹³. However, most of the literature has focused on nodal versions of centrality. The centrality of a node is often used to measure its importance within a graph's structure. Certain centrality measures can be extended to evaluate the importance of a group of vertices that share a common property or form a specific structure, such as a path.

The Centrality measures^{94,95,96,97,98} are used to find the most important vertices in a graph. Centrality metrics include Degree Centrality, Closeness Centrality, Betweenness Centrality, Eigenvector Centrality, Katz Centrality, and the local clustering coefficient-dependent degree centrality.

By analysing the centrality of different parts of the Comet assay's graph, we can pinpoint exposures that are critical for a Comet assays morphology. We have modeled the graph structure form Comet assays images and then calculated various topological indices for these graphs^{99,100}. This approach allows for the quantitative analysis of the properties of the Comet assays images structure, potentially helping to understand the impact of exposures¹⁰¹.

Centralization^{102,103} is a general method for calculating a graph-level centrality score based on node-level centrality measure. The formula for this is

$$C(G) = \left(\sum_v \max_w c_w - c_v \right)$$

where c_v is the centrality of vertex v .

The graph-level centralization measure can be normalized by dividing by the maximum theoretical score for a graph with the same number of vertices, using the same parameters, e.g. directedness, whether we consider loop edges, etc.

For degree, closeness and betweenness the most centralized structure is some version of the star graph, in-star, out-star or undirected star.

For eigenvector centrality the most centralized structure is the graph with a single edge (and potentially many isolates).

This study focuses on deriving relationships between the topological indices and the structural features of the Comet assays images model.

The most used concept is centrality¹⁰⁴. This measurement counts direct connections. However, this approach is not always effective, as many direct connections may be inactive. Another important centrality measure is betweenness centrality¹⁰⁵. In this concept, several shortest paths between any pair of vertices through the specified node is considered for measuring betweenness centrality. In this technique, the strength of the path is not analysed. Other centrality measurements in the literature include closeness centrality¹⁰⁶ and eigenvector centrality¹⁰⁷.

Closeness centrality is an idea ¹⁰⁸ to capture how close in terms of shortest path distance two vertices are in a graph.

Empirical global clustering coefficient

In graph theory, a clustering coefficient is a measure of how much vertices in a graph usually cluster together¹⁰⁹.

Goal in graph data analysis is to characterize the structural topology of a graph. Various summary metrics have been formulated to measure the graph topology. These metrics help with the comparative analysis of graphs and their categorization according to salient structural features. The global clustering coefficient was designed to indicate the clustering in the graph. The global clustering coefficient is among the most often applied graph statistics^{110, 111, 112}. It provides insights into the local link density and the tendency of vertices to cluster together in a graph. It is defined as the number of closed triplets over the number of triplets. The global clustering coefficient measures the extent to which vertices in a graph usually cluster together. The empirical global clustering coefficient¹¹³ is a measure of how clustered a graph is, calculated as the ratio of “closed triplets” (three vertices connected by two or three edges, forming triangles) to “total triplets” (all combinations of three vertices with two or three edges). It provides an overall picture of the graph's structure, where a value near 1 indicates a highly clustered graph and a value near 0 indicates a sparse one.

Results

Table 13 The empirical and the theoretical global clustering coefficient

	exposure	The empirical global clustering coefficient	The theoretical global clustering coefficient
KPAH4K	0	0.1104	0.0163
KPAH4L	1	0.0694	0.0142
KPAH4M	8	0.0865	0.0144
KPAH4H	16	0.0478	0.0088
KPAH4PK		0.0147	0.0154
KPAH24K	0	0.0941	0.0115
KPAH24L	1	0.0366	0.0057
KPAH24M	8	0.0252	0.0085
KPAH24H	16	0.0487	0.0067
KPAH24PK		0.0561	0.0079

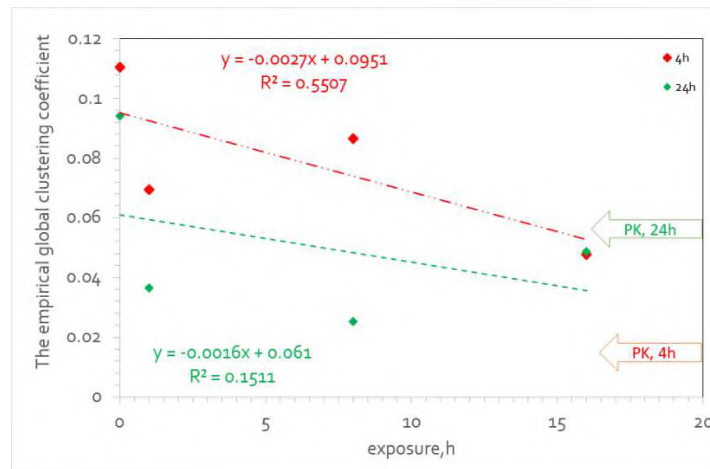


Figure 160 The Empirical global clustering coefficient at 4h (red) and 24h (green).

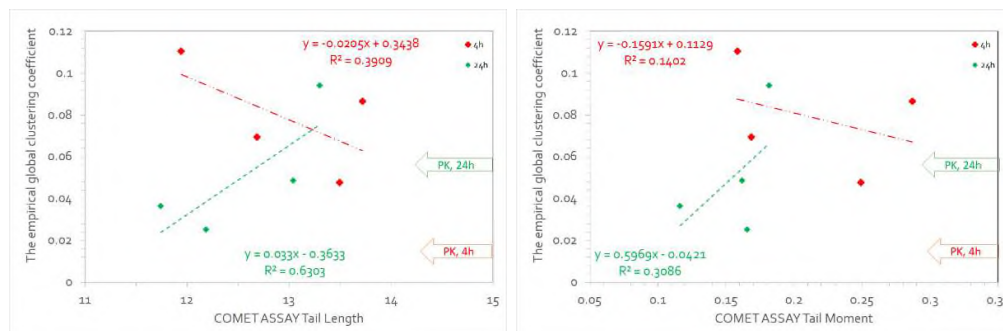


Figure 161 The Empirical global clustering coefficient versus COMET ASSAY Tail Length (left) and versus COMET ASSAY Tail Moment (right) at 4h (red) and 24h (green).

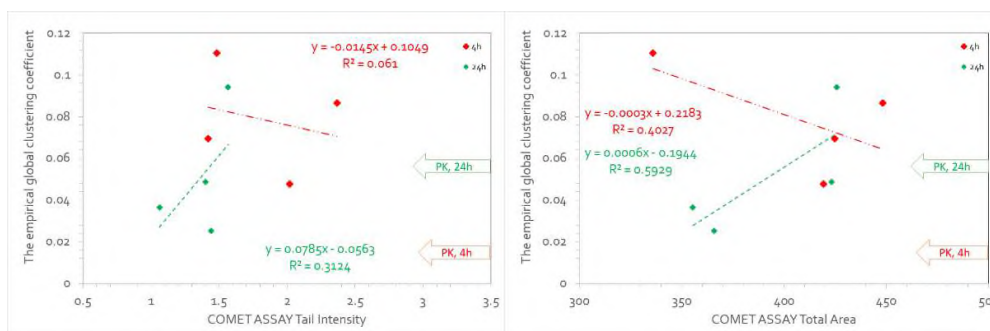


Figure 162 The Empirical global clustering coefficient versus COMET ASSAY Tail Intensity (left) and versus COMET ASSAY Total Area (right) at 4h (red) and 24h (green).

Theoretical global clustering coefficient

The theoretical global clustering coefficient is the proportion of closed triplets to the number of triplets in a graph, representing the overall tendency for vertices to cluster together. It is calculated as the number of closed triplets (or triangles) divided by the number of triplets (both open and closed). This value ranges from 0 (no clustering) to 1 (maximum clustering, where every node is part of a triangle).

Results

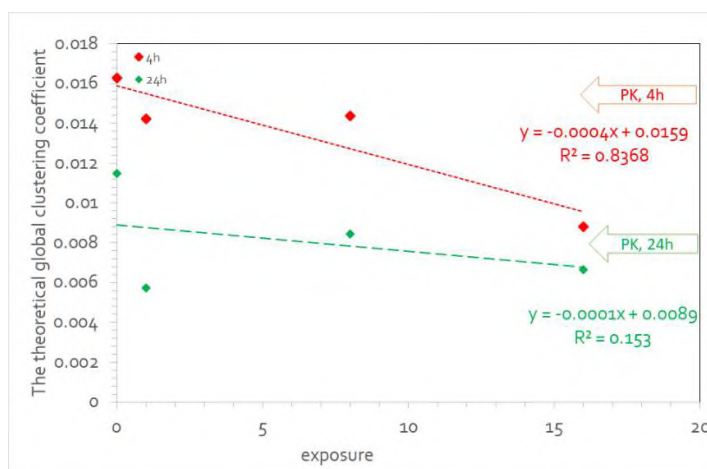


Figure 163 The Theoretical global clustering coefficient at 4h (red) and 24h (green). (see Table 13)

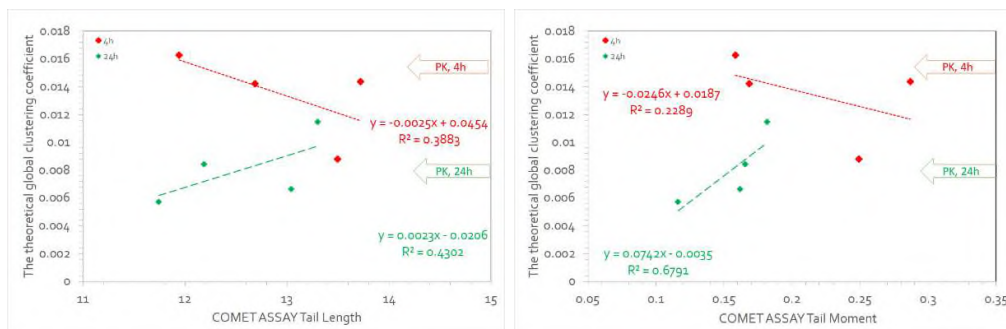


Figure 164 The Theoretical global clustering coefficient versus COMET ASSAY Tail Length (left) and versus COMET ASSAY Tail Moment (right) at 4h (red) and 24h (green). (see Table 13)

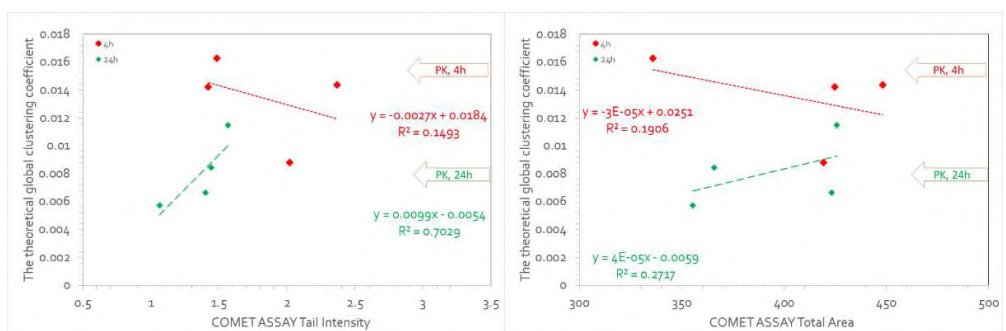


Figure 165 The Theoretical global clustering coefficient versus COMET ASSAY Tail Intensity (left) and versus COMET ASSAY Total Area (right) at 4h (red) and 24h (green). (see Table 13)

Degree-based Centrality Metrics

Centrality measures analyse the structure of a graph, such as how many direct connections a node has (degree centrality), or how often it is on the shortest path between other vertices (betweenness centrality).

Degree centrality: Measures the number of direct connections a node has or measure of vertices next to a vertex (degree)

Results

Table 14 Degree Centrality

	exposure	Degree Centrality
KPAH4K	0	0.0068
KPAH4L	1	0.0043
KPAH4M	8	0.0051
KPAH4H	16	0.0037
KPAH4PK		0.0048
KPAH24K	0	0.0119
KPAH24L	1	0.0068
KPAH24M	8	0.0027
KPAH24H	16	0.0079
KPAH24PK		0.0091

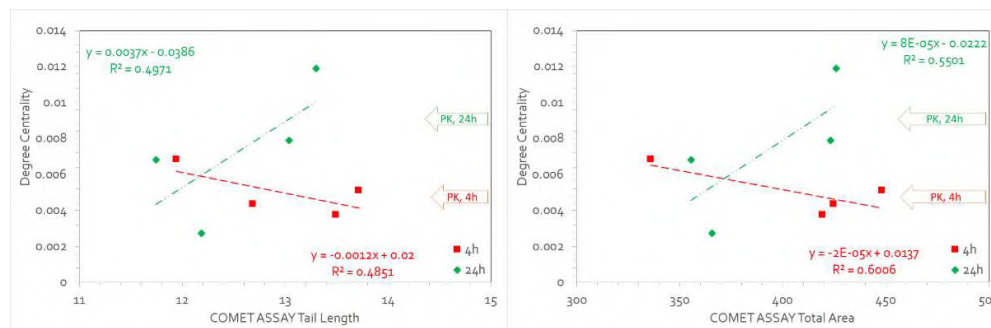


Figure 166 Degree Centrality versus COMET ASSAY Tail Length (left) and versus COMET ASSAY Total Area (right) at 4h (red, ■) and 24h (green, ◆)

Betweenness centrality

Betweenness centrality measures how often a node lies on the shortest path between other vertices or measure of shortest paths a node is part of.

Results

Betweenness centrality

Table 15 Betweenness centrality

	exposure	Betweenness centrality
KPAH4K	0	0.2135
KPAH4L	1	0.1893
KPAH4M	8	0.3029
KPAH4H	16	0.1694
KPAH4PK		0.1342
KPAH24K	0	0.1803
KPAH24L	1	0.3127
KPAH24M	8	0.2085
KPAH24H	16	0.1455
KPAH24PK		0.2960

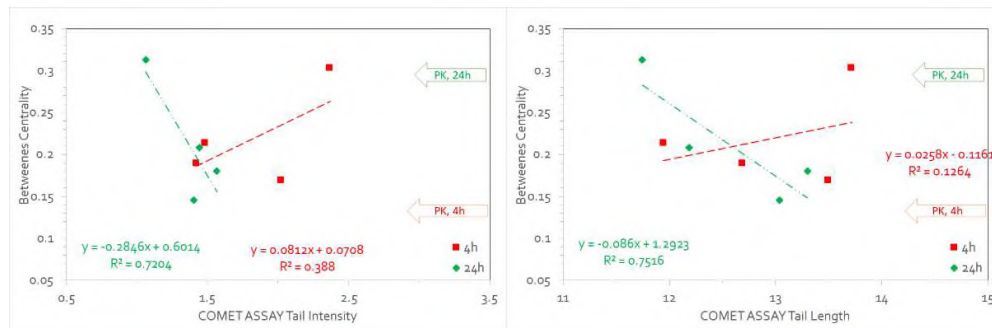


Figure 167 Betweenness Centrality versus COMET ASSAY Tail Intensity (left) and versus COMET ASSAY Tail Length (right) at 4h (red, ■) and 24h (green, ◆)

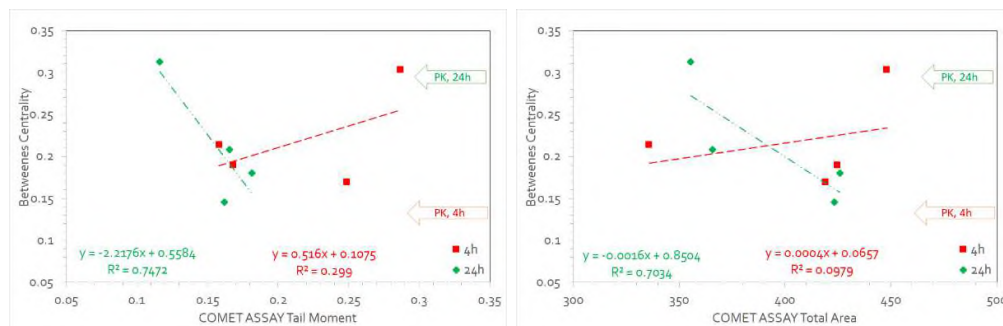


Figure 168 Betweenness Centrality versus COMET ASSAY Tail Moment (left) and versus COMET ASSAY Total Area (right) at 4h (red, ■) and 24h (green, ◆)

Shannon Entropy of Normalized Betweenness Centrality

Table 16 Shannon Entropy of Normalized Betweenness Centrality

	exposure	Shannon Entropy of Normalized Betweenness Centrality
KPAH4K	0	5.4653
KPAH4L	1	5.8821
KPAH4M	8	5.6012
KPAH4H	16	6.1536
KPAH4PK		5.6734
KPAH24K	0	6.0674
KPAH24L	1	6.9586
KPAH24M	8	6.4015
KPAH24H	16	6.5657
KPAH24PK		6.4204

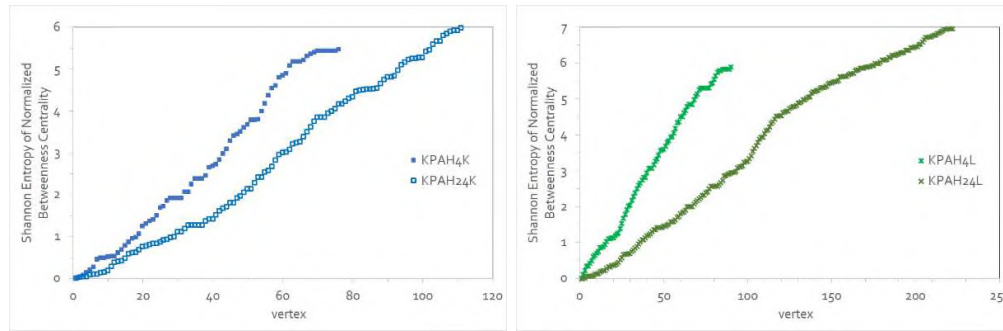


Figure 169 Shannon Entropy of Normalized Betweenness Centrality, control (K) (left) and low exposure (L) (right): 4h (filled symbols) and 24h (empty symbols)

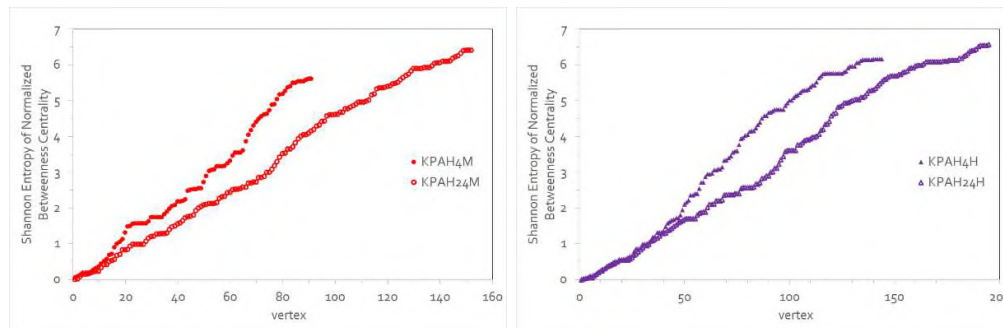


Figure 170 Shannon Entropy of Normalized Betweenness Centrality, middle exposure (M) (left) and high exposure (H) (right): 4h (filled symbols) and 24h (empty symbols)

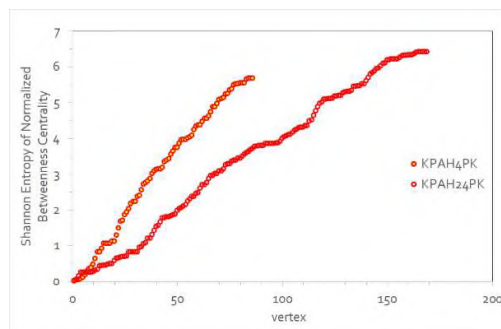


Figure 171 Shannon Entropy of Normalized Betweenness Centrality, positive control (PK): 4h (filled symbols) and 24h (empty symbols)

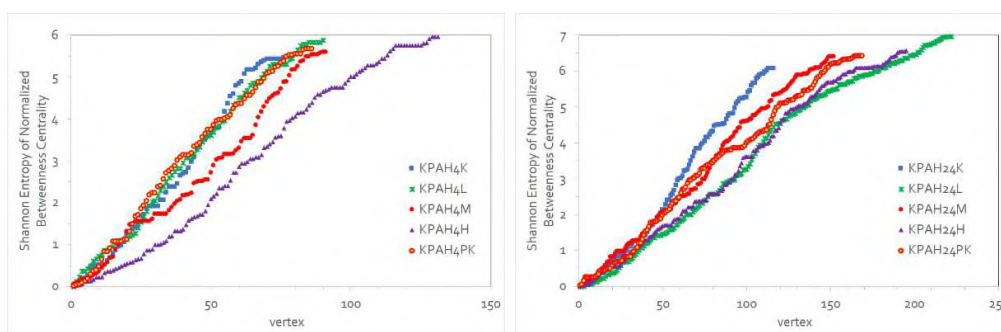


Figure 172 Shannon Entropy of Normalized Betweenness Centrality at 4h (left) and 24h (right)

Edge Betweenness Centrality

Edge Betweenness is the number of shortest paths between vertices with the edge.

Highlight Edge Betweenness Centrality

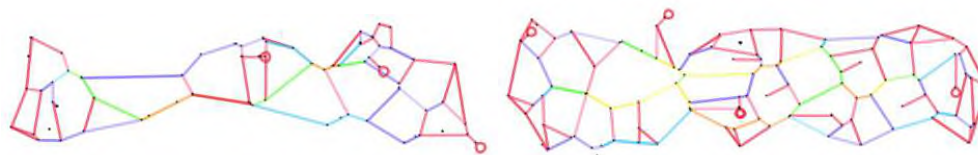


Figure 33 Highlight Edge Betweenness Centrality of control (K): 4h (left) and 24h (right)

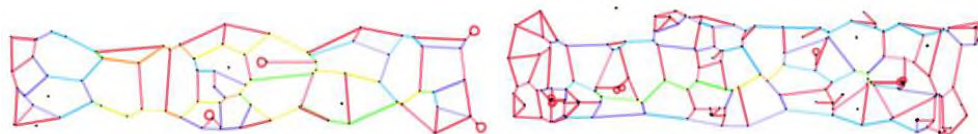


Figure 34 Highlight Edge Betweenness Centrality of low exposure (L): 4h (left) and 24h (right)

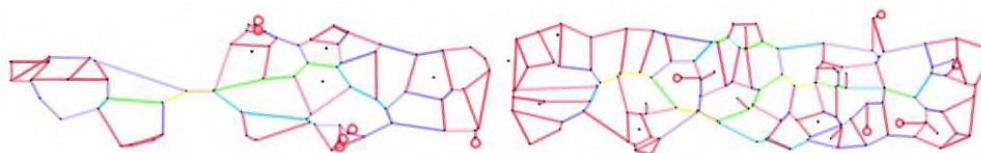


Figure 35 Highlight Edge Betweenness Centrality of middle exposure (M): 4h (left) and 24h (right)

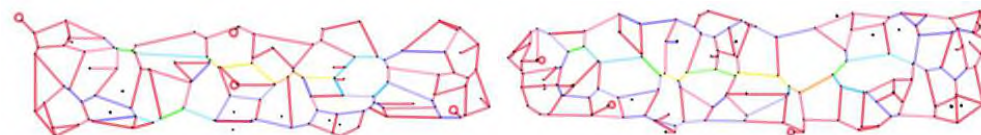


Figure 36 Highlight Edge Betweenness Centrality of high exposure (H): 4h (left) and 24h (right)

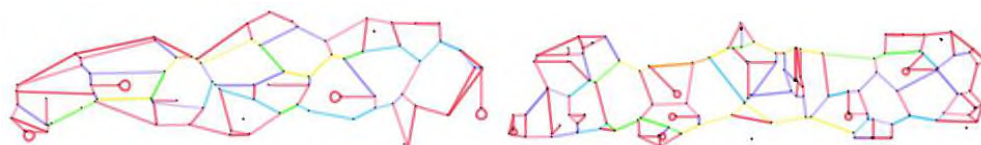


Figure 37 Highlight Edge Betweenness Centrality of positive control (PK): 4h (left) and 24h (right)

Shannon Entropy of Edge Betweenness Centrality

Table 17 Shannon Entropy of Edge Betweenness Centrality

	exposure	Shannon Entropy of Edge Betweenness Centrality
KPAH4K	0	5.8600
KPAH4L	1	6.2632
KPAH4M	8	6.1363
KPAH4H	16	6.6240
KPAH4PK		6.1419
KPAH24K	0	6.5386
KPAH24L	1	7.3898
KPAH24M	8	6.8719
KPAH24H	16	7.0280
KPAH24PK		6.9040

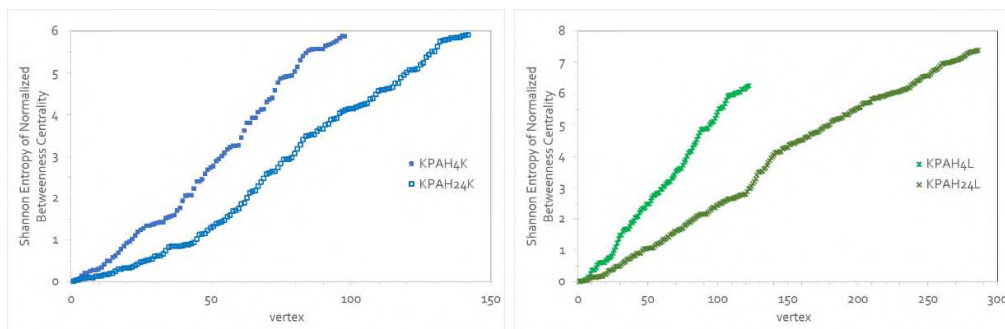


Figure 173 Shannon Entropy of Edge Betweenness Centrality, control (K) (left) and low exposure (L) (right): 4h (filled symbols) and 24h (empty symbols)

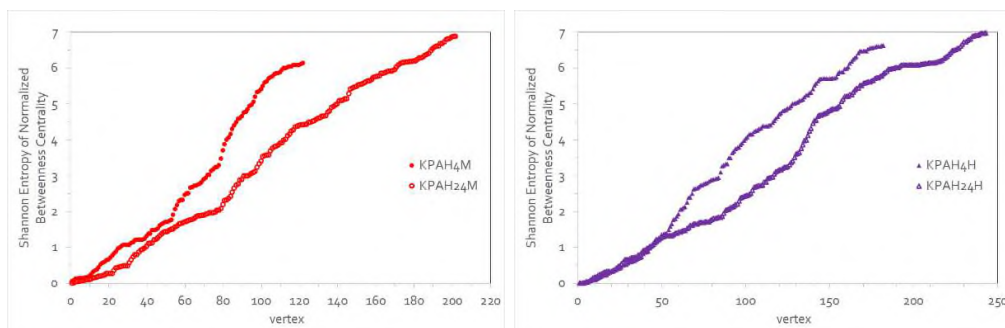


Figure 174 Shannon Entropy of Edge Betweenness Centrality, middle exposure (M) (left) and high exposure (H) (right): 4h (filled symbols) and 24h (empty symbols)

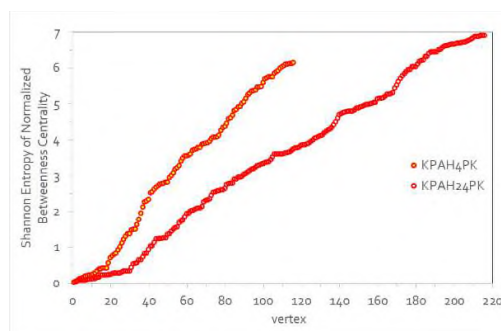


Figure 175 Shannon Entropy of Edge Betweenness Centrality, positive control (PK): 4h (filled symbols) and 24h (empty symbols)

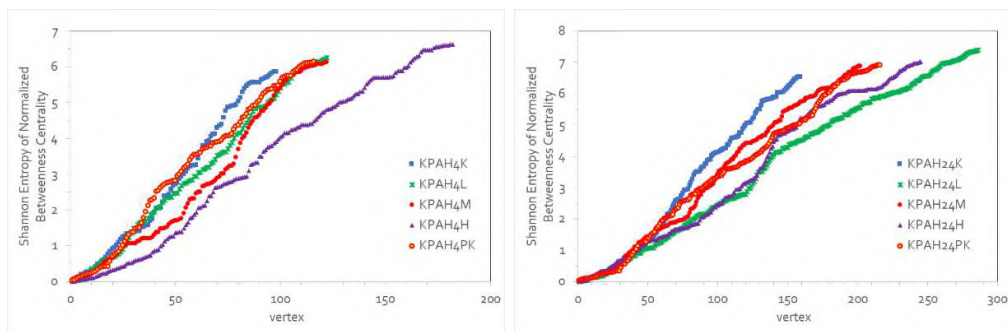


Figure 176 Shannon Entropy of Edge Betweenness Centrality at 4h (left) and 24h (right)

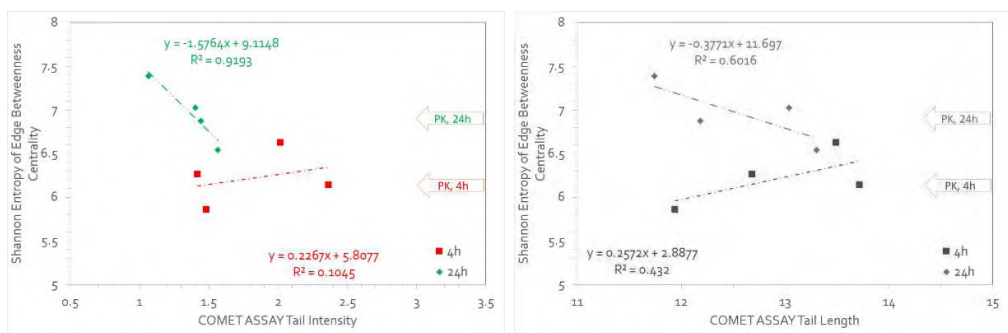


Figure 177 Eigenvector centrality versus COMET ASSAY Tail Intensity (left) and versus COMET ASSAY Tail Length (right) at 4h (red, ■) and 24h (green, ◆)

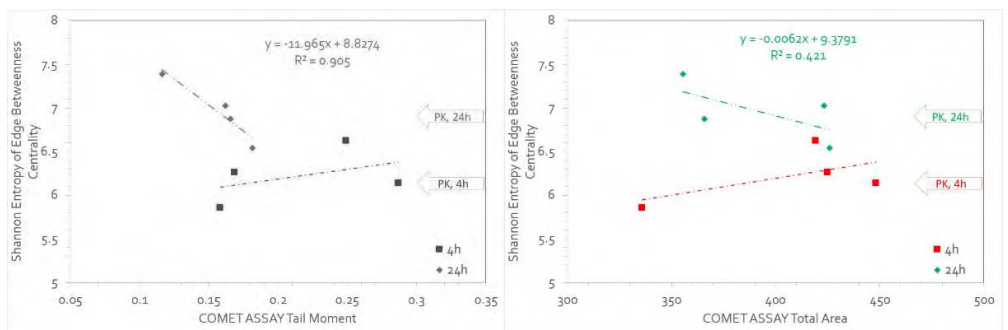


Figure 178 Eigenvector centrality versus COMET ASSAY Tail Moment (left) and versus COMET ASSAY Total Area (right) at 4h (red, ■) and 24h (green, ◆)

Eigenvector centrality

Eigenvector centrality^{114, 115} introduces a refined perspective by considering the quantity of connections and the quality of those connections. A node's eigenvector centrality score is influenced by its connections to other highly influential vertices. The eigenvector is best known within graph analysis as a measure of graph centrality:

$$Ax = \lambda x$$

λ is the principal eigenvalue for an adjacency matrix A , and x is the eigenvector associated with the eigenvalue. The eigenvector of the largest eigenvalue of an adjacency matrix could make a good graph centrality measure. Usually, the first dimension captures the „global“ aspects of distances among actors; second and further dimensions capture more specific and local sub-structures. This eigenvector is interpretable as a measure of centrality¹¹⁶. Eigenvector Centrality is a measure of the degree of the vertex as well as the degree of its neighbours. Eigenvector centrality assigns importance based on the importance of a node's neighbours. The eigenvector centrality of a vertex of proportional to the sum of the centralities of its neighbours (*by definition*). The constant of proportionality is the eigenvalue. Eigenvector Centrality gives a list of centralities, c , that can be expressed as a weighted sum of centralities of its neighbours. With λ_1 being the largest eigenvalue of the adjacency matrix for the graph G , you have: Eigenvector Centrality (G)= $c=1/\lambda_1 A^T.c$

Eigenvector centrality extends degree centrality. In the degree centrality, all node connections are credited of equal importance. Eigenvector centrality provides a relative score to each node depending on the vertices (high-scoring and low-scoring) it is connected to. Eigenvector Centrality will give high centralities to vertices connected to many other well-connected vertices. Google's PageRank¹¹⁷ and Katz centrality¹¹⁸ are based on the same idea. Katz centrality measures the relative influence of each node in a graph by considering its immediate neighbouring vertices and non-immediate neighbouring vertices connected through immediate neighbouring vertices.

Results

Table 18 Eigenvector centrality

	exposure	Eigenvector centrality	Shannon Entropy of Eigenvector centrality
KPAH4K	0	0.2135	4.7014
KPAH4L	1	0.1893	5.0145
KPAH4M	8	0.9801	5.1681
KPAH4H	16	0.7833	5.5812
KPAH4PK		0.5191	4.6398
KPAH24K	0	0.8634	5.1147
KPAH24L	1	0.9602	5.8456
KPAH24M	8	0.7651	5.4560
KPAH24H	16	0.9880	7.0438
KPAH24PK		0.9123	5.4932

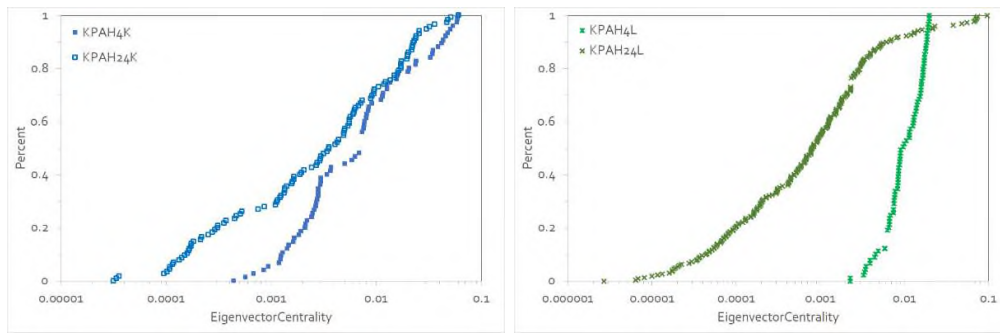


Figure 179 Eigenvector centrality of control (K) (left) and low exposure (L) (right) at 4h and 24h

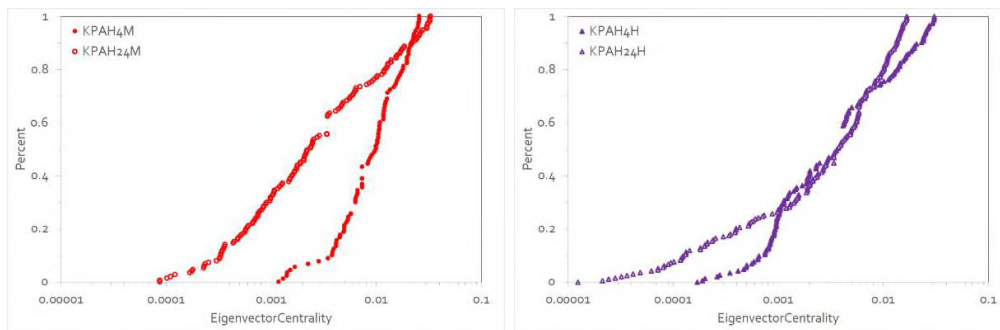


Figure 180 Eigenvector centrality of middle exposure (M) (left) and high exposure (H) (right) at 4h and 24h

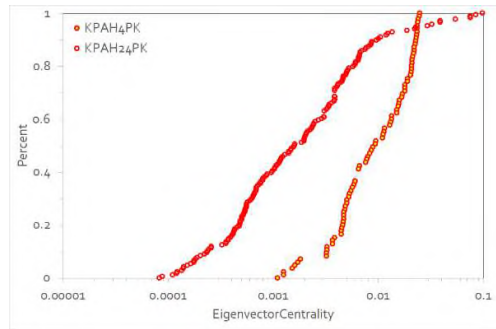


Figure 181 Eigenvector centrality of positive control (PK) at 4h and 24h

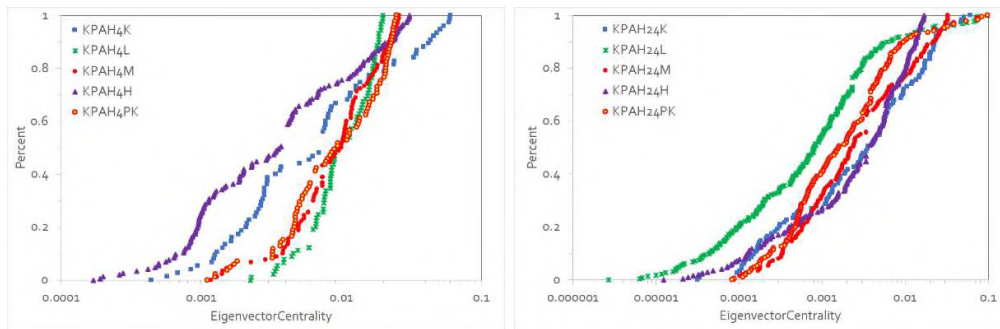


Figure 182 Eigenvector centrality at 4h (left) and 24h (right)

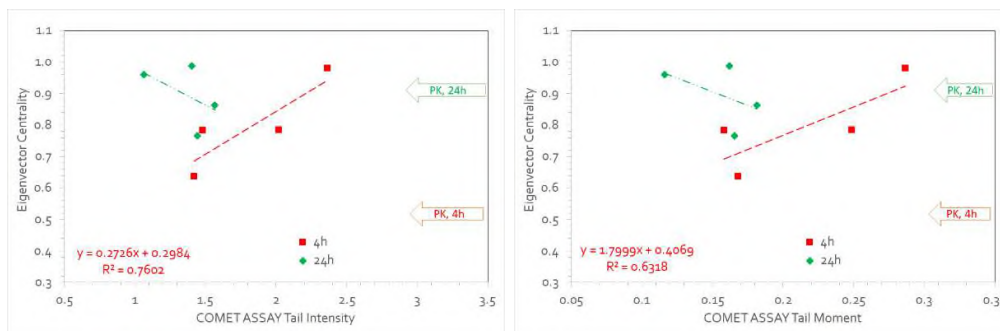


Figure 183 Eigenvector centrality versus COMET ASSAY Tail Intensity (left) and versus COMET ASSAY Tail Moment (right) at 4h (red, ■) and 24h (green, ◆)

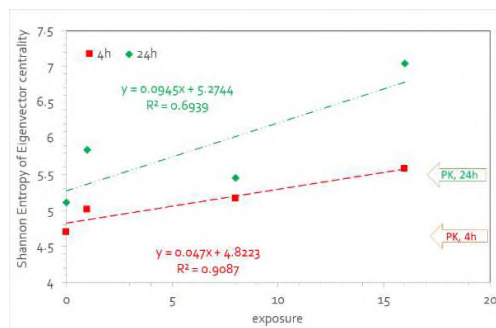


Figure 184 Shannon Entropy of Eigenvector centrality at 4h (red, ■) and 24h (green, ◆)

Information centrality

The information centrality is based on the concept of efficient propagation of information over the graph. Information centrality^{119,120,121} is a graph analysis metric that measures how central a node is to the efficient flow of information throughout the graph, unlike traditional closeness centrality which only considers shortest paths. It is often based on the concept of current-flow or effective resistance, where a node's importance is calculated by how much graph efficiency drops when that node is removed. A higher information centrality score means a node is more critical for communication, as its removal would have a greater negative impact on the graph's efficiency. Information centrality is a variation of closeness centrality but uses a different model of information spread. While closeness centrality uses the average shortest distance to all other vertices, information centrality considers the total „information“ across all paths.

The definition of “information” comes from the theory of statistical estimation. It is supposed there exists noise in every link, causing losing during every single transition of the information. The longer the path is, the more the loss is. Mathematically, the total information sent on a path equals the reciprocal of the path length. The quantity of the information that can be transmitted between a pair of vertices v_i and v_j equals the summation of the information sent through every possible path between them, denoted as q_{ij} . Information index considers all possible paths in the connected graph^{122,123}. Characteristic Polynomial gives the characteristic polynomial for the matrix of Information Centrality.

Results

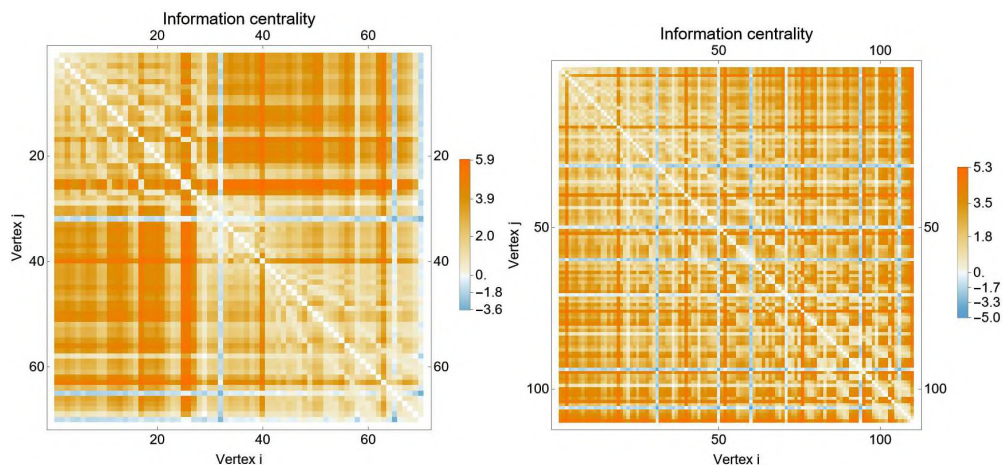


Figure 185 Matrix plot of Information centrality of control (K): 4h (left) and 24h (right)

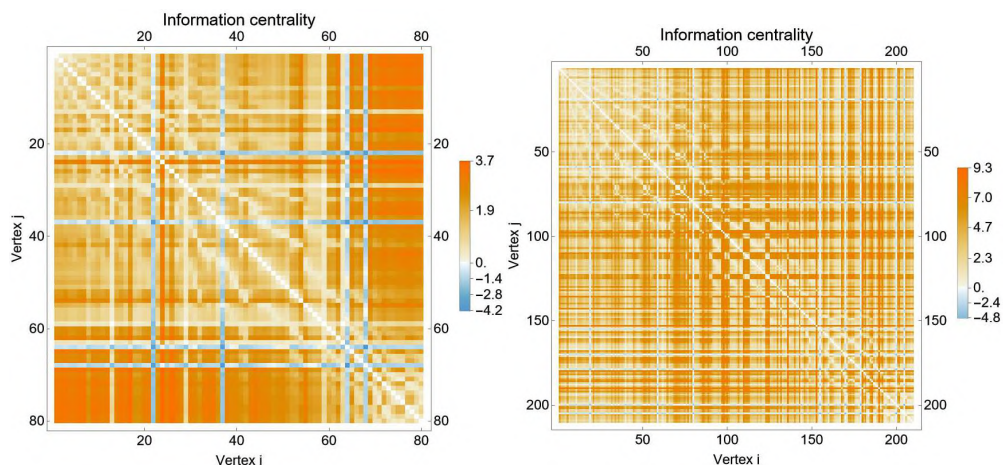


Figure 186 Matrix plot of Information centrality of low exposure (L): 4h (left) and 24h (right)

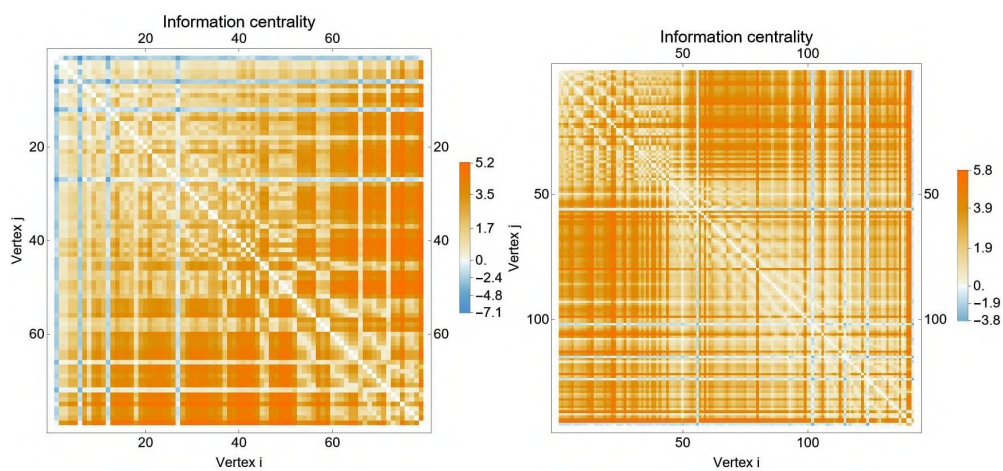


Figure 187 Matrix plot of Information centrality of middle exposure (M): 4h (left) and 24h (right)

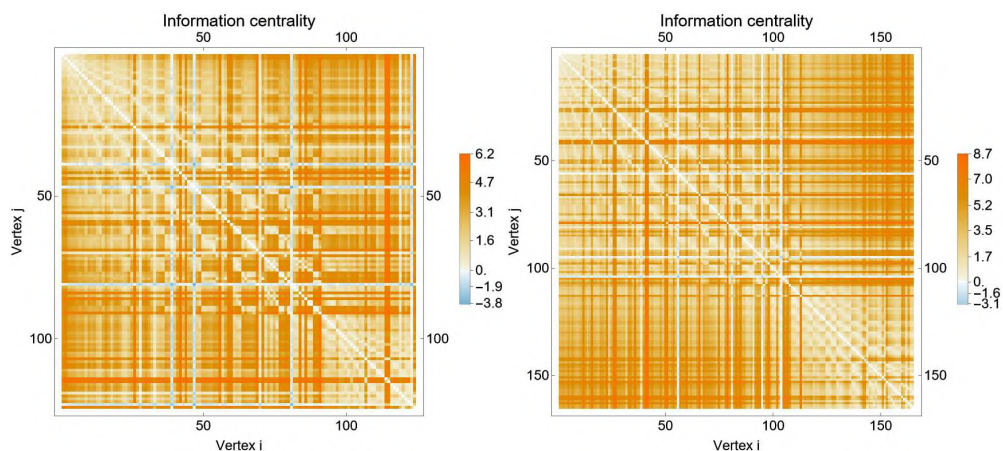


Figure 188 Matrix plot of Information centrality of high exposure (H): 4h (left) and 24h (right)

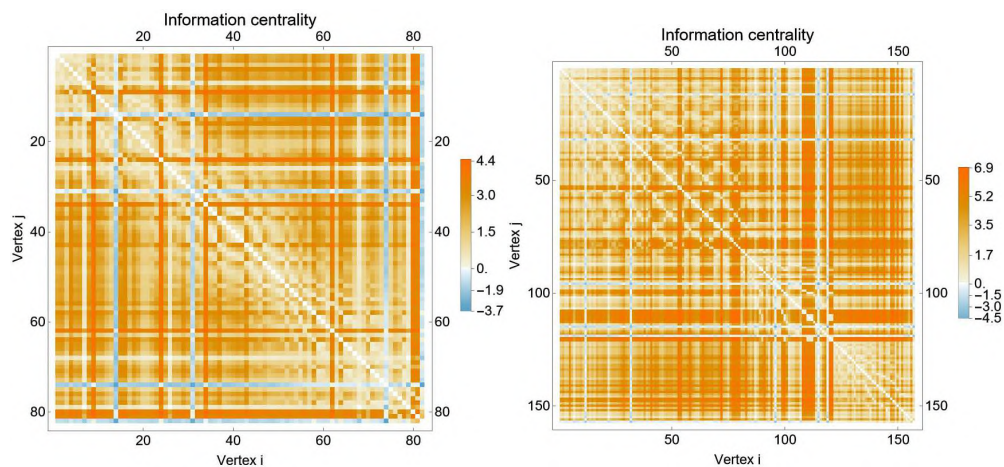


Figure 189 Matrix plot of Information centrality of positive control (PK): 4h (left) and 24h (right)

Eigenvalues of Information centrality

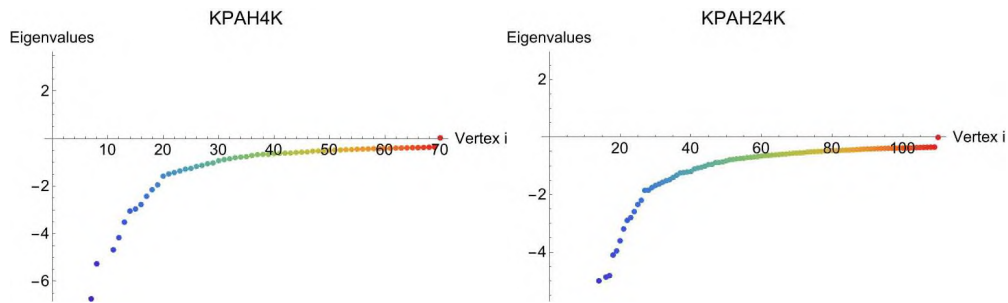


Figure 190 Eigenvalues of Information centrality of control (K): 4h (left) and 24h (right)

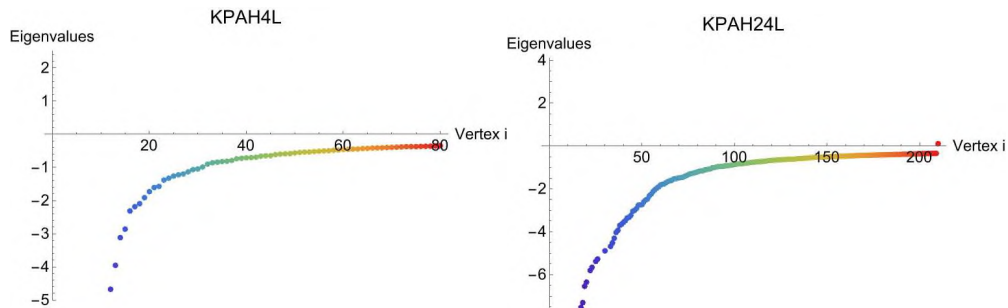


Figure 191 Eigenvalues of Information centrality of low exposure (L): 4h (left) and 24h (right)

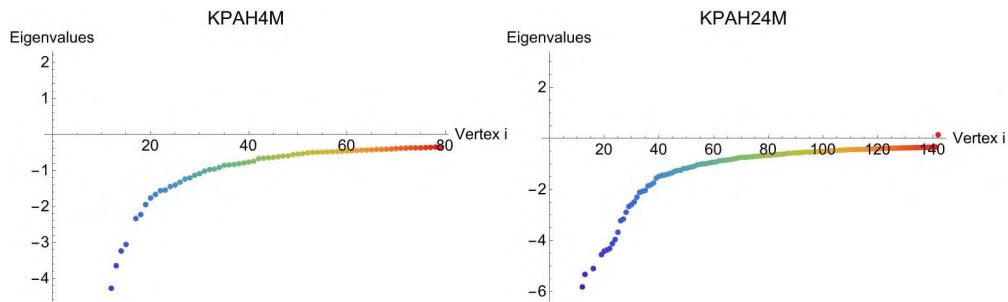


Figure 192 Eigenvalues of Information centrality of middle exposure (M): 4h (left) and 24h (right)

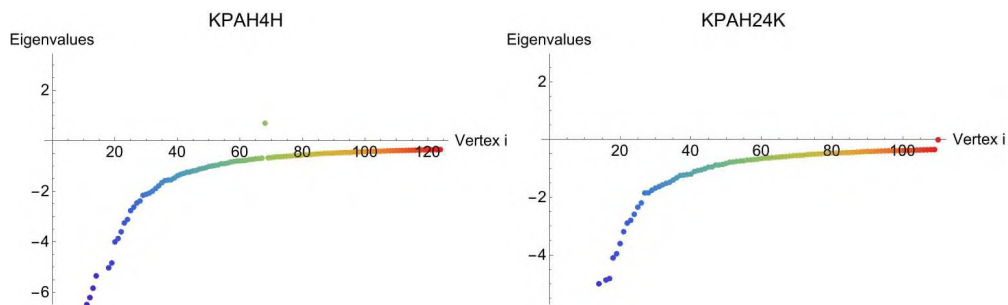


Figure 193 Eigenvalues of Information centrality of high exposure (H): 4h (left) and 24h (right)

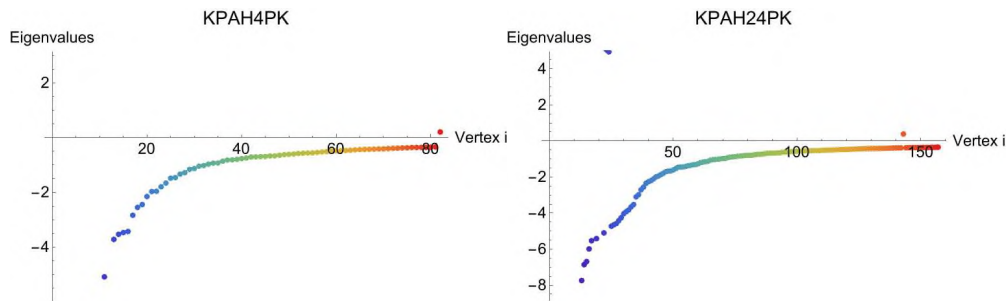


Figure 194 Eigenvalues of Information centrality of positive control (PK): 4h (left) and 24h (right)

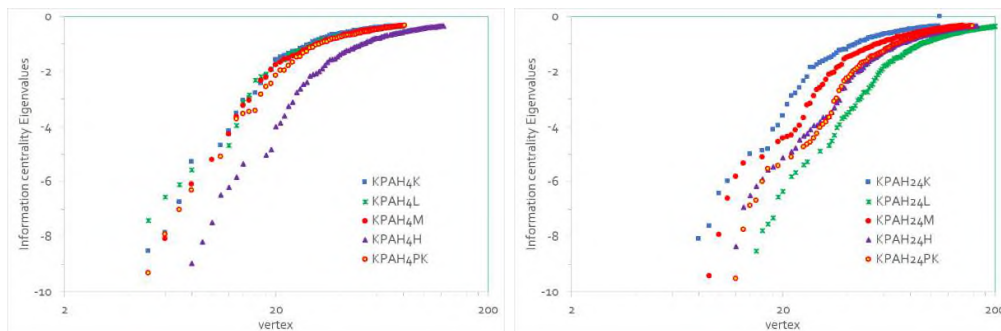


Figure 195 Eigenvalues of Information centrality: 4h (left) and 24h (right)

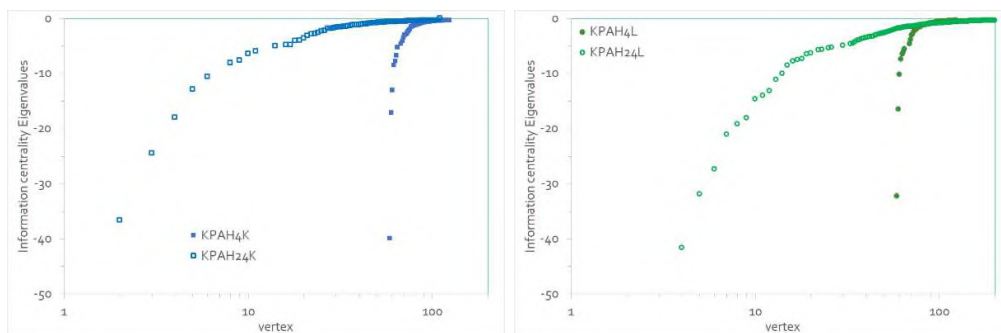


Figure 196 Eigenvalues of Information centrality: 4h (full symbols) and 24h (empty symbols), control (K) (left) and low exposure (L) (right)

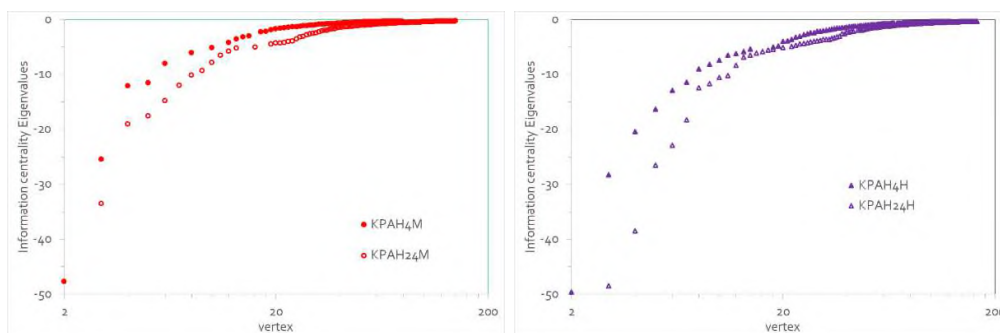


Figure 197 Eigenvalues of Information centrality: 4h (full symbols) and 24h (empty symbols), middle exposure (M) (left) and high exposure (L) (right)

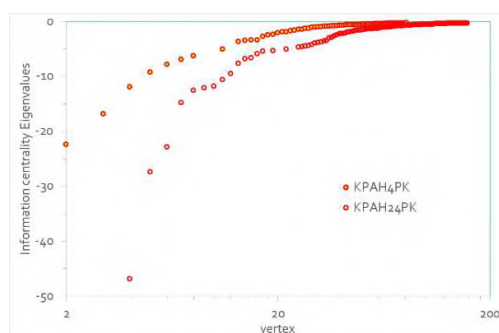


Figure 198 Eigenvalues of Information centrality: 4h (full symbols) and 24h (empty symbols), positive control (PK)

Eccentricity

The eccentricity¹²⁴ of a vertex is the maximum shortest distance from that vertex to any other vertex in the graph. This value is calculated by finding the maximum distance from a vertex to all other reachable vertices in the graph. Eccentricity centrality is a measure of the centrality of a node in a graph based on having a small maximum distance from a node to every other reachable node (i.e. the graph eccentricities¹²⁵).

Eccentricity indicates the centrality of each node and collectively encodes graph properties: the radius and the diameter, the minimum and maximum eccentricity, respectively, over all the vertices in the graph. Eccentricity not only reflects the centrality of each node, a node's small eccentricity denotes its central location in the graph, but also collectively encodes fundamental global properties of a graph: the radius of the graph is the minimum eccentricity over all the vertices while the diameter the maximum. A high eccentricity centrality score means the node has a small maximum distance to all other vertices, suggesting it is "central" and well-connected.

A low eccentricity centrality score means the node has a large maximum distance to some other vertices, suggesting it is more on the "edge" of the graph. So, the eccentricities over all the vertices in the graph, called the *eccentricity distribution*, becomes an important descriptor of the graph¹²⁶.

Graph invariants dependent on vertex distances in graphs of Comet assays images and eccentricity-related topological descriptors of graphs correlate with several properties of Comet assays¹²⁷. These indices exhibit remarkable predictive capabilities.

Results

Table 19 Eccentricity Centrality, central vertex

	exposure	Eccentricity Centrality, central vertex	Sahnnon Entropy of Eccentricity Centrality
KPAH4K	0	1.9326	2.2536
KPAH4L	1	1.8892	1.5012
KPAH4M	8	1.9108	1.7271
KPAH4H	16	1.9120	1.3492
KPAH4PK		1.9690	3.2660
KPAH24K	0	1.9742	3.2606
KPAH24L	1	1.9584	1.8231
KPAH24M	8	0.9551	2.3403
KPAH24H	16	0.9431	2.3775
KPAH24PK		0.4798	4.2474

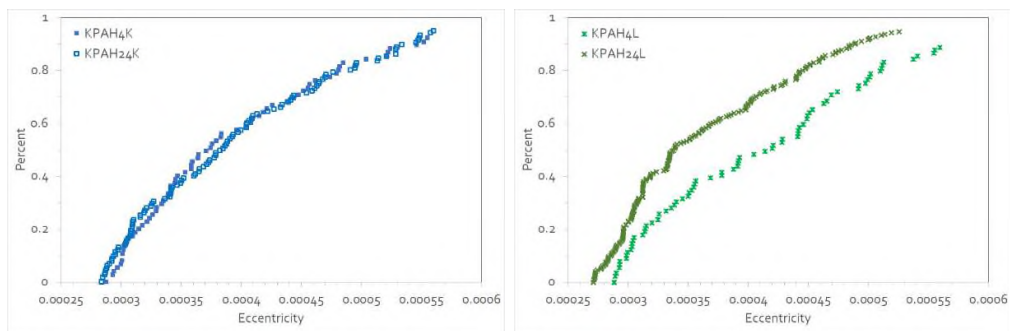


Figure 199 Eccentricity Centrality of control (K) (left) and low exposure (L) (right) at 4h and 24h

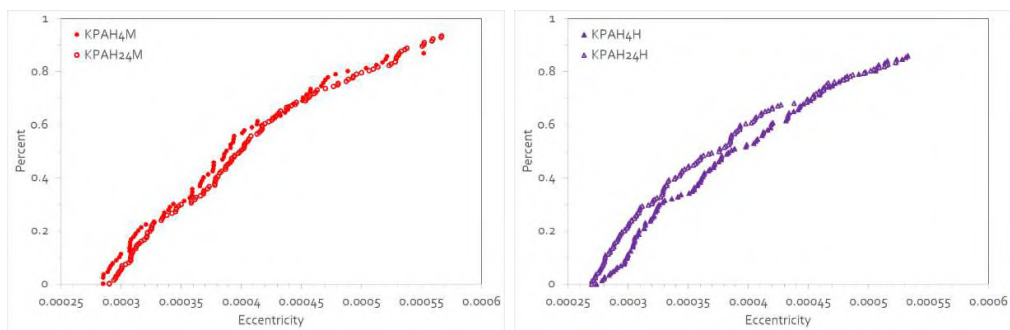


Figure 200 Eccentricity Centrality of middle exposure (M) (left) and high exposure (H) (right) at 4h and 24h

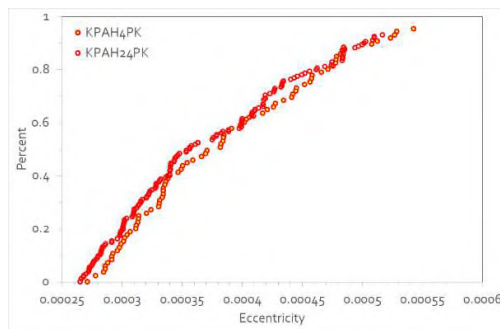


Figure 201 Eccentricity Centrality of positive control (PK) at 4h and 24h

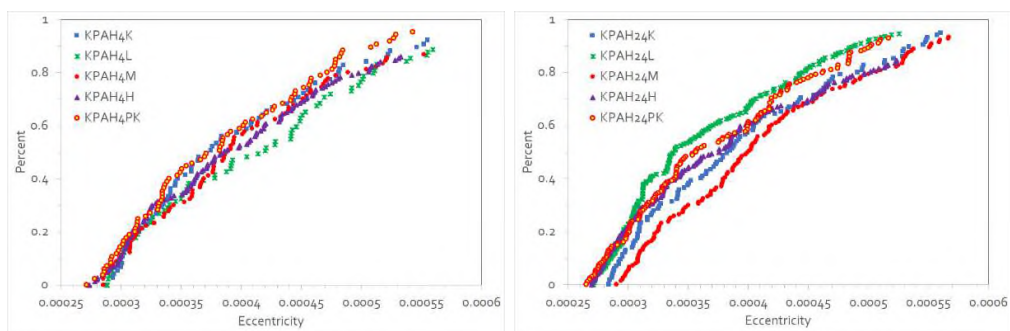


Figure 202 Eccentricity Centrality at 4h (left) and 24h (right)

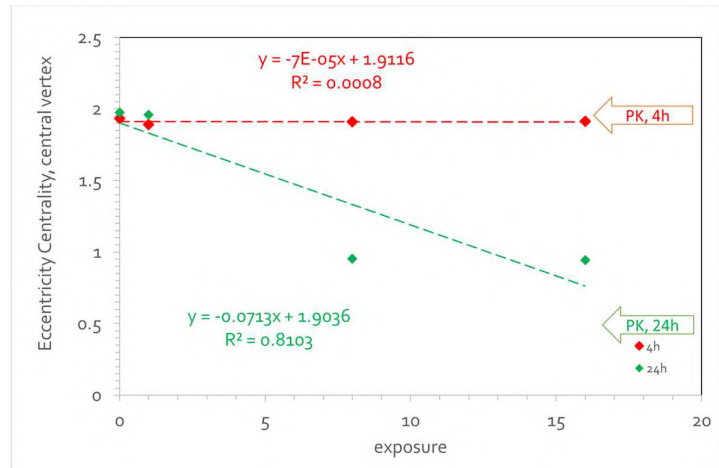


Figure 203 Eccentricity Centrality, central vertex at 4h (red) and 24h (green).

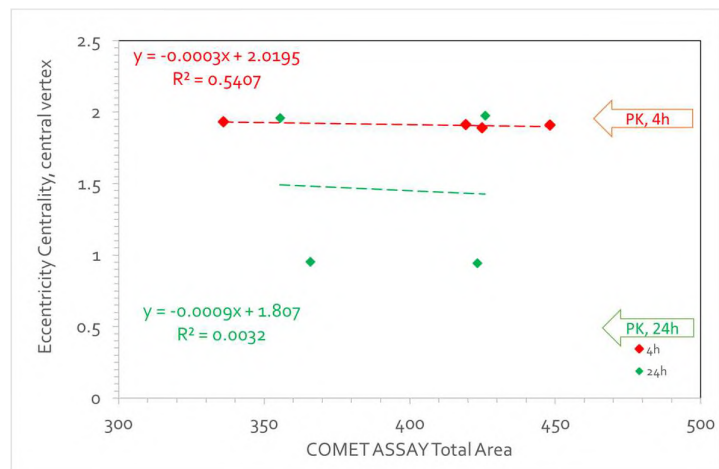


Figure 204 Eccentricity Centrality, central vertex vs. COMET ASSAY Total Area at 4h (red) and 24h (green).

Assortativity

A graph is called assortative if the vertices with higher degree connect with other vertices that also have substantial connectivity^{128,129,130,131}. Assortativity measures the tendency of vertices in a graph to connect with other vertices that are similar to them based on a certain property, such as degree. Assortativity measure is a Pearson correlation coefficient of the degree-degree correlation. $r > 0$ (degree assortativity) indicates a tendency of high degree vertices to connect to other high degree ones. $r < 0$ (degree disassortativity) indicates the tendency of high degree vertices to be connected to low degree ones. Assortativity is measured by an assortativity coefficient, which ranges from -1 to +1. A coefficient near +1 indicates assortative mixing, where similar vertices connect, while a coefficient near -1 indicates disassortative mixing, where dissimilar vertices connect. A coefficient of 0 means there is no assortative mixing.

Results

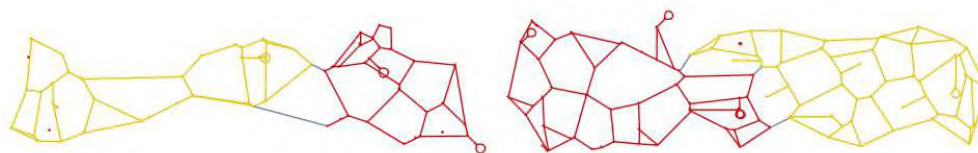


Figure 205 Graph Assortativity of control (K): 4h (left) and 24h (right)

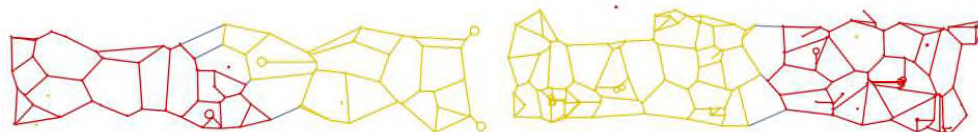


Figure 206 Graph Assortativity of low exposure (L): 4h (left) and 24h (right)

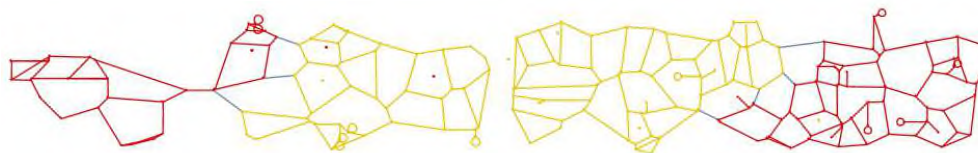


Figure 207 Graph Assortativity of middle exposure (M): 4h (left) and 24h (right)

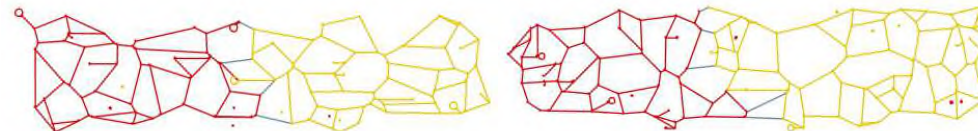


Figure 208 Graph Assortativity of high exposure (H): 4h (left) and 24h (right)

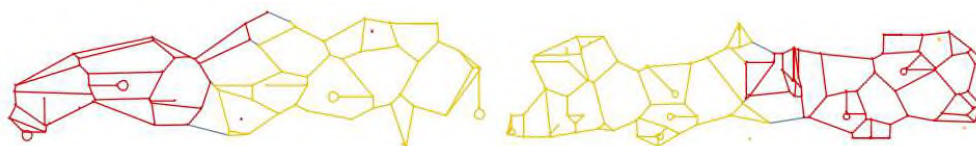


Figure 209 Graph Assortativity of positive control (PK): 4h (left) and 24h (right)

Table 20 Graph Assortativity

	exposure	Graph Assortativity, r
KPAH4K	0	0.5305
KPAH4L	1	0.3179
KPAH4M	8	0.3863
KPAH4H	16	0.3207
KPAH4PK		0.6472
KPAH24K	0	0.3902
KPAH24L	1	0.5438
KPAH24M	8	0.5405
KPAH24H	16	0.4918
KPAH24PK		0.4621

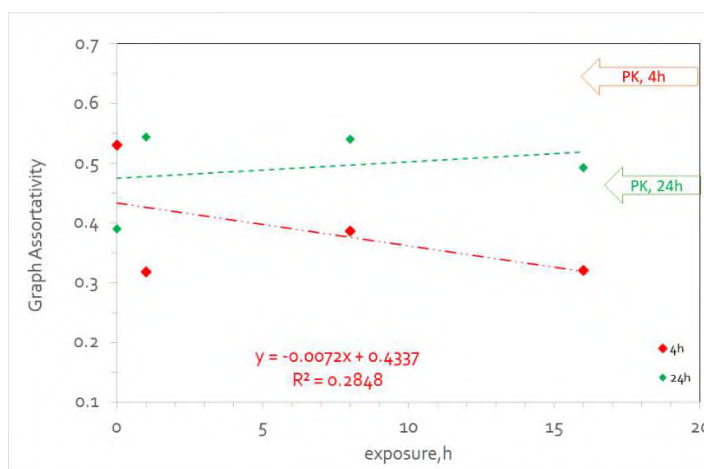


Figure 210 Graph Assortativity at 4h (red) and 24h (green).

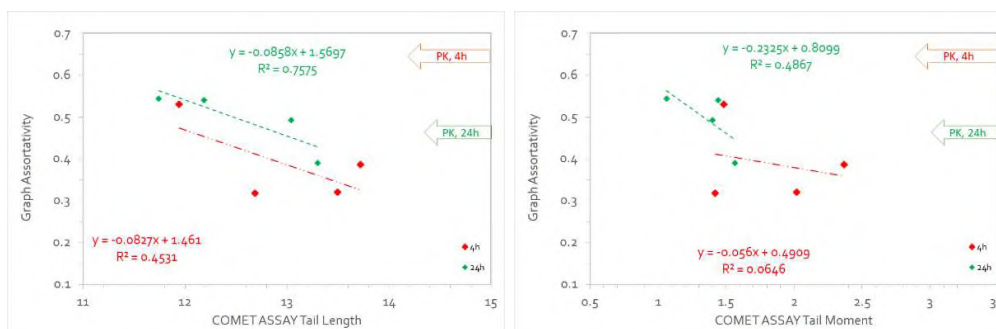


Figure 211 Correlation of Graph Assortativity vs. COMET ASSAY Tail Moment (left) and of Graph Assortativity vs. COMET ASSAY Tail Length (right) at 4h (red) and 24h (green).

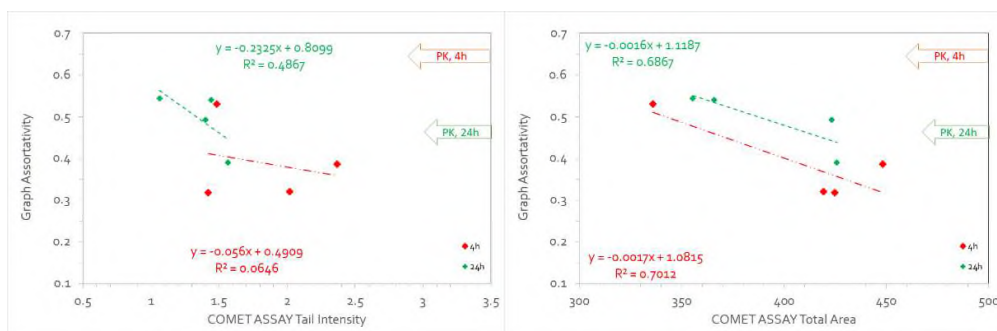


Figure 212 Correlation of Graph Assortativity vs. COMET ASSAY Tail Intensity (left) and of Graph Assortativity vs. COMET ASSAY Total Area (right) at 4h (red) and 24h (green).

Adjacency matrix

The adjacency matrix, sometimes also called the connection matrix, of H is the $|V| \times |V|$ matrix. a_{ij} = the number of edges joining v_i and v_j . A Weighted Adjacency Matrix^{132, 133} of a simple graph can also be defined for a real positive symmetric function $f(d_i, d_j)$ on the vertex degrees d_i of a graph.

A measure of the information content of the spectrum of the adjacency matrix, where $p_j = (\exp(\lambda_j) / (\text{Estrada index of a graph}))$, where the Estrada index of a graph is a topological index defined as the sum of the exponentials of its eigenvalues, where the eigenvalues come from the graph's adjacency matrix. It can be calculated using the formula:

$$\text{Estrada index of a graph} = \sum_{i=1}^n e^{\lambda_i}$$

Results

Table 21 The Estrada index of a graph

	exposure	The Estrada index of a graph
KPAH4K	0	231.53
KPAH4L	1	287.38
KPAH4M	8	290.73
KPAH4H	16	426.36
KPAH4PK		268.43
KPAH24K	0	376.93
KPAH24L	1	664.58
KPAH24M	8	465.17
KPAH24H	16	574.44
KPAH24PK		507.06

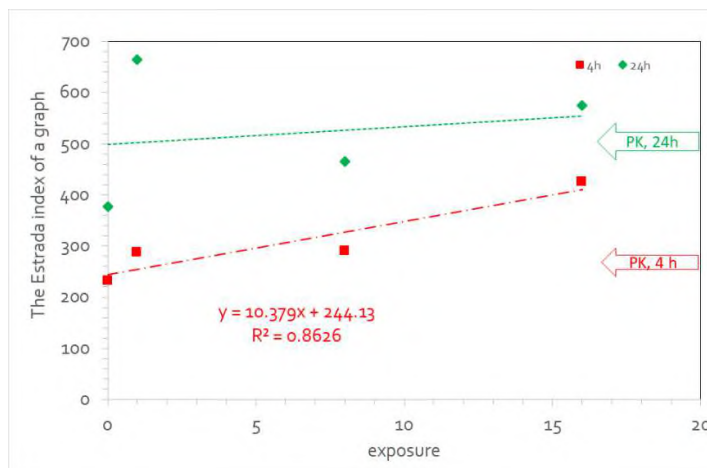


Figure 213 The Estrada index of a graph at 4h (red) and 24h (green).

Adjacency matrix

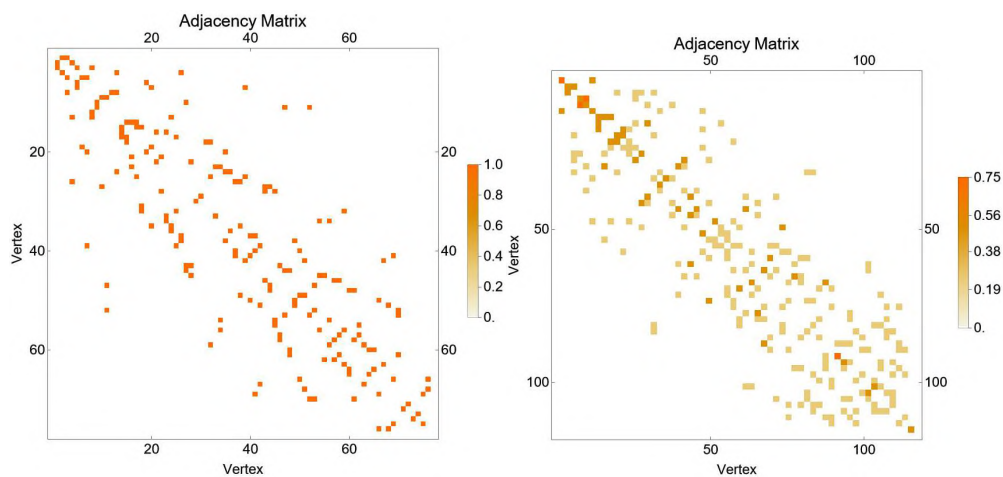


Figure 214 Adjacency matrix of control (K): 4h (left) and 24h (right)

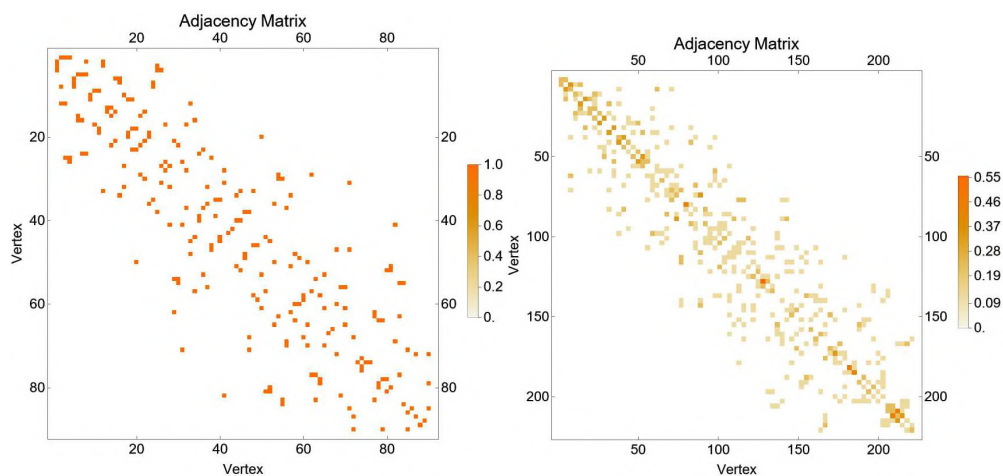


Figure 215 Adjacency matrix of low exposure (L): 4h (left) and 24h (right)

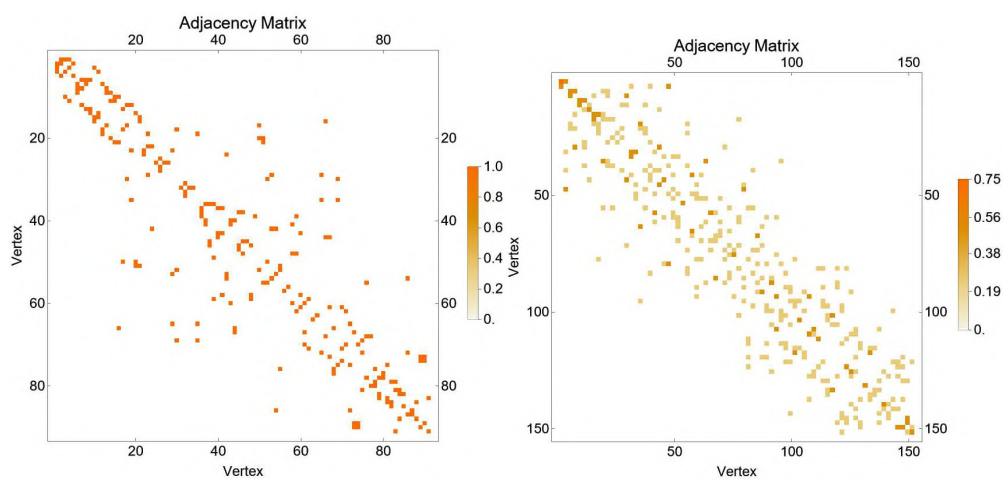


Figure 216 Adjacency matrix of middle exposure (M): 4h (left) and 24h (right)

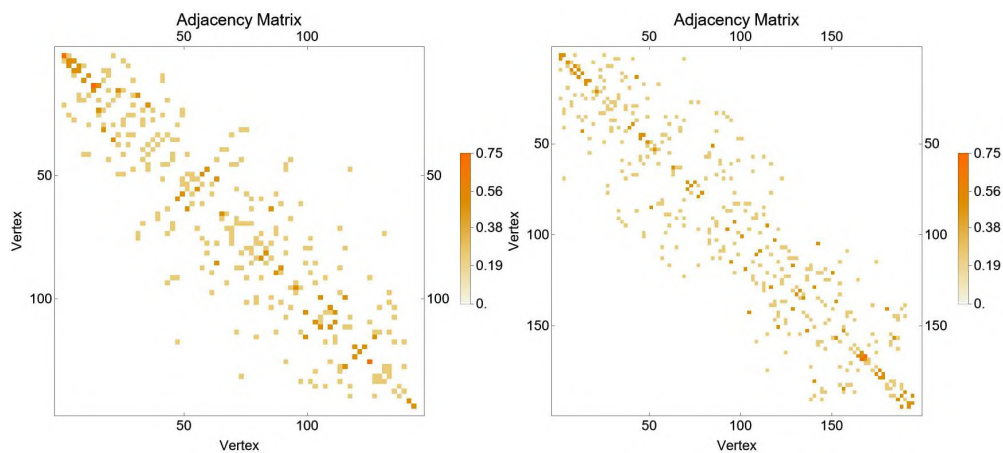


Figure 217 Adjacency matrix of high exposure (H): 4h (left) and 24h (right)

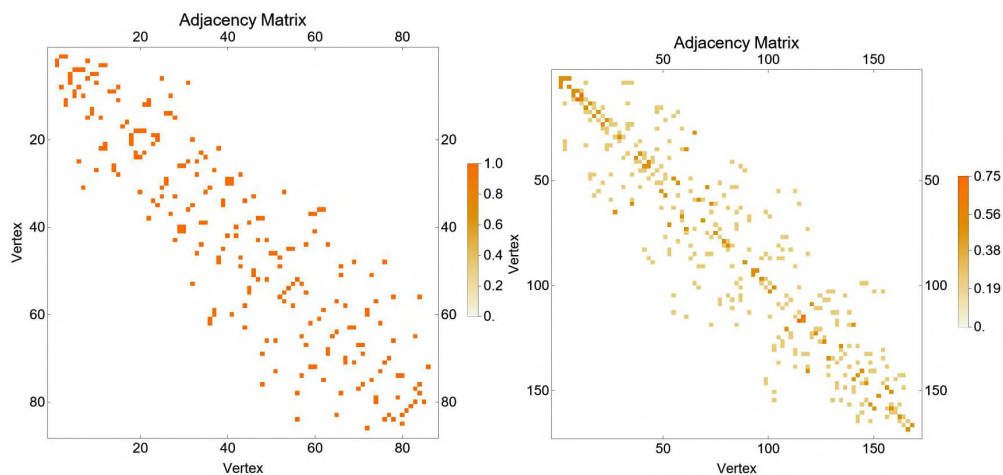


Figure 218 Adjacency matrix of positive control (PK): 4h (left) and 24h (right)

Table 22 Shannon entropy (or base 2 entropy) of Entropy Adjacency Matrix, and Shannon entropy (or base 2 entropy) of Eigenvalues of Adjacency Matrix distribution, Estimated Normal Distribution of Eigenvalues of Adjacency Matrix parameters

	exposure	Shannon entropy (or base 2 entropy) of		Estimated Normal Distribution of Eigenvalues of Adjacency Matrix	
		Adjacency Matrix	Eigenvalues of Adjacency Matrix distribution	mean (μ , mu)	σ , sigma
KPAH4K	0	6.2479	2.5400	0.0395	1.5931
KPAH4L	1	6.4390	2.5636	0.0444	1.6324
KPAH4M	8	6.4033	2.5356	0.0659	1.6159
KPAH4H	16	7.1230	2.5558	0.0278	1.5809
KPAH4PK		6.4030	2.5715	0.0465	1.6276
KPAH24K	0	6.8170	2.6254	0.0431	1.6421
KPAH24L	1	7.7730	2.5660	0.0360	1.5935
KPAH24M	8	7.2479	2.5685	0.0329	1.6199
KPAH24H	16	7.4893	2.5812	0.0154	1.5803
KPAH24PK		7.3328	2.5553	0.0355	1.5873

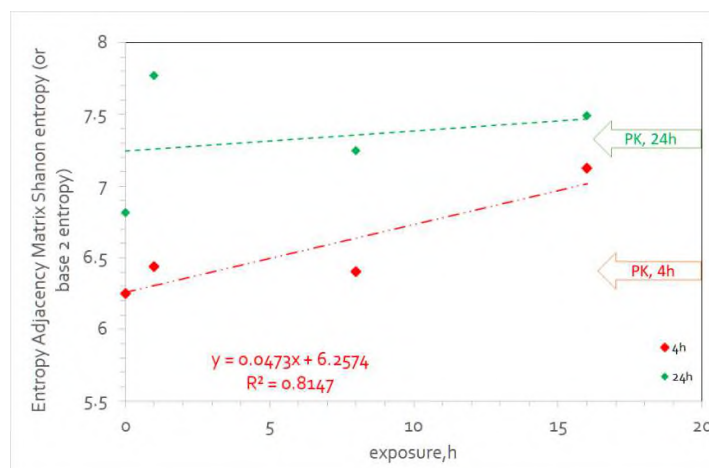


Figure 219 Shannon entropy (or base 2 entropy) of Entropy Adjacency Matrix at 4h (red) and 24h (green).

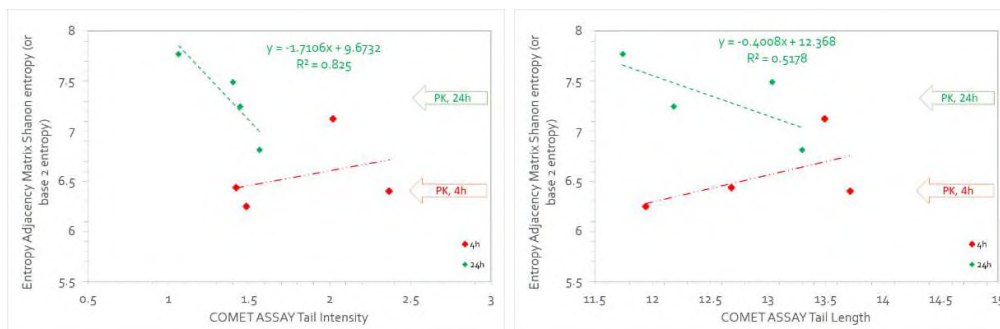


Figure 220 Shannon entropy (or base 2 entropy) of Entropy Adjacency Matrix versus COMET ASSAY Tail Intensity (left) and versus COMET ASSAY Tail Length (right) at 4h (red) and 24h (green).

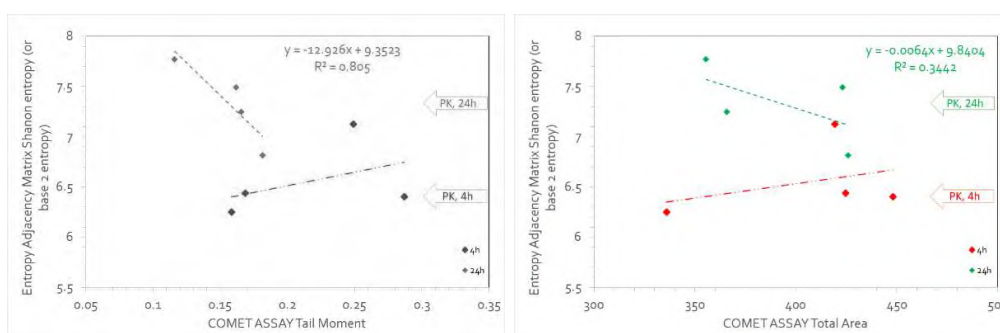


Figure 221 Shannon entropy (or base 2 entropy) of Entropy Adjacency Matrix versus COMET ASSAY Tail Moment (left) and versus COMET ASSAY Total Area (right) at 4h (red) and 24h (green).

Exact Eigenvalue Adjacency Matrix

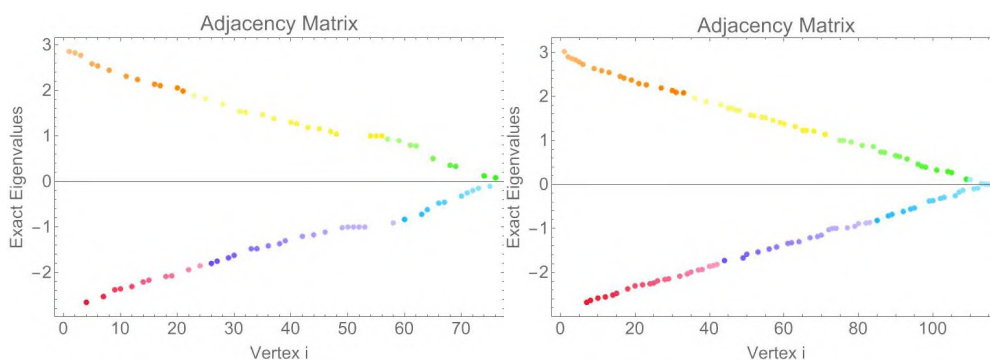


Figure 222 Exact Eigenvalue Adjacency Matrix of control (K): 4h (left) and 24h (right)

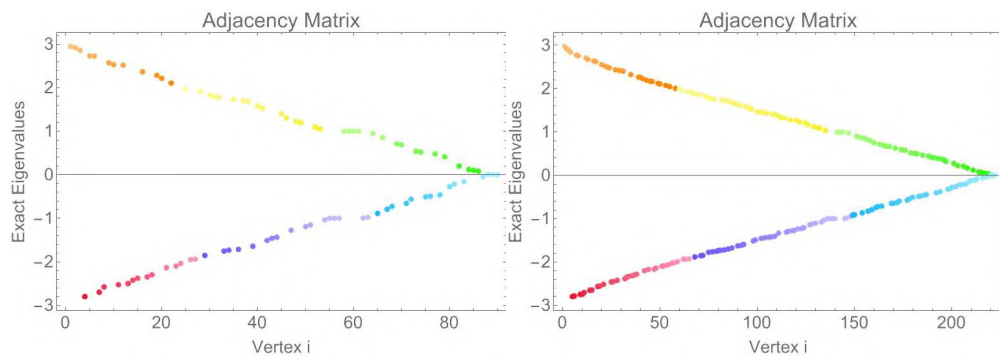


Figure 223 Exact Eigenvalue Adjacency Matrix of low exposure (L): 4h (left) and 24h (right)

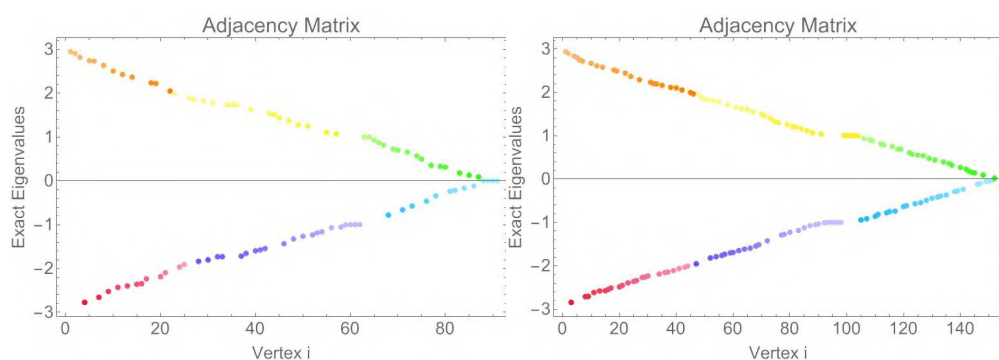


Figure 224 Exact Eigenvalue Adjacency Matrix of middle exposure (M): 4h (left) and 24h (right)

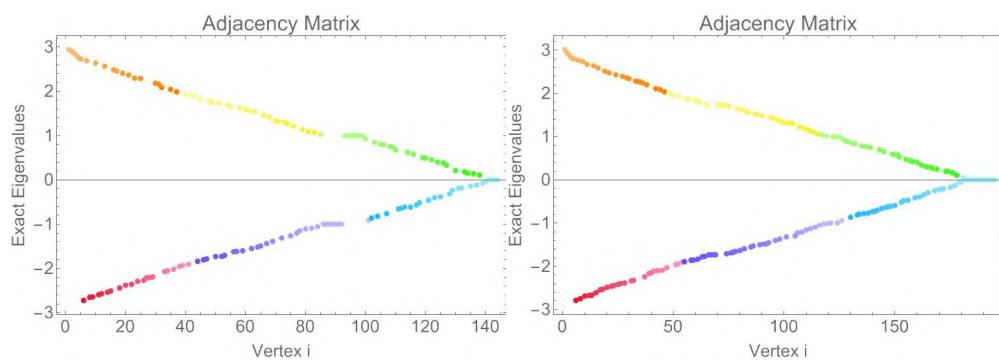


Figure 225 Exact Eigenvalue Adjacency Matrix of high exposure (H): 4h (left) and 24h (right)

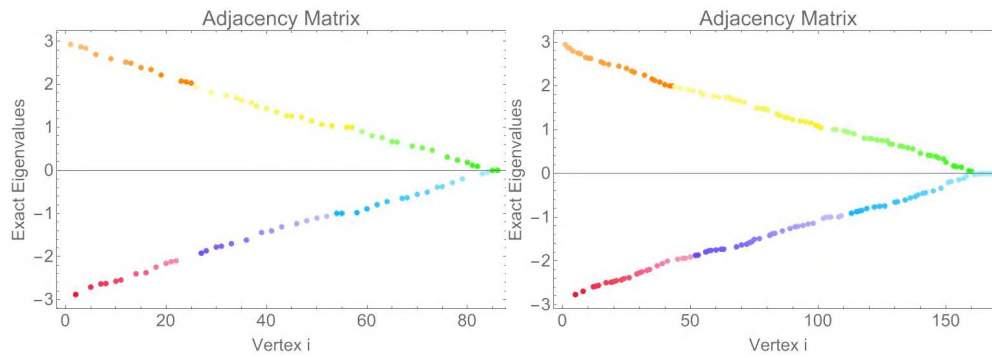


Figure 226 Exact Eigenvalue Adjacency Matrix of positive control (PK): 4h (left) and 24h (right)

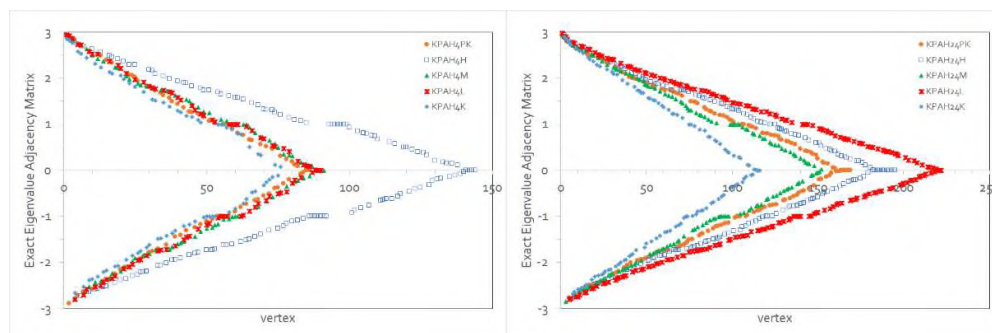


Figure 227 Exact Eigenvalue Adjacency Matrix at 4h (left) and 24h (right)

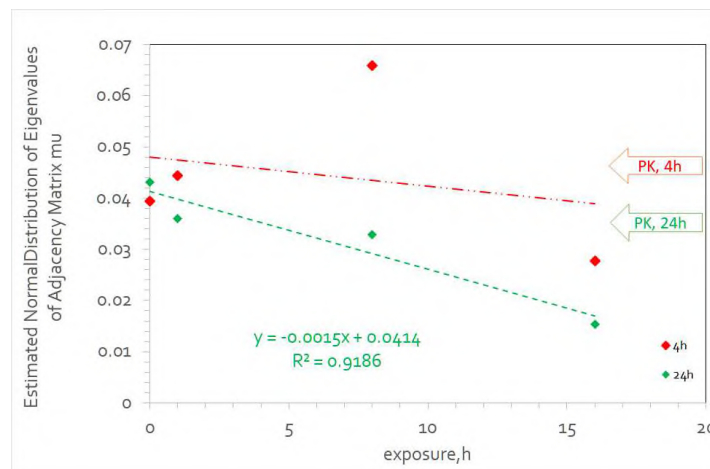


Figure 228 Estimated Normal Distribution mean (μ , μ_u) of exact Eigenvalues of Adjacency Matrix at 4h (left) and 24h (right)

Spectral entropy

This involves using the eigenvalues of a graph's Laplacian or adjacency matrix. The calculated eigenvalues are normalized and then by using a formula like the Shannon entropy the entropy is computed from the resulting probability distribution.

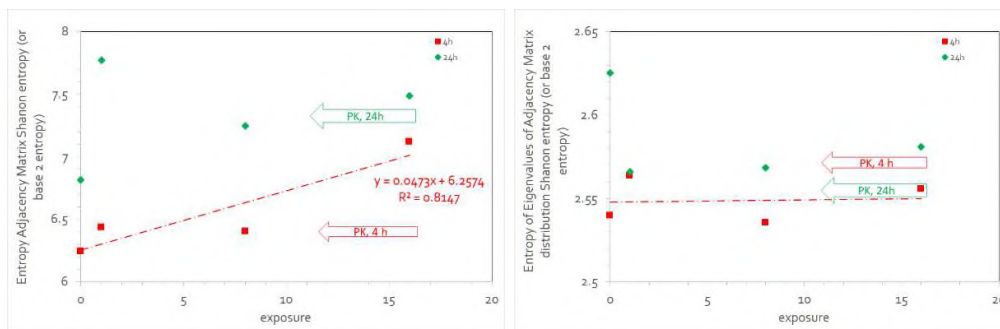


Figure 229 Shannon entropy (or base 2 entropy) of Adjacency Matrix (left) and Shannon entropy (or base 2 entropy) of Exact Eigenvalue of Adjacency Matrix distribution (right) at 4h (red, ■) and 24h (green, ◆)

The Estrada index of a graph

Table 23 The Estrada index of a graph

	exposure	The Estrada index of a graph
KPAH4K	0	231.53
KPAH4L	1	287.38
KPAH4M	8	290.73
KPAH4H	16	426.36
KPAH4PK		268.43
KPAH24K	0	376.93
KPAH24L	1	664.58
KPAH24M	8	465.17
KPAH24H	16	574.44
KPAH24PK		507.06

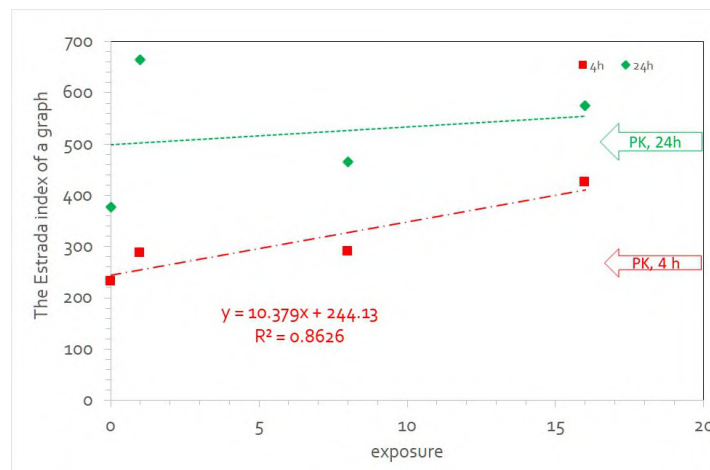


Figure 230 The Estrada index of a graph at 4h (red) and 24h (green).

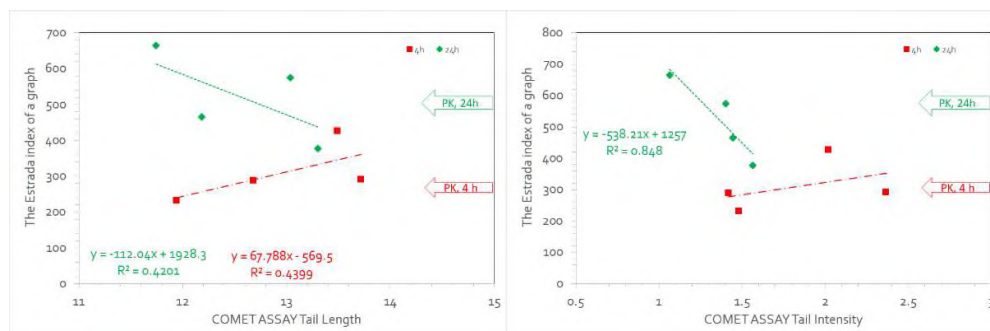


Figure 231 The Estrada index of a graph versus COMET ASSAY Tail Intensity (left) and versus COMET ASSAY Tail Length (right) at 4h (red) and 24h (green).

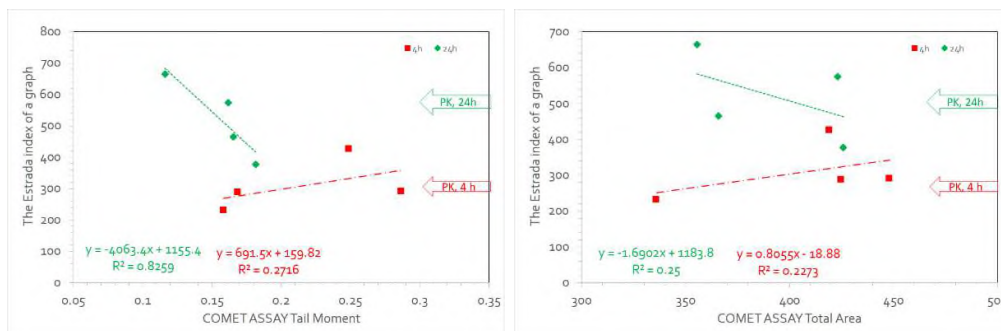


Figure 232 The Estrada index of a graph versus COMET ASSAY Tail Moment (left) and versus COMET ASSAY Total Area (right) at 4h (red) and 24h (green).

Weighted adjacency matrix

A weighted adjacency matrix is a square matrix used to represent a graph where each edge has a weight. For a graph with V vertices, the matrix is $V \times V$. If an edge exists between vertex i and vertex j , the entry in the matrix at row i , column j is the weight of that edge; if no edge exists, the value is zero. For an undirected graph, the matrix is symmetric, meaning the entry for (i,j) is the same as for (j,i) . If there is no edge between vertex i and vertex j , the value is 0.

Results

Weighted Adjacency matrix

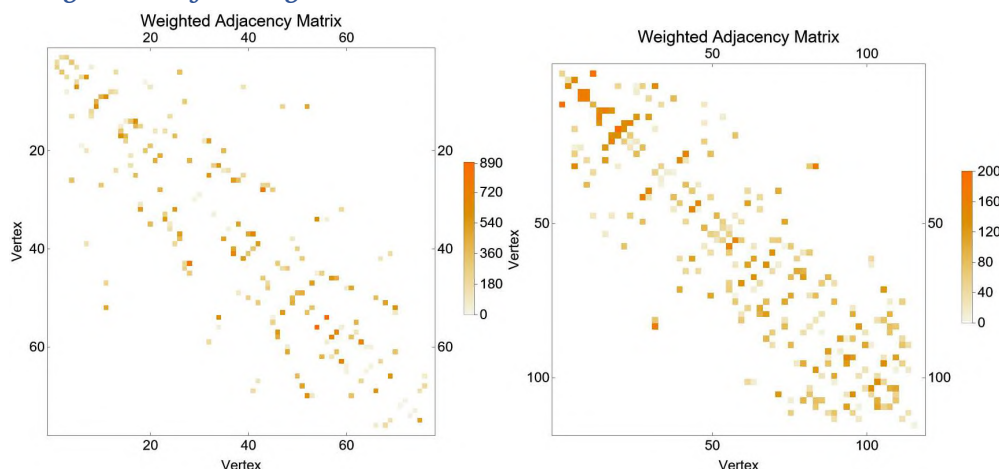


Figure 233 Weighted Adjacency matrix of control (K): 4h (left) and 24h (right)

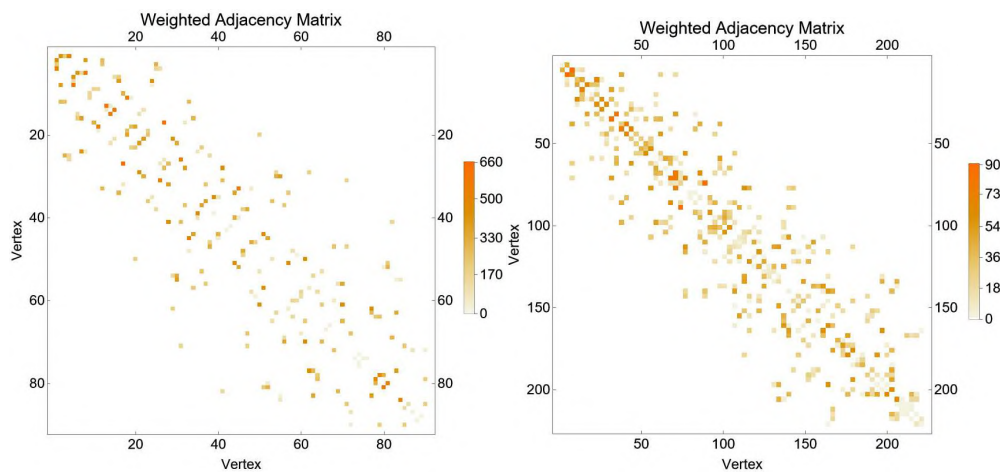


Figure 234 Weighted Adjacency matrix of low exposure (L): 4h (left) and 24h (right)

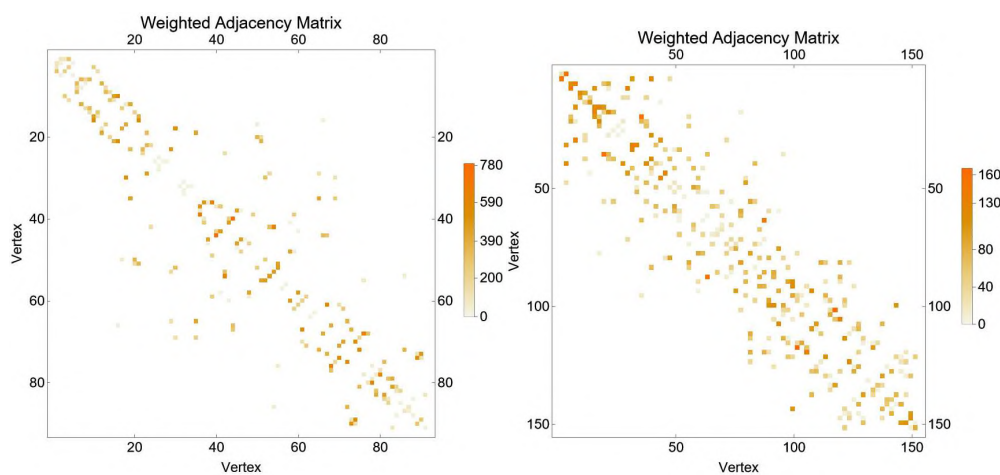


Figure 235 Weighted Adjacency matrix of middle exposure (M): 4h (left) and 24h (right)

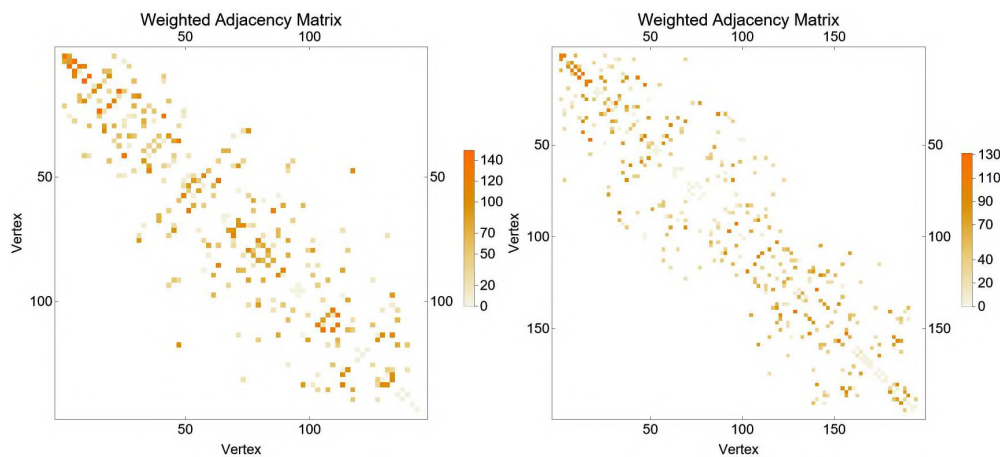


Figure 236 Weighted Adjacency matrix of high exposure (H): 4h (left) and 24h (right)

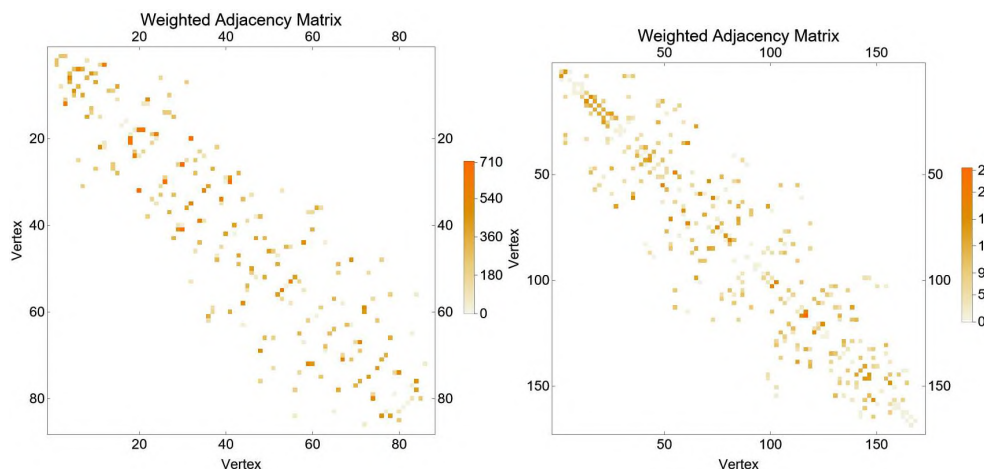


Figure 237 Weighted Adjacency matrix of positive control (PK): 4h (left) and 24h (right)

Exact Eigenvalue Weighted Adjacency Matrix

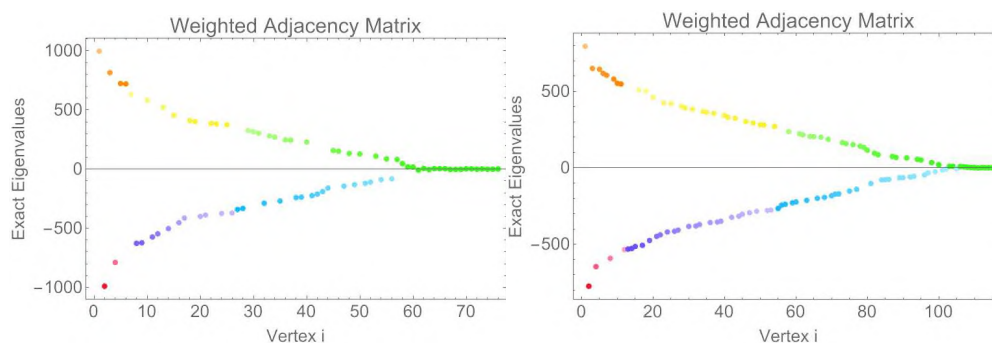


Figure 238 Exact Eigenvalue Weighted Adjacency Matrix of control (K): 4h (left) and 24h (right)

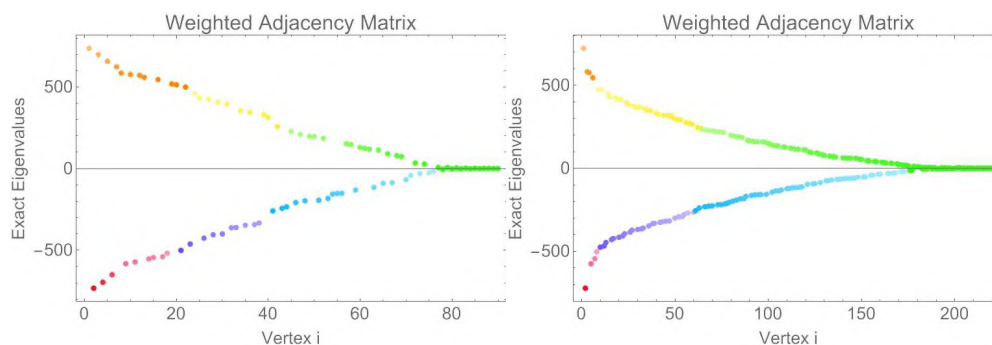


Figure 239 Exact Eigenvalue Weighted Adjacency Matrix of low exposure (L): 4h (left) and 24h (right)

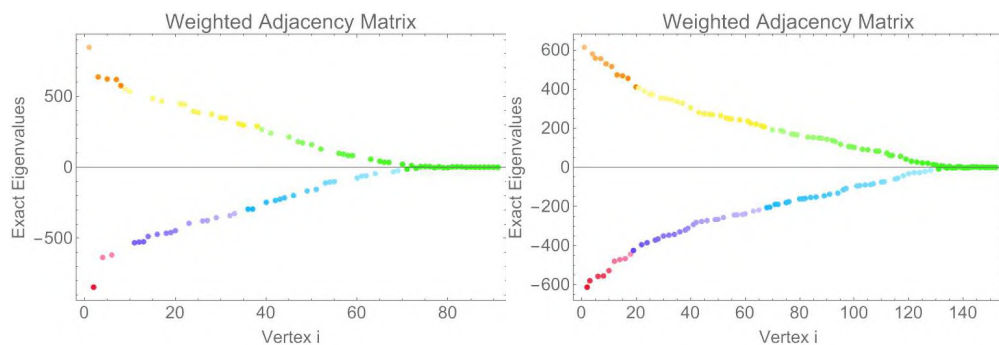


Figure 240 Exact Eigenvalue Weighted Adjacency Matrix of middle exposure (M): 4h (left) and 24h (right)

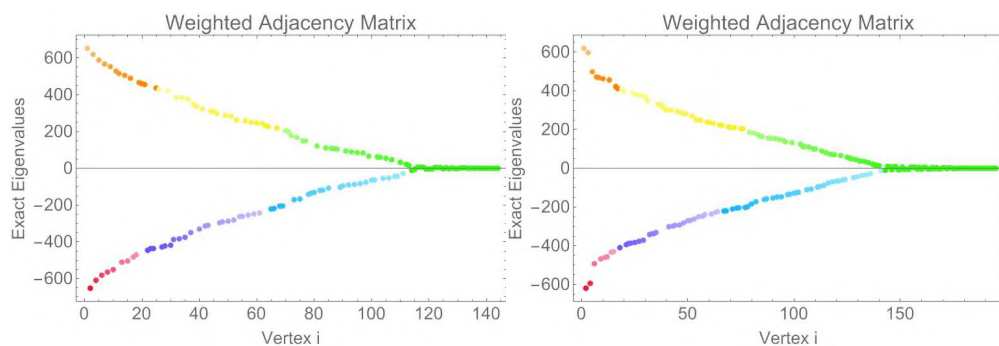


Figure 241 Exact Eigenvalue Weighted Adjacency Matrix of high exposure (H): 4h (left) and 24h (right)

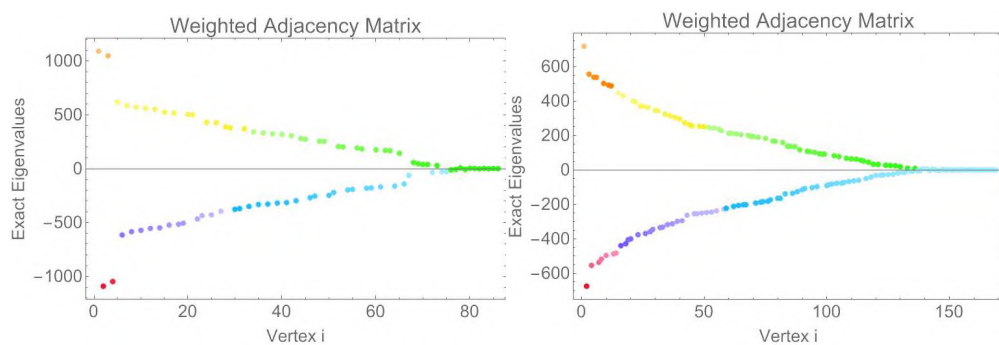


Figure 242 Exact Eigenvalue Weighted Adjacency Matrix of positive control (PK): 4h (left) and 24h (right)

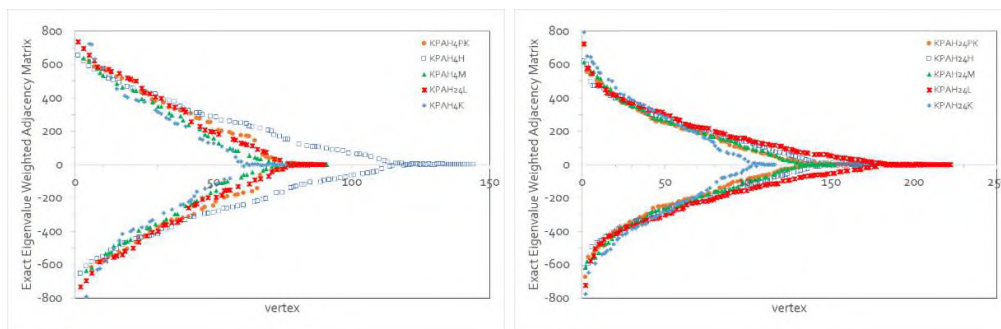


Figure 243 Exact Eigenvalue Weighted Adjacency Matrix at 4h (left) and 24h (right)

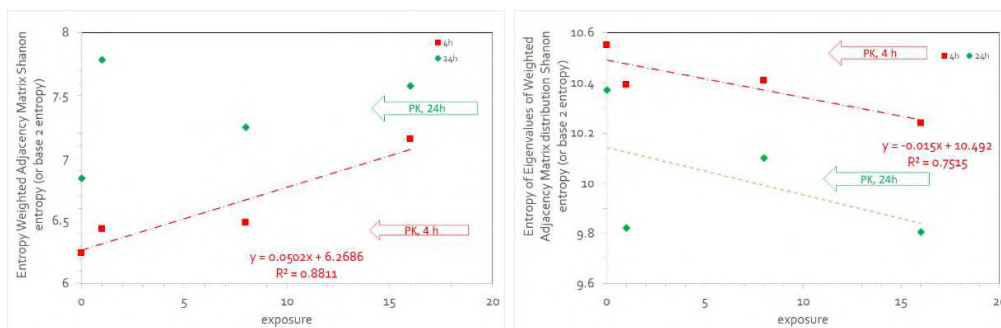


Figure 244 Shannon entropy (or base 2 entropy) of Weighted Adjacency Matrix (left) and Shannon entropy (or base 2 entropy) of Exact Eigenvalue Weighted Adjacency Matrix distribution (right) at 4h (red, ■) and 24h (green, ◆)

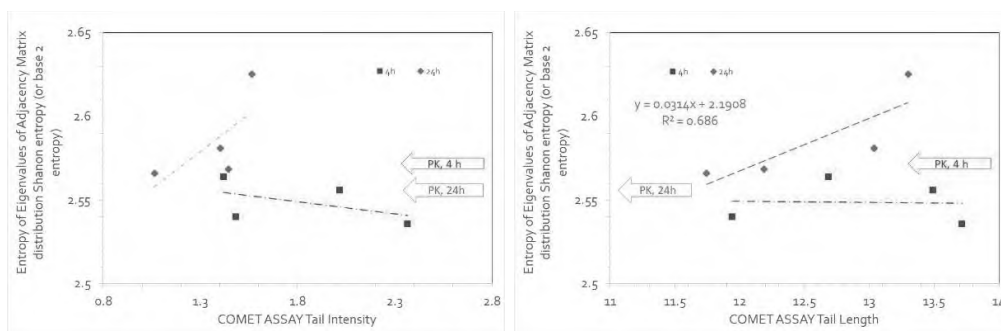


Figure 245 Correlations of Shannon entropy (or base 2 entropy) of Exact Eigenvalue Weighted Adjacency Matrix distribution and **Comet Assay Tail Intensity** (left) of Shannon entropy (or base 2 entropy) of Exact Eigenvalue Weighted Adjacency Matrix distribution and **Comet Assay Tail Length** (right) at 4h (red, ■) and 24h (green, ◆)

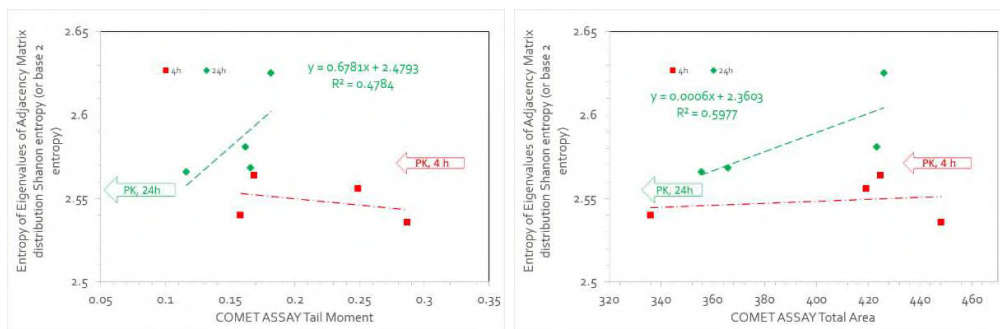


Figure 246 Correlations of Shannon entropy (or base 2 entropy) of Exact Eigenvalue Weighted Adjacency Matrix distribution and **Comet Tail Moment** (left) of Shannon entropy (or base 2 entropy) of Exact Eigenvalue Weighted Adjacency Matrix distribution and **Comet Assay Total Area** (right) at 4h (red, ■) and 24h (green, ◆)

Incidence Matrix

Incidence Matrix¹³⁴ is a $|V| \times |E|$ matrix that gives the vertex-edge incidence matrix of the graph H , with entries a_{ij} , the number of times that vertex v_i is incident with the edge e_j given by 1 if a vertex is incident to an edge and 0 otherwise.

Results

Incidence Matrix

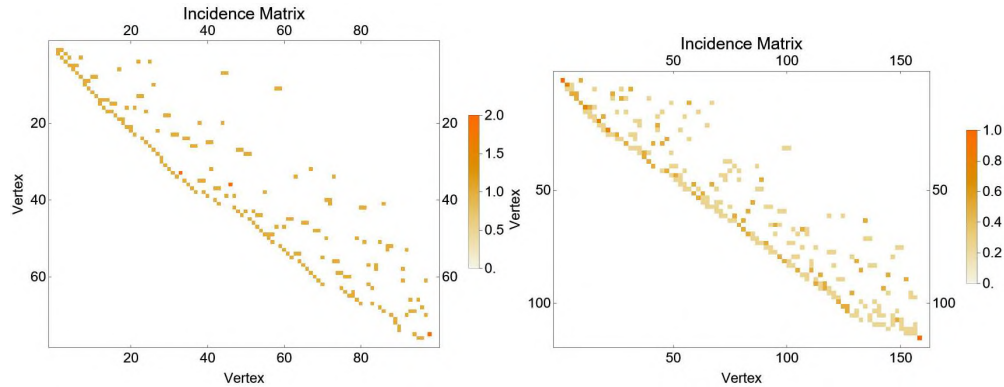


Figure 247 Incidence Matrix of control (K): 4h (left) and 24h (right)

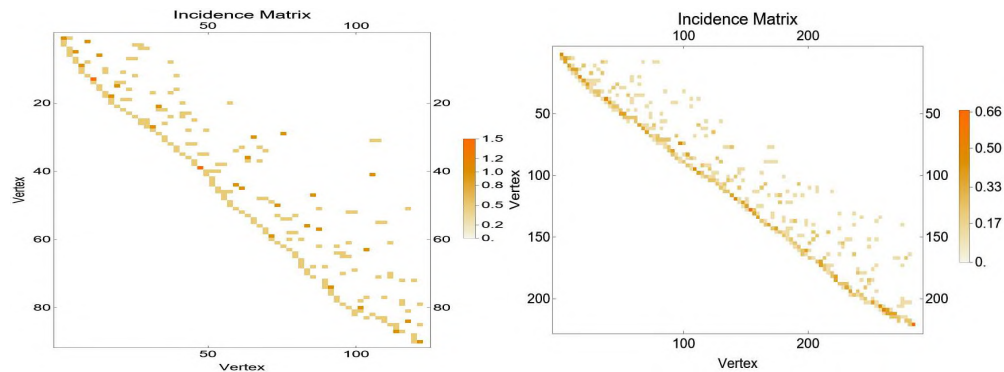


Figure 248 Incidence Matrix of low exposure (L): 4h (left) and 24h (right)

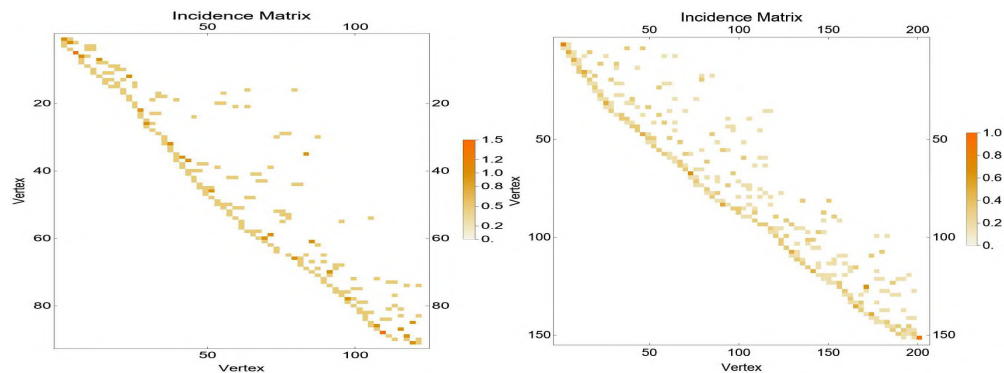


Figure 249 Incidence Matrix of middle exposure (M): 4h (left) and 24h (right)

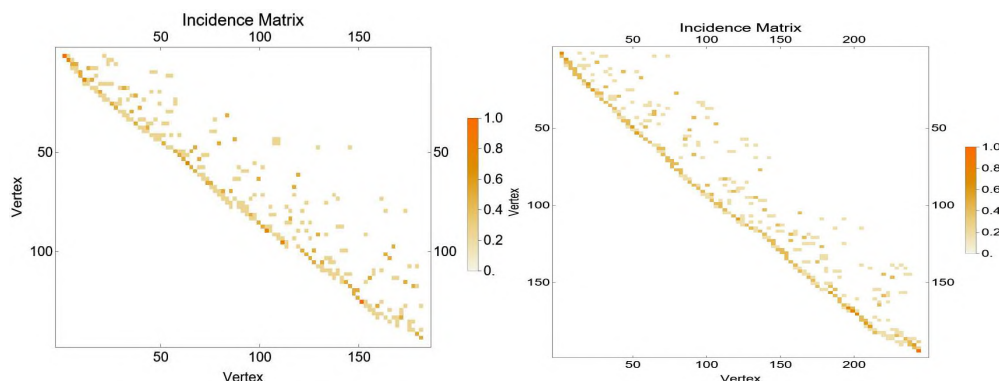


Figure 250 Incidence Matrix of high exposure (H): 4h (left) and 24h (right)

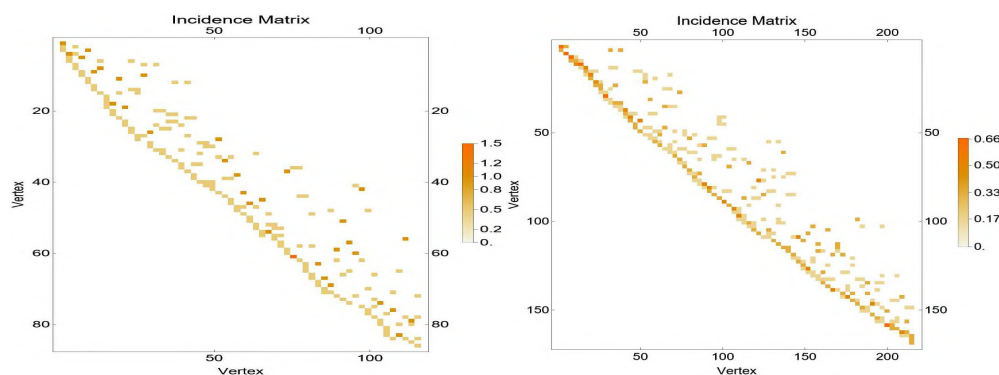


Figure 251 Incidence Matrix of positive control (PK): 4h (left) and 24h (right)

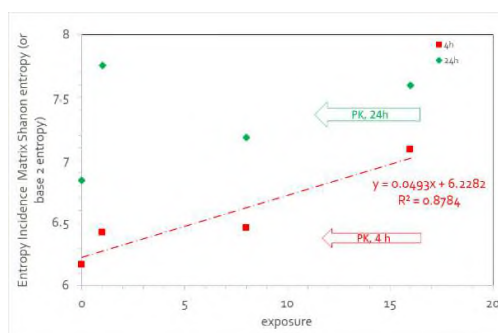


Figure 252 Shannon entropy (or base 2 entropy) of Incidence Matrix (right) at 4h (red, ■) and 24h (green, ◆)

Joint Degree Matrix

A Joint degree matrix^{135, 136} is a matrix used in graph theory that specifies the number of edges between vertices of a degree i and vertices of a degree j . It provides more detailed information than a simple degree sequence alone by capturing the degree correlations in a graph, showing how vertices of one degree connect to vertices of another degree. The Joint degree matrix is a symmetric matrix for undirected graphs, and its row/column sums reveal the total degree of each group of vertices. The joint degree distribution gives the

exact number of edges between vertices of degrees i and j , $m = |E|$, and the number of vertices $n = |V|$ in the graph.

Joint degree matrix represents structure of connections. It defines the specific count of edges between different “degree groups”, which are sets of vertices with the same degree. Joint degree matrix measures how vertices of different degrees are interconnected (degree correlations), a property that is important for understanding graph structures like assortativity (e.g., a preference for high-degree vertices to connect to other high-degree vertices). A Joint degree matrix implies the degree sequence and the number of vertices and edges in a graph.

Consider an undirected graph $G = (V, E)$, with $n = |V|$ vertices and $m = |E|$ edges. Let $\deg(v)$ be node v . Let V_k be the set of vertices with degree k , also called degree group k . The Joint Degree Matrix is defined as the number of edges connecting vertices of degree k with vertices of degree l and describes the degree correlations.

$$JDM(k, l) = \sum_{v \in V_k} \sum_{w \in V_l} l_{\{(v, w) \in E\}}$$

Results

Joint Degree Matrix

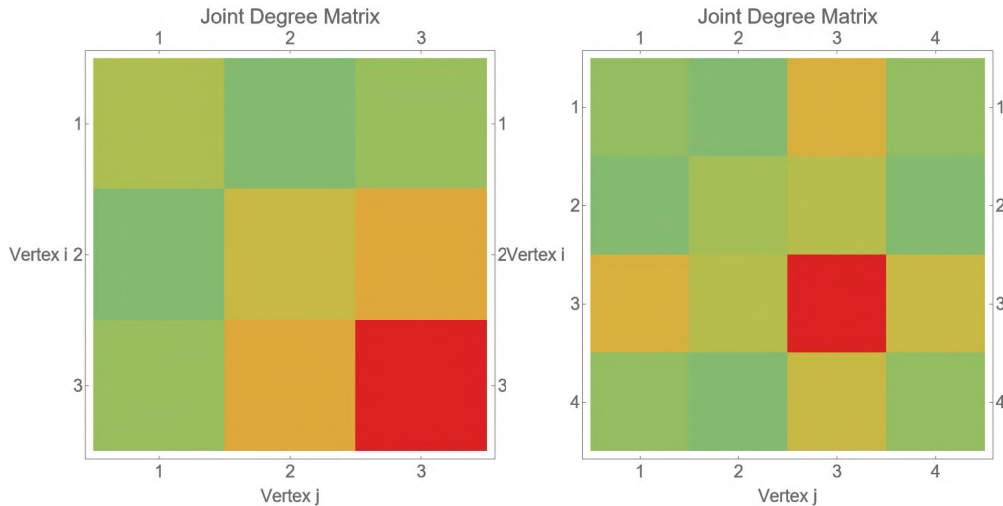


Figure 253 Joint Degree Matrix of control (K): 4h (left) and 24h (right)

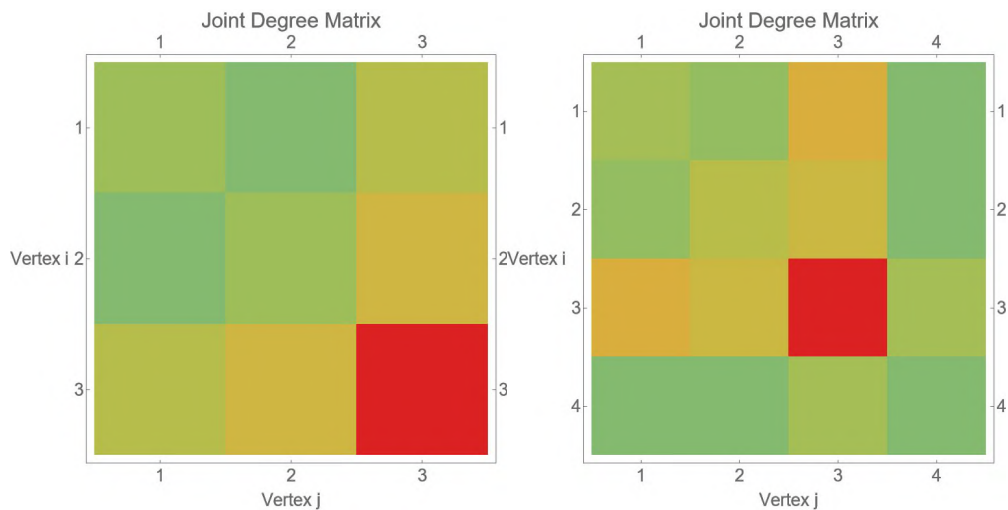


Figure 254 Joint Degree Matrix of low exposure (L): 4h (left) and 24h (right)

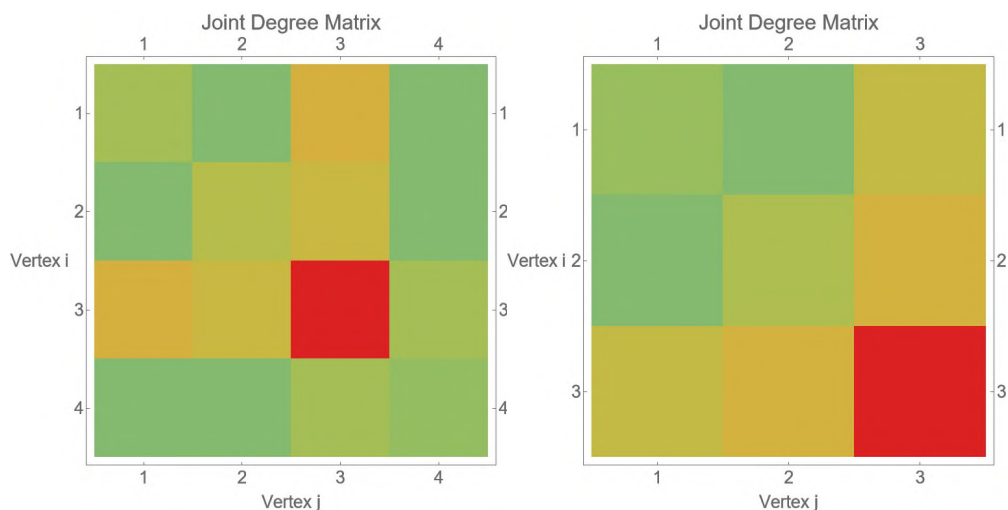


Figure 255 Joint Degree Matrix of middle exposure (M): 4h (left) and 24h (right)

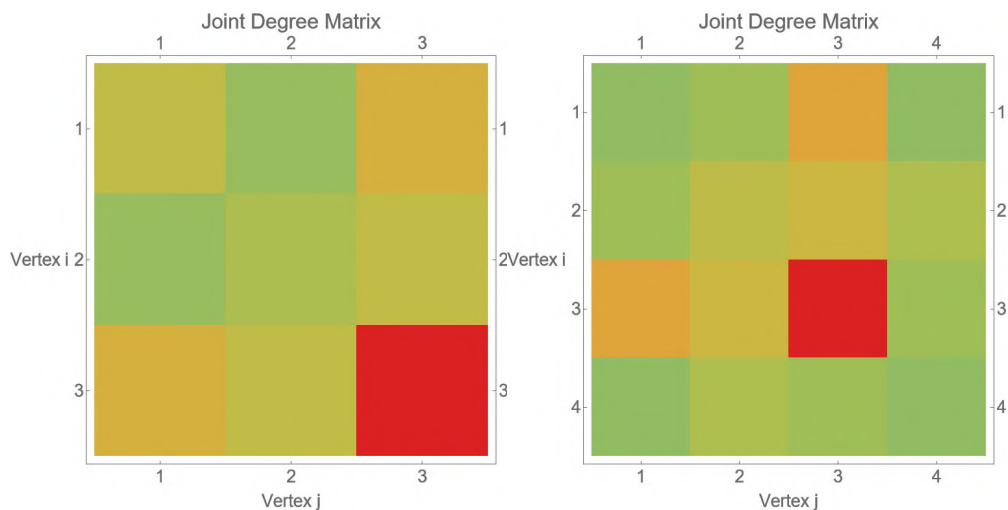


Figure 256 Joint Degree Matrix of high exposure (H): 4h (left) and 24h (right)

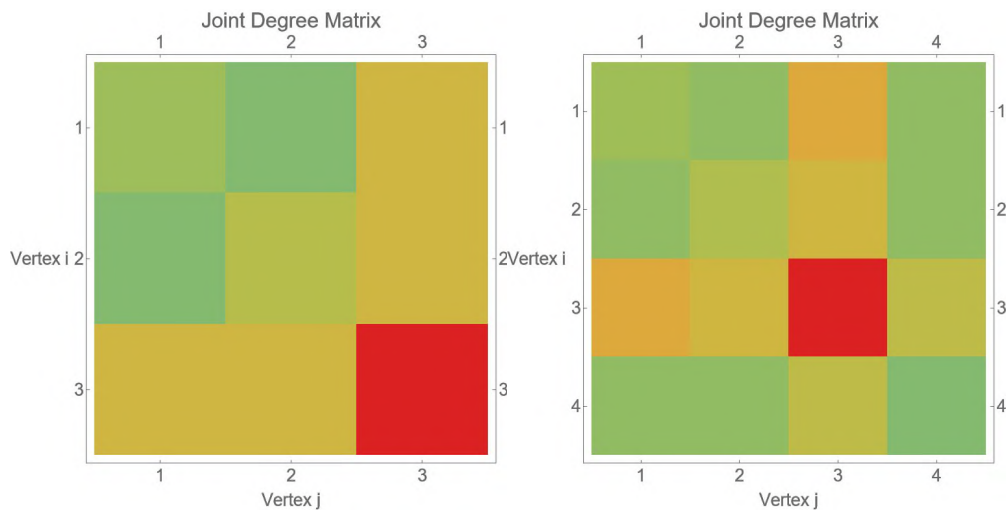


Figure 257 Joint Degree Matrix of positive control (PK): 4h (left) and 24h (right)

Eigenvalues of Joint Degree Matrix

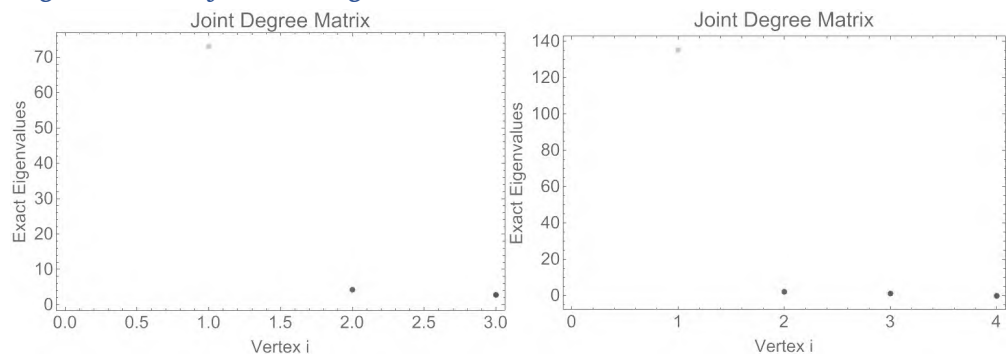


Figure 258 Eigenvalues of Joint Degree Matrix of control (K): 4h (left) and 24h (right)

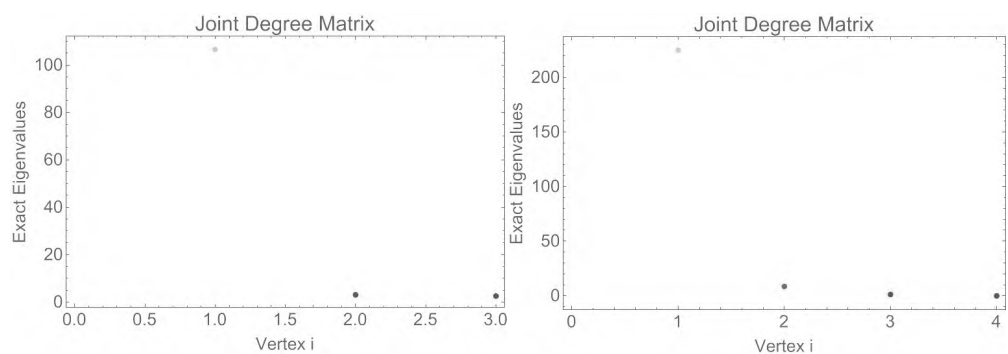


Figure 259 Eigenvalues of Joint Degree Matrix of low exposure (L): 4h (left) and 24h (right)

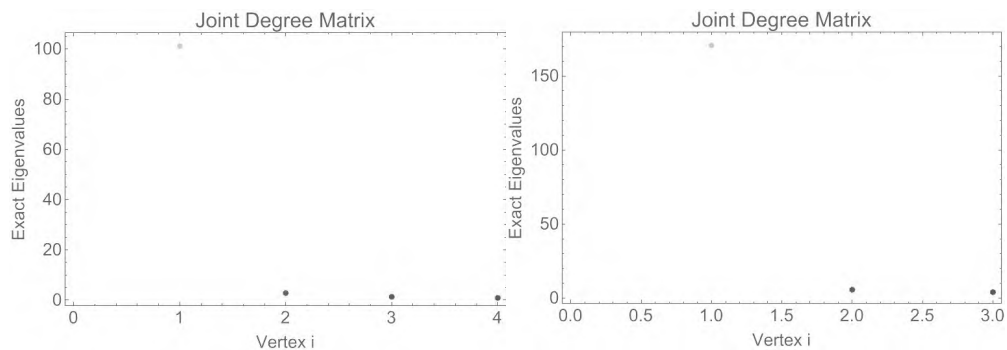


Figure 260 Eigenvalues of Joint Degree Matrix of middle exposure (M): 4h (left) and 24h (right)

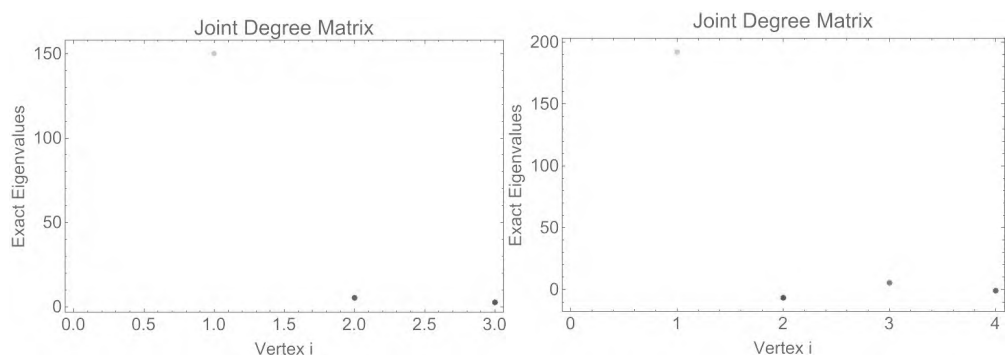


Figure 261 Eigenvalues of Joint Degree Matrix of high exposure (H): 4h (left) and 24h (right)

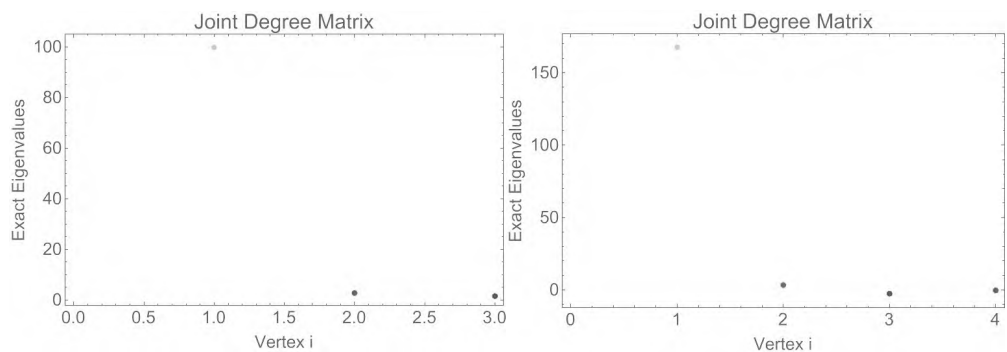


Figure 262 Eigenvalues of Joint Degree Matrix of positive control (PK): 4h (left) and 24h (right)

Table 24 Shannon entropy (or base 2 entropy) of Eigenvalues of Joint Degree Matrix distribution

	exposure	Shannon entropy (or base 2 entropy) of Eigenvalues of Joint Degree Matrix distribution
KPAH4K	0	6.64386
KPAH4L	1	7.56215
KPAH4M	8	7.45513
KPAH4H	16	7.64386
KPAH4PK		6.64386
KPAH24K	0	8.14386
KPAH24L	1	9.14386
KPAH24M	8	7.64386
KPAH24H	16	8.64386
KPAH24PK		8.64386

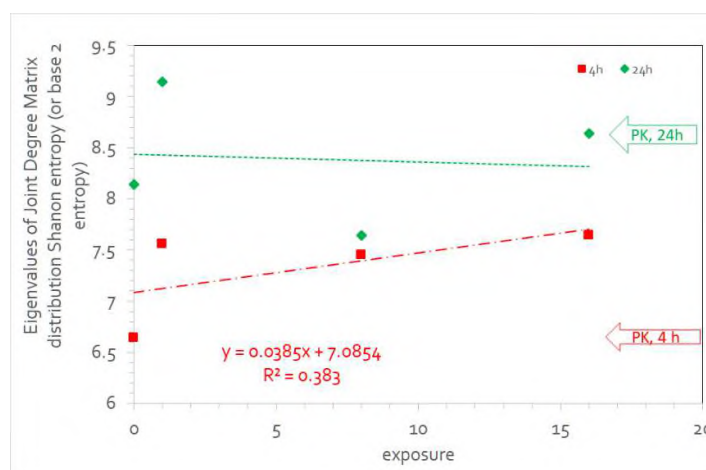


Figure 263 Shannon entropy (or base 2 entropy) of Eigenvalues of Joint Degree Matrix distribution at 4h (red, ■) and 24h (green, ◆)

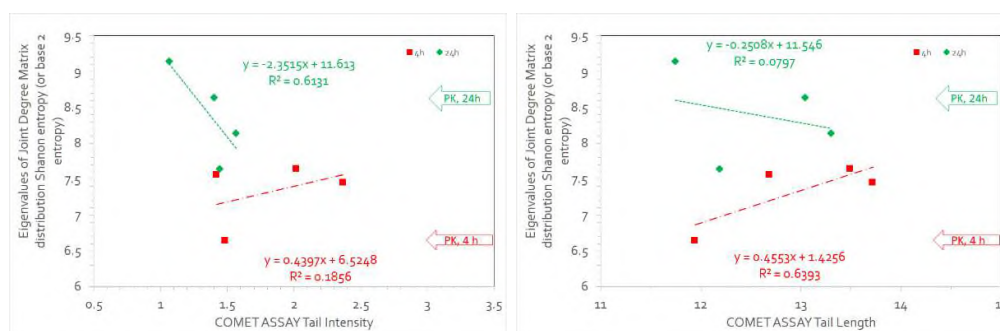


Figure 264 Correlations of Shannon entropy (or base 2 entropy) of Exact Eigenvalue Joint Degree Matrix distribution and **Comet Assay Tail Intensity** (left) and of Shannon entropy (or base 2 entropy) of Exact Eigenvalue Joint Degree Matrix distribution and **Comet Assay Tail Length** (right) at 4h (red, ■) and 24h (green, ◆)

Kirchhoff Matrix

The Kirchhoff Matrix is also known as the Laplacian matrix¹³⁷, also known as the Laplacian matrix, admittance matrix, or discrete Laplacian. Kirchhoff Matrix which takes a graph object as input and returns the $(L=D-A)$ matrix, where (D) is the degree matrix and (A) is the adjacency matrix.

Spectrum of a matrix is a set eigenvalue and their multiplicities. Let λ_i denote the eigenvalues of the Laplacian matrix. A graph, G , is connected if its second smallest eigenvalue is nonzero. $\lambda_2 > 0$ if and only if G is connected. The eigenvalue λ_2 is called the algebraic connectivity of a graph¹³⁸.

Entropy of a Kirchhoff matrix

To find the Entropy of a Kirchhoff matrix, you typically calculate the von Neumann entropy of the normalized Laplacian matrix's eigenvalues, denoted $S_L(G)$ using the formula

$$S_L(G) = \sum_{i=1}^n \left(\frac{\lambda_i}{d_i} \right) \log \left(\frac{\lambda_i}{d_i} \right)$$

where λ_i are the Laplacian eigenvalues¹³⁹ and d_i are the vertex degrees. This involves first constructing the Kirchhoff matrix (Laplacian matrix), finding its eigenvalues, and then applying the Shannon entropy formula. All the eigenvalues of Kirchhoff matrix are non-negative, less than or equal to the number of vertices, and less than or equal to twice the maximum vertex degree.

Results

Kirchhoff Matrix

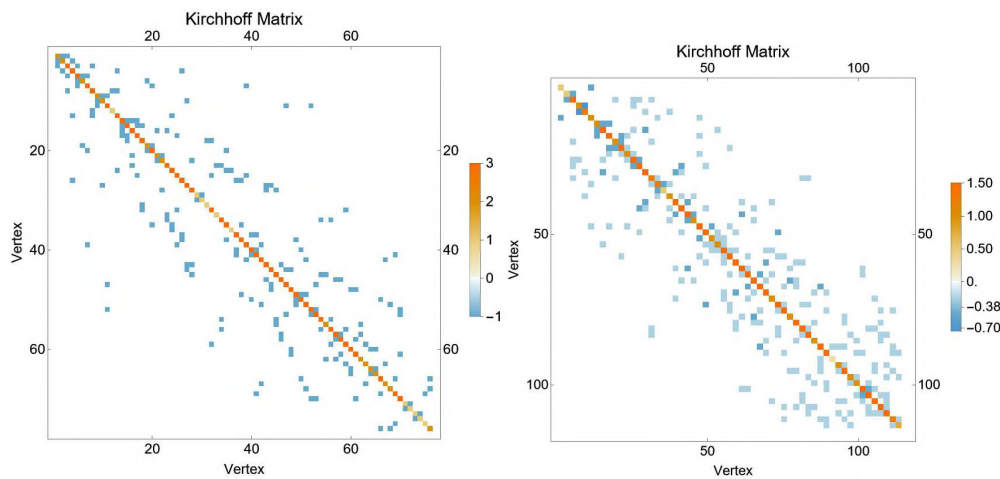


Figure 33 Kirchhoff Matrix of control (K): 4h (left) and 24h (right)

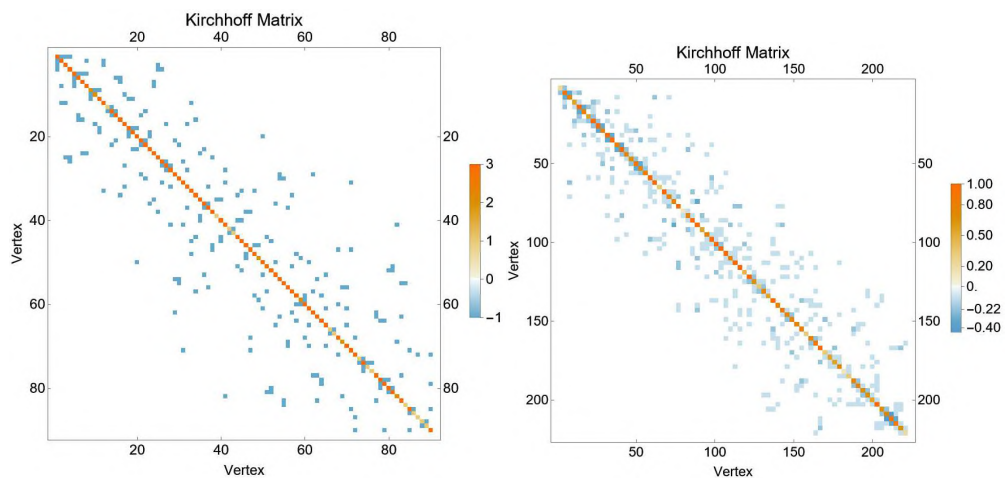


Figure 34 Kirchhoff Matrix of low exposure (L): 4h (left) and 24h (right)

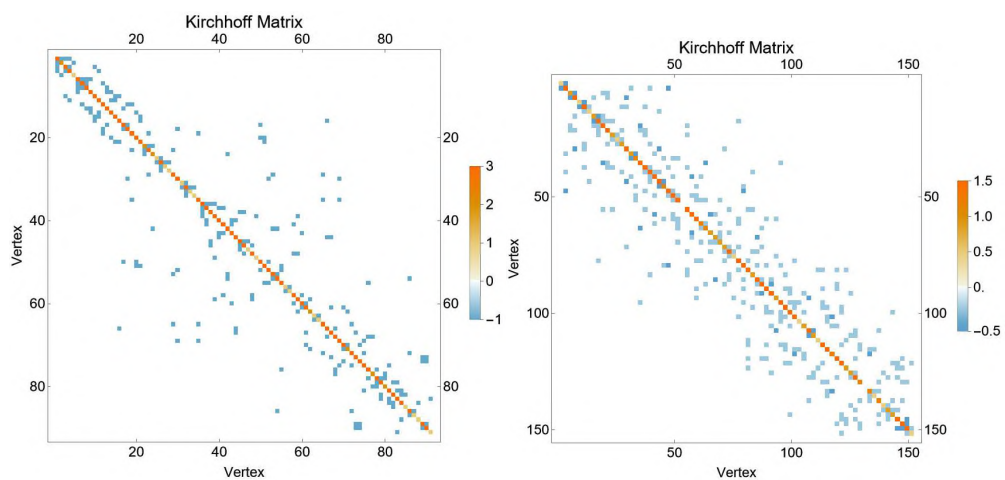


Figure 35 Kirchhoff Matrix of middle exposure (M): 4h (left) and 24h (right)

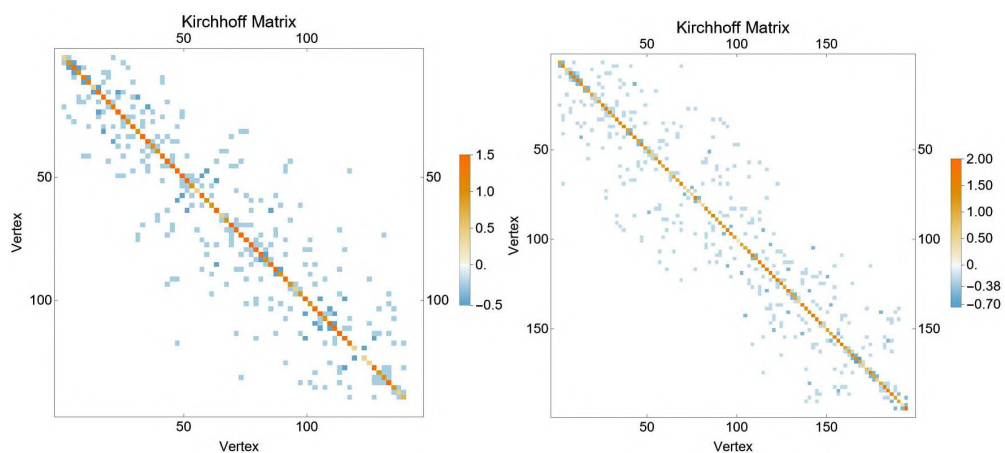


Figure 36 Kirchhoff Matrix of high exposure (H): 4h (left) and 24h (right)

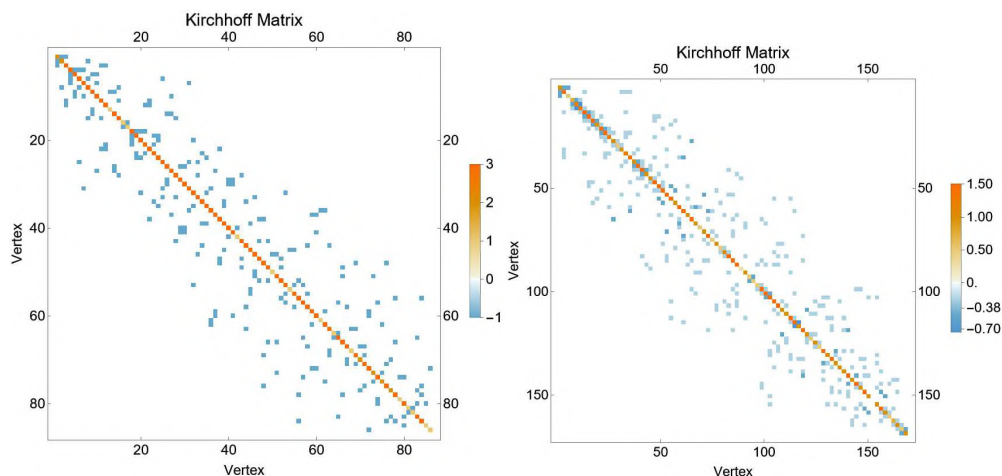


Figure 37 Kirchhoff Matrix of positive control (PK): 4h (left) and 24h (right)

Exact Eigenvalue Kirchhoff Matrix

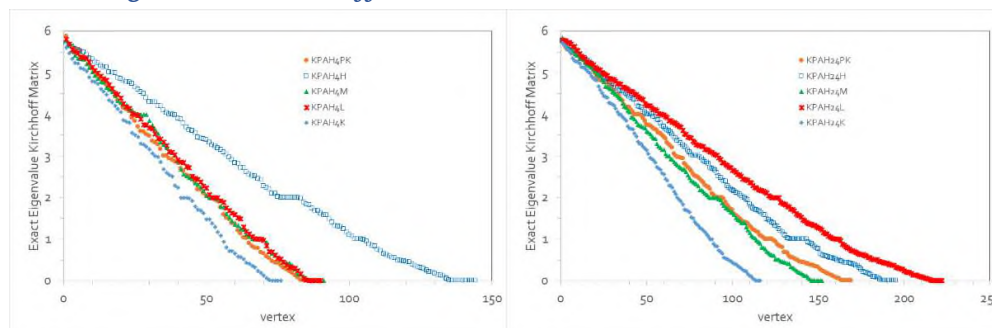


Figure 265 Exact Eigenvalue Kirchhoff Matrix at 4h (left) and 24h (right)

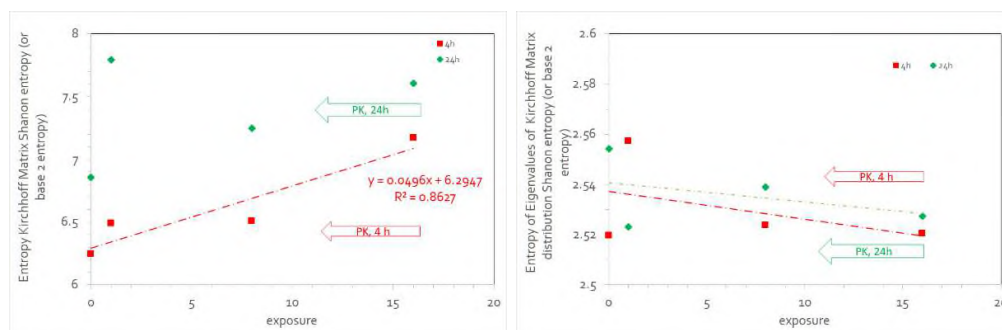


Figure 266 Shannon entropy (or base 2 entropy) of Kirchhoff Matrix (left) and Shannon entropy (or base 2 entropy) of Exact Eigenvalue Kirchhoff Matrix distribution (right) at 4h (red, \blacksquare) and 24h (green, \blacklozenge)

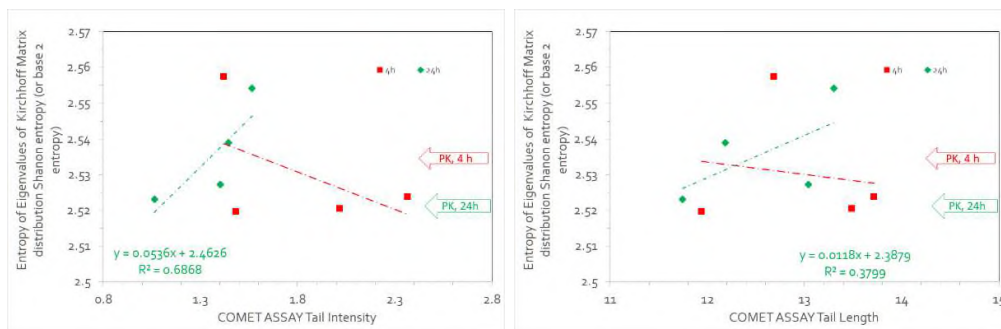


Figure 267 Correlations of Shannon entropy (or base 2 entropy) of Exact Eigenvalue Kirchoff Matrix distribution and **Comet Assay Tail Intensity** (left) and of Shannon entropy (or base 2 entropy) of Exact Eigenvalue Kirchoff Matrix distribution and **Comet Assay Tail Length** (right) at 4h (red, ■) and 24h (green, ◆)

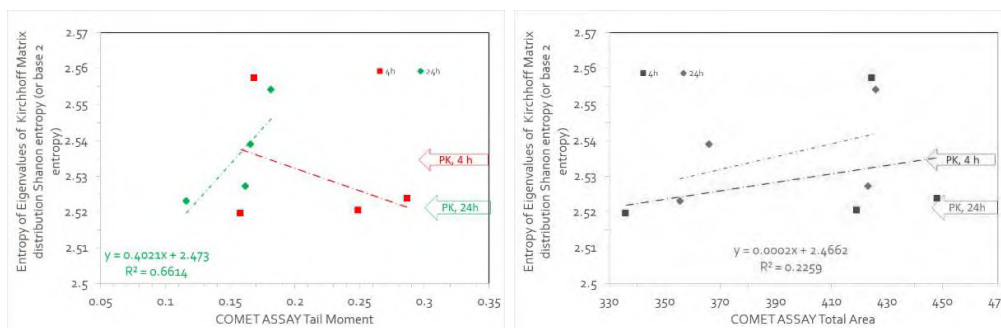


Figure 268 Correlations of Shannon entropy (or base 2 entropy) of Exact Eigenvalue Kirchoff Matrix distribution and **Comet Tail Moment** (left) and of Shannon entropy (or base 2 entropy) of Exact Eigenvalue Kirchoff Matrix distribution and **Comet Assay Total Area** (right) at 4h (red, ■) and 24h (green, ◆)

Radiality centrality

Radiality (integration centrality) is a graph centrality measure that scores vertices based on how close they are to all other vertices relative to the graph's diameter. A node with high radiality is close to every other node in its accessible neighbourhood, suggesting a central position, while a node with low radiality is more peripheral. This is calculated using a formula involving the shortest path distances between a node and all others, and can be computed even for graphs that are not strongly connected.

Radiality is based on „reverse distance“, which is the graph's diameter plus one, minus the shortest path distance between two vertices. By summing the reverse distances from a node to all others, you get a high score if the node is short paths away from many other vertices, especially compared to the overall graph size.

The formula for radiality $C_{rad}(i)$ of a node i is the sum of (diameter+1-distance(i,j)) for all other vertices (j) in the graph, divided by the number of other vertices ($n-1$). Key features: It can be calculated for weakly connected or even unconnected graphs, unlike closeness centrality, which is limited to strongly connected ones. It provides a measure of a node's ability to reach other vertices in the graph.

In a graph, vertices with significantly higher radiality than the graph average would be crucial for regulating other vertices, as it can be reached quickly from many other parts of the graph.

Basically, as the diameter is the maximal possible distance between vertices, subtracting systematically from the diameter the shortest paths between the node v and its neighbours will give high values if the paths are short and low values if the paths are long. Radiality Centrality will give high centralities to vertices that are a short distance to every other vertex in its reachable neighbourhood compared to its diameter. Overall, if the radiality is high this means that, with respect to the diameter, the node is generally closer to the other vertices, whereas, if the radiality is low, this means that the node is peripheral. Also here, “high” and “low” values are more meaningful when compared to the average radiality of the graph G calculated by averaging the radiality values of all vertices in the graph. As for the closeness, the radiality value should be considered as an “average tendency to node proximity or isolation”, not definitively informative on the centrality of the individual node.

Results

Table 25 Radiality centrality

	exposure	Sahnnon Entropy of Radiality centrality	Sahnnon Entropy of normalized Radiality centrality
KPAH4K	0	6.1297	4.7014
KPAH4L	1	6.3361	5.0145
KPAH4M	8	6.3328	5.1681
KPAH4H	16	6.9695	5.5812
KPAH4PK		6.3549	4.6398
KPAH24K	0	6.7818	5.1147
KPAH24L	1	7.7113	5.8456
KPAH24M	8	7.1451	5.4560
KPAH24H	16	8.1643	6.5241
KPAH24PK		4.2474	5.4933

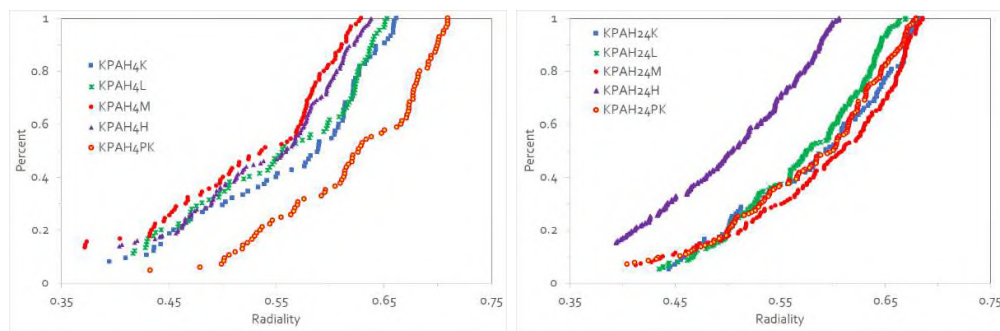


Figure 269 Radiality Centrality at 4h (left) and 24h (right)

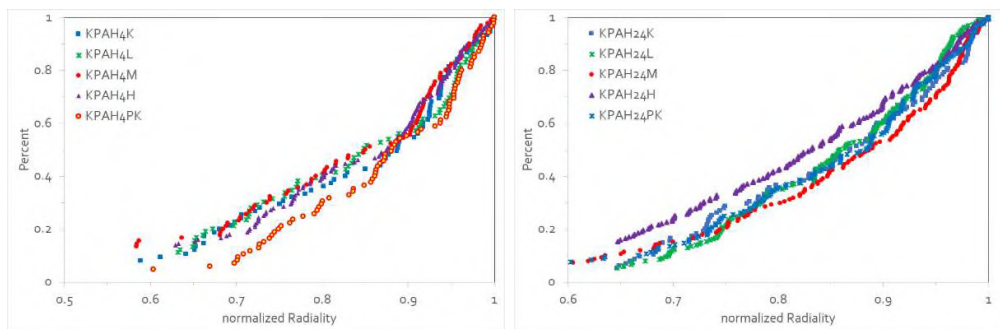


Figure 270 Normalized Radiality Centrality at 4h (left) and 24h (right)

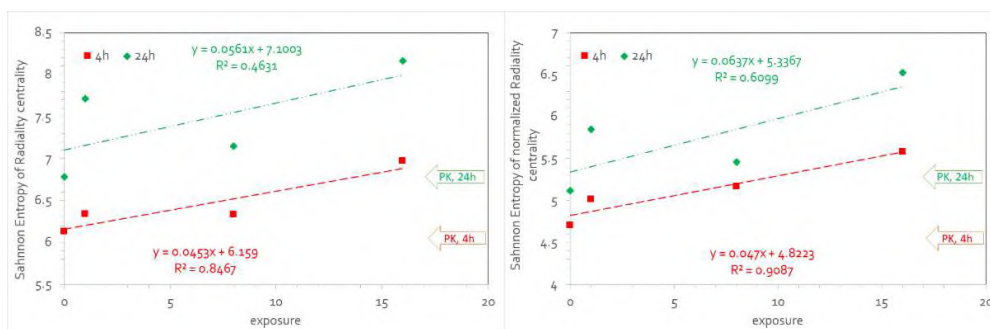


Figure 271 Shannon Entropy of Radiality Centrality (left) and Shannon Entropy of normalized Radiality centrality (right) at 4h (red, ■) and 24h (green, ◆)

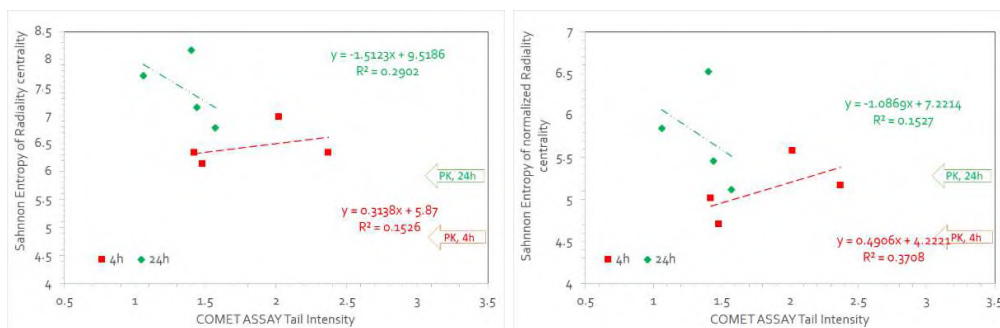


Figure 272 Shannon Entropy of Radiality Centrality (left) and Shannon Entropy of normalized Radiality centrality (right) versus COMET ASSAY Tail Intensity at 4h (red, ■) and 24h (green, ◆)

Voronoi entropy for a nearest-neighbour graph

A Voronoi diagram is created from a set of points (seeds, i.e., comets), and each seed is surrounded by a corresponding comets including all points closer to that seed than any other. A nearest-neighbour graph connects each point to its nearest neighbours. A Voronoi diagram can come from this graph, and its properties are used to measure order.

The Voronoi entropy for a nearest-neighbour graph refers to the Voronoi entropy, a measure of the disorder or orderliness of the point distribution in a Voronoi diagram derived from that graph. It is calculated by analysing the statistical distribution of edges (n) for each cell, using the formula for Shannon entropy:

$$S = - \sum_n p_n \log_2 (p_n)$$

where p_n is the probability of a cell having n edges. A perfectly ordered structure, where all cells are the same, has zero Voronoi entropy, while a completely random distribution of points in 2D has a maximum entropy value of about 1.71. A value in-between indicates a certain degree of order or disorder.

First approach is by given query point q , scan through all n sample points with the r shortest distances seen and computing (squared) distances to q . The second approach is scan through all n sample points with the k -nearest neighbour seen and computing (squared) distances to q (an order- k Voronoi diagram with a cell for each possible k -nearest neighbours)

Results

Table 26 Voronoi entropy for a r-nearest-neighbour graph and Voronoi entropy for a k-nearest-neighbour graph

	exposure	Voronoi entropy for a r-nearest-neighbour graph	Voronoi entropy for a k-nearest-neighbour graph
KPAH4K	0	5.6595	7.9551
KPAH4L	1	5.8051	8.2953
KPAH4M	8	5.9764	8.3175
KPAH4H	16	6.7105	9.2395
KPAH4PK	24	5.7071	8.2039
KPAH24K	0	6.1291	8.8054
KPAH24L	1	7.7284	10.1077
KPAH24M	8	6.8090	9.3480
KPAH24H	16	7.3132	9.8477
KPAH24PK	24	7.1148	9.5607

Voronoi entropy for a r-nearest-neighbour graph

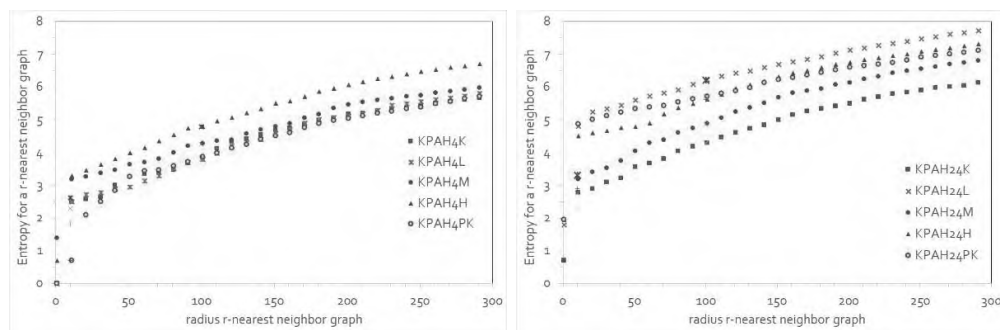


Figure 273 Voronoi entropy for a r-nearest-neighbour graph: 4h (left) and 24h (right).

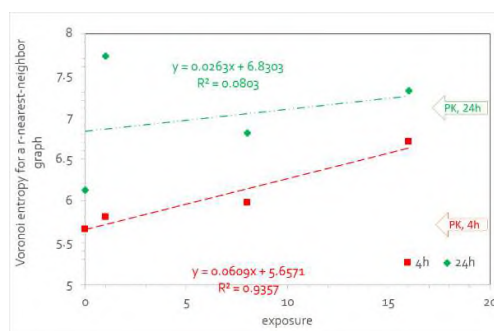


Figure 274 Voronoi entropy for a k-nearest-neighbour graph: 4h (red symbols) and 24h (green symbols).

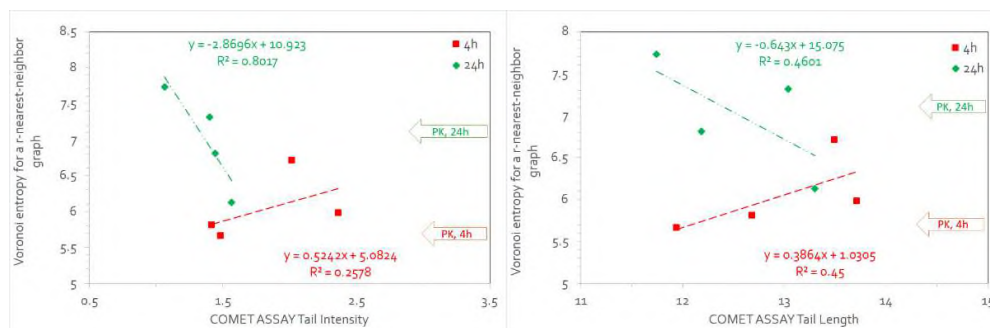


Figure 275 Voronoi entropy for a k-nearest-neighbour graph versus COMET ASSAY Tail Intensity (left) and versus COMET ASSAY Tail Length: 4h (red symbols) and 24h (green symbols).

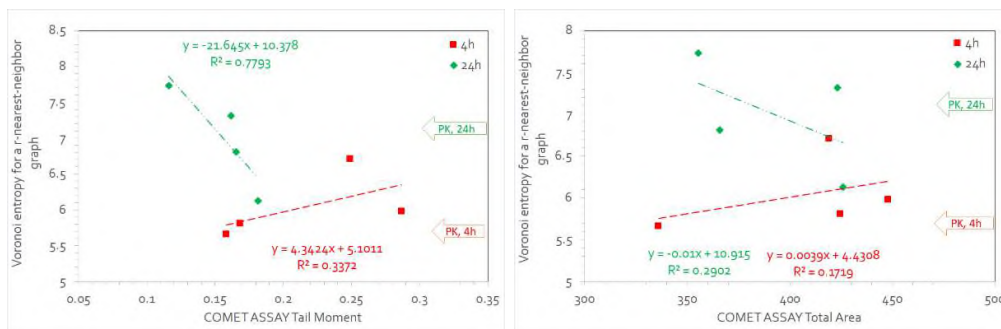


Figure 276 Voronoi entropy for a k-nearest-neighbour graph versus COMET ASSAY Tail Moment (left) and versus COMET ASSAY Total Area: 4h (red symbols) and 24h (green symbols).

Voronoi entropy for a k-nearest-neighbour graph

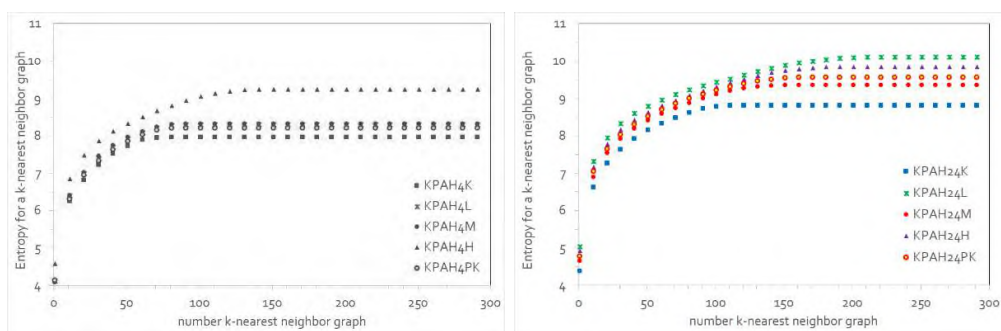


Figure 277 Voronoi entropy for a k-nearest-neighbour graph: 4h (left) and 24h (right).

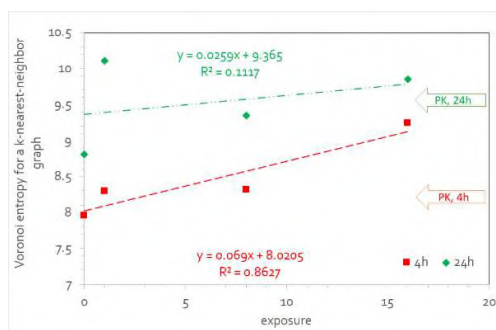


Figure 278 Voronoi entropy for a k-nearest-neighbour graph: 4h (red symbols) and 24h (green symbols).

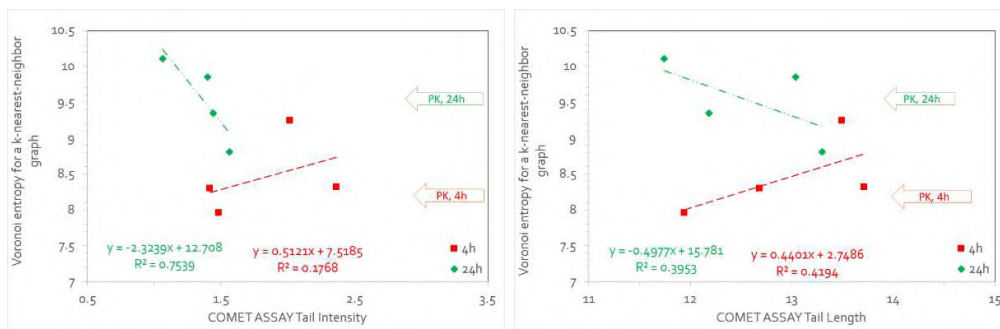


Figure 279 Voronoi entropy for a k-nearest-neighbour graph versus COMET ASSAY Tail Intensity (left) and versus COMET ASSAY Tail Length: 4h (red symbols) and 24h (green symbols).

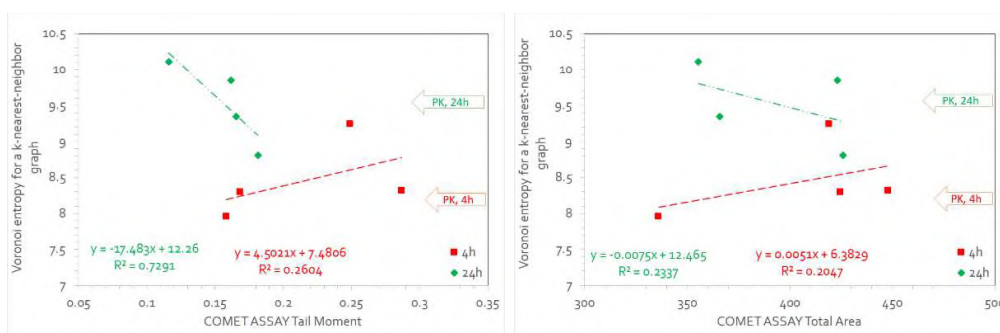


Figure 280 Voronoi entropy for a k-nearest-neighbour graph versus COMET ASSAY Tail Moment (left) and versus COMET ASSAY Total Area: 4h (red symbols) and 24h (green symbols).

Persistent homology

Representing the image as a graph where pixels are nodes and edges represent connectivity enables topological invariants to be computed with graph theory tools. Persistent homology is an emerging tool in studying graphs^{140, 141, 142}. Persistent homology can derive information for the visualization of graphs. Persistent homology has recently been a strong descriptor of graphs¹⁴³, and it has a few key qualities that make it ideal for applications. The calculated, from a graph without the need to select parameters, Persistent homology features are in the form of 0-dimensional homological groups. The Persistent homology features can be measured as persistence. Persistent homology features provide compressed description of the graph using a persistence barcode.

Persistent homology can be understood from three parts.

First, it is the relation between a graph and a simplicial complex. Mathematically, a graph, which consists of only nodes (0-simplexes) and edges (1-simplexes), is a simplicial complex. A simplicial complex K can be viewed as a set of simplexes that satisfy two conditions. First, any face of a simplex from K is also in K . *Secondly*, the intersection of any two simplexes in K is empty or a shared face¹⁴⁴. Other than 0- and 1-simplexes, it also includes 2-simplexes (solid triangles), 3-simplexes (tetrahedrons), and other higher-dimensional components.

Second, it is about geometric measurements and topological invariants. In persistent homology, the data is characterized by Betti numbers, including β_0 , β_1 , β_2 and higher order topological invariants^{145, 146, 147}. These measurements differ significantly from previous geometric measurements, like distances, angles, areas, etc.

Third, it is the difference between single scale model and multi-scale representation. Essentially, a series of related simplicial complexes is considered in persistent homology and they provide a multiscale representation that balances geometry and topology.

Geometrically, β_0 indicates the number of connected components, β_1 corresponds to the number of circles, rings or loops, and β_2 represents the number of voids or cavities. The key concept in persistent homology is the filtration. For example, given a point cloud data, we can associate each point with an identical-sized sphere and assign its radius as the filtration parameter. As the filtration value is increased, these spheres will systematically enlarge and merge with each other to form simplexes. Roughly speaking, an edge between two points is formed when the two corresponding spheres overlap. A triangle is formed when each of two spheres (of the three corresponding spheres from triangle vertices) overlap. A tetrahedron is formed when each three spheres (of the four corresponding spheres from tetrahedron vertices) overlap. At each filtration value, all the simplexes, i.e., vertices, edges, triangles, tetrahedrons, form a simplicial complex. From it, topological invariants, i.e., Betti numbers, can be calculated. The persistent homology hierarchically increases the complexity in data representation by systematically incorporating higher order simplices as the filtration proceeds. This enables a multiscale representation of topological invariants from simplicial complexes. In this way, a systematic variation of the filtration parameter leads to a series of simplicial complexes at different scales. Some topological invariants persist longer in these simplicial complexes, while

others disappear quickly as the filtration value is increased. The length of the β_1 bar defines the “lifespan” of the topological invariants (circles, loops, etc) and provides a natural geometric measurement. The lifespan, known as the persistence, measures how “large” are the circles, loops and voids in the system. We denote a filtration value at which a topological invariant formed or killed as birth time and death time respectively. In this way, each topological invariant has a “lifespan” defined by its birth and death time. Essentially, the lifespan provides a geometric measurement of the topological invariant. If we use a one-dimensional bar, which starts at a birth time and ends at a death time, to represent each homology generator, a barcode representation is generated. A general idea (not always confirmed) in the study of topological persistence is that significant topological attributes must have long life-times, and topological features with short life-times are considered “noise”.

Persistent homology has been identified as one of the most promising mathematical tools available for data analysis¹⁴⁸. For example, PH has been employed analysis of percolating surfaces and porous materials¹⁴⁹. Further, Persistent homology has been an effective method for analysing datasets related to medicine such as the homology of brain network^{150, 151, 152}, brain artery tree structure¹⁵³, orthodontics¹⁵⁴, RNA¹⁵⁵.

Homology is an algebraic tool that for our purposes constructs groups $H_n(S)$ that describes the holes lying in the simplicial complex (Mesh Regions, S). The generators for these groups are essentially the holes in the data. H_0 describes 0-dimensional “holes” because it is generated by the connected parts of S , H_1 describes 1-dimensional holes because it is generated by loops of S . Persistent Homology^{156, 157, 158, 159} constructs a sequence of meshes (known as a filtration) around the data (here COMET ASSAY positions derived by Morphological Graph) and analyses these to find the “holes”. The meshes are formed by growing balls of increasing radii around the data points (nucleus positions) and placing edges, triangles, etc when they overlap pairwise.

Even though persistent entropy has been a powerful tool for the characterization of “topological disorder”, its physical meaning is usually unclear, thus hinders its further application in chemical, physical and biological systems. In this paper, we use a Shannon persistent entropy based on the Betti numbers and probability distribution.

Betti numbers are the numbers of holes in a space with different dimensions. Persistent homology is one of the most used techniques for computing the topological invariants of a topological space. It returns a parametrized version of the Betti numbers: the Betti barcodes. Diaz et al., defined an entropy based on the persistent barcode. The persistent entropy, based on the persistent barcode, has been proposed^{160, 161, 162, 163} to measure the system disorder. For the k -th dimensional barcodes, it is defined as,

$$S_{Sh,k} = \sum_i^{N_k} -p_{k,j} \times \log_2(p_{k,j})$$

Persistent Life=Death-Birth

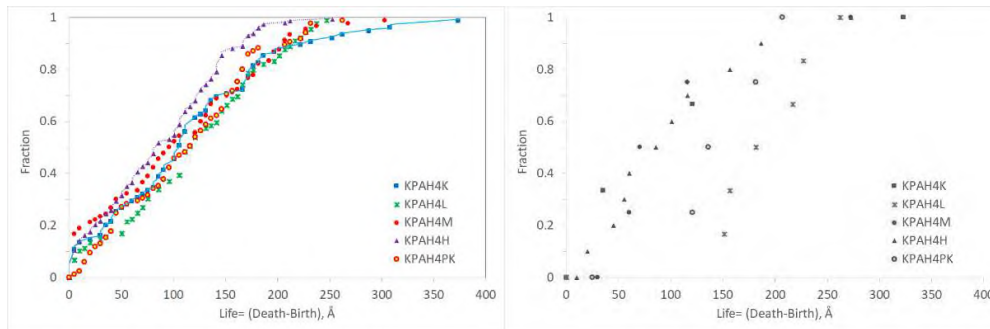


Figure 281 Fraction of H0 persistent Life (=Death-Birth) (Å) (left) and Fraction of H1 persistent Life (=Death-Birth) (Å) (right)

Persistence diagrams: Death vs. Birth

The way to read a persistence diagram is to note that the further away a point is from the diagonal line $y = x$, the longer it has persisted.

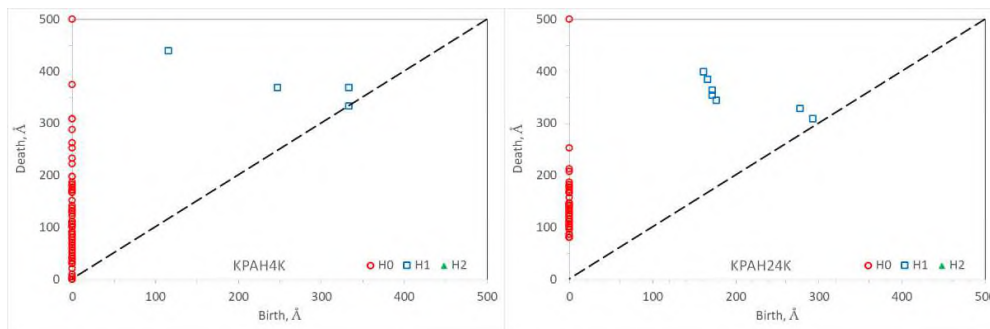


Figure 282 Persistence diagrams: Death vs. Birth of control (K): 4h (left) and 24h (right)

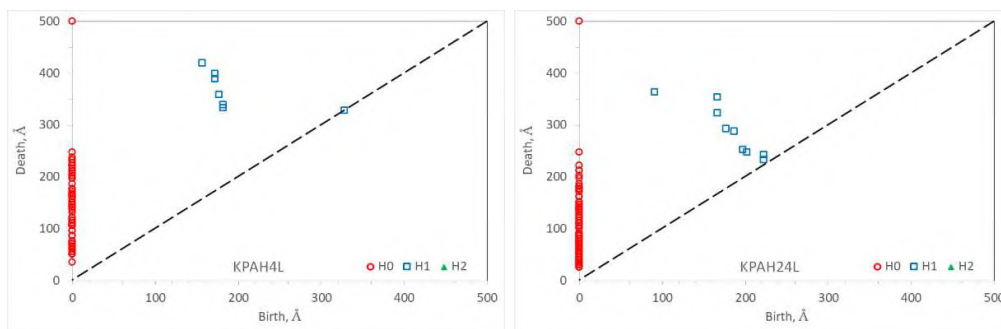


Figure 283 Persistence diagrams: Death vs. Birth of low exposure (L): 4h (left) and 24h (right)

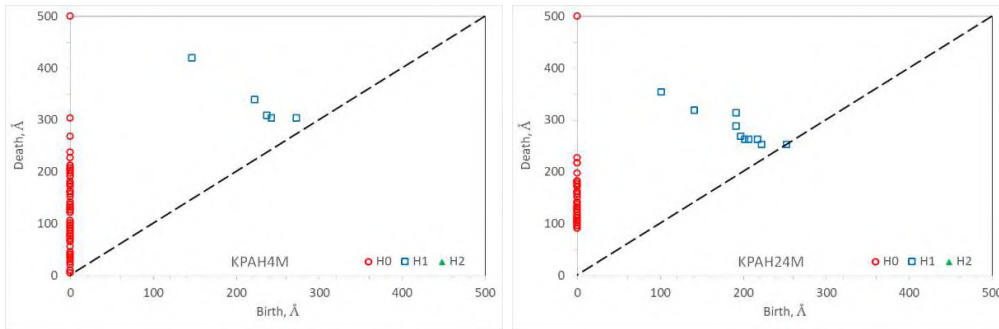


Figure 284 Persistence diagrams: Death vs. Birth of middle exposure (M): 4h (left) and 24h (right)

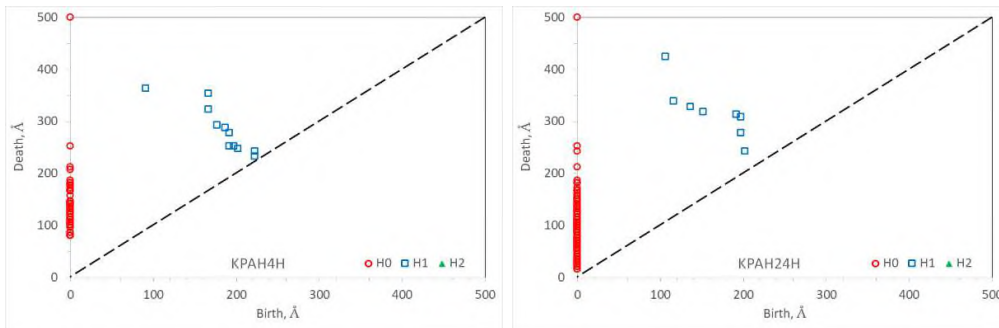


Figure 285 Persistence diagrams: Death vs. Birth of high exposure (H): 4h (left) and 24h (right)

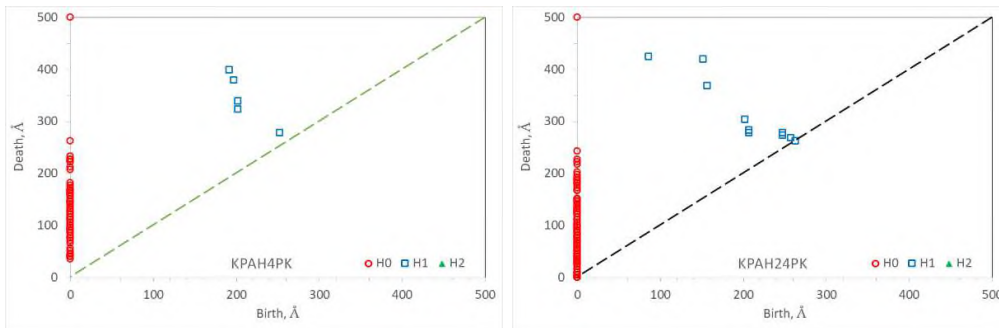


Figure 286 Persistence diagrams: Death vs. Birth of positive control (PK): 4h (left) and 24h (right)

Persistence diagrams: Life vs. Birth

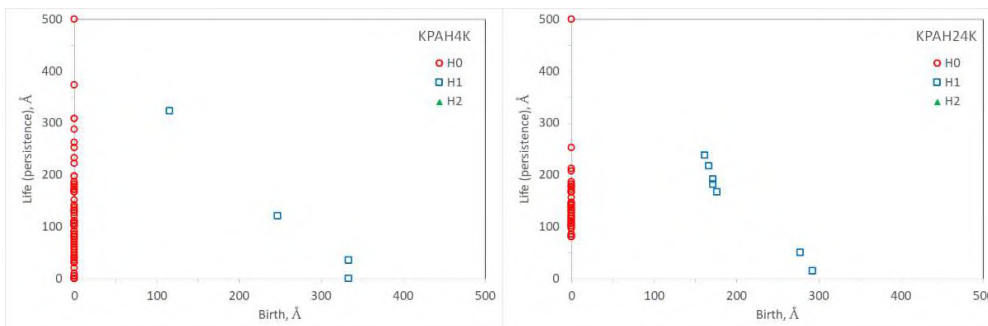


Figure 287 Persistence diagrams: Life vs. Birth of control (K): 4h (left) and 24h (right)

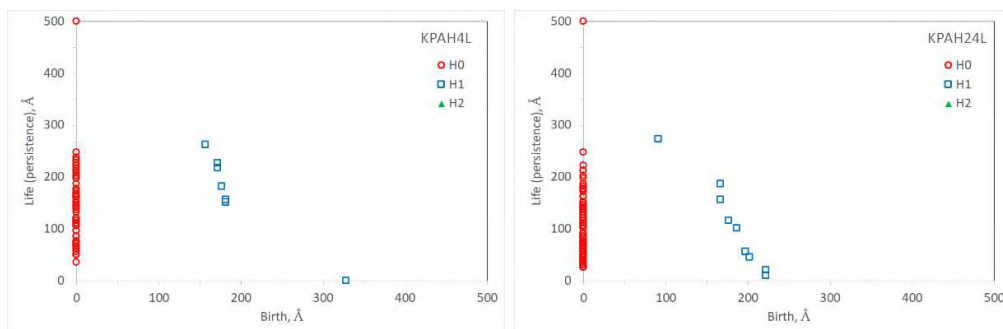


Figure 288 Persistence diagrams: Life vs. Birth of low exposure (L): 4h (left) and 24h (right)

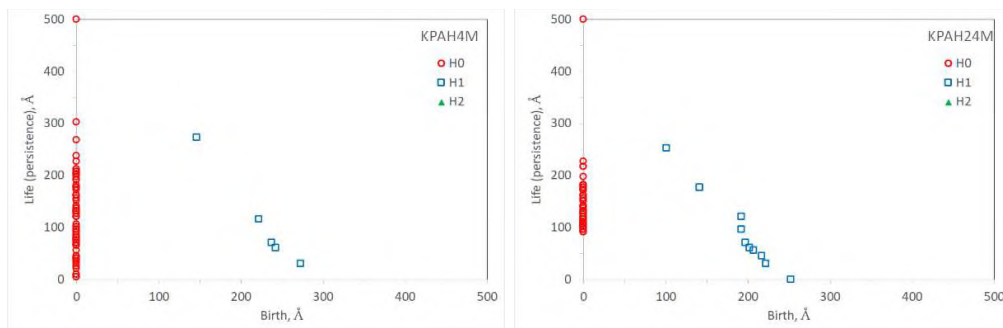


Figure 289 Persistence diagrams: Life vs. Birth of middle exposure (M): 4h (left) and 24h (right)

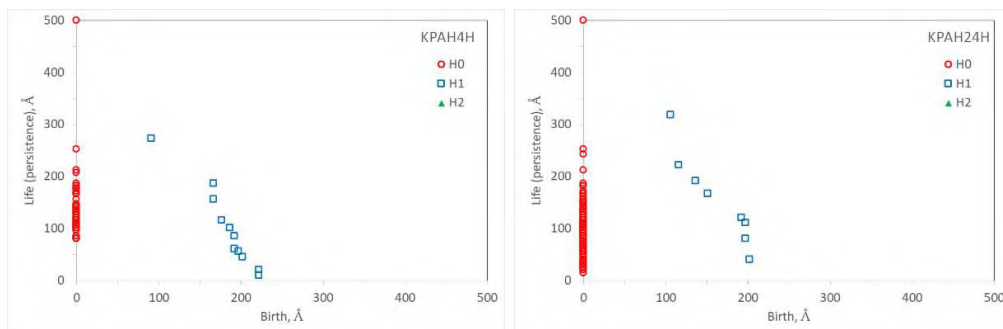


Figure 290 Persistence diagrams: Life vs. Birth of high exposure (H): 4h (left) and 24h (right)

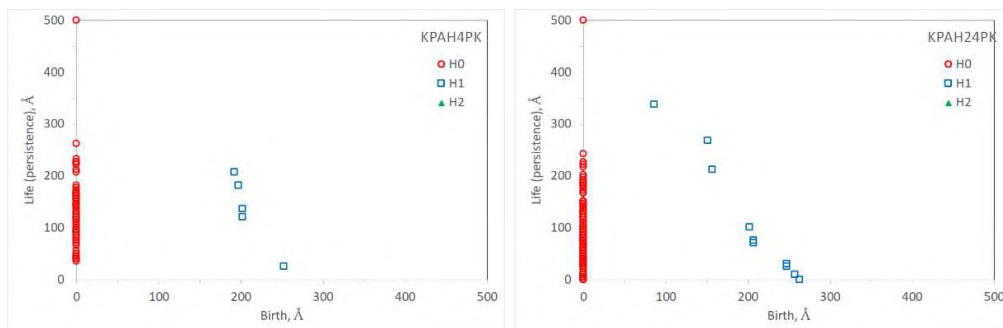


Figure 291 Persistence diagrams: Life vs. Birth of positive control (PK): 4h (left) and 24h (right)

Persistence diagrams

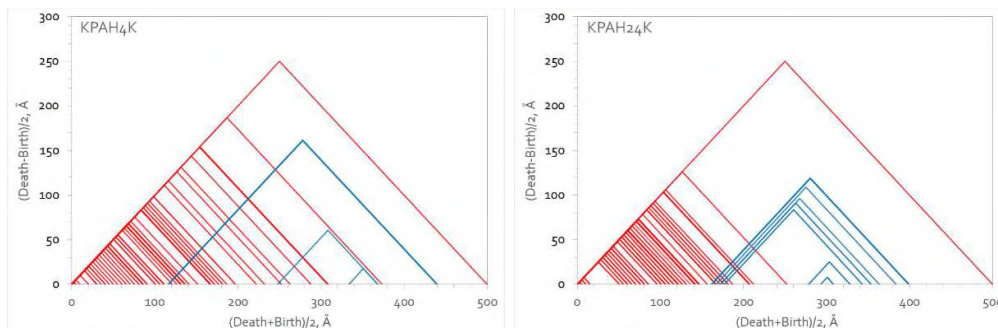


Figure 292 Persistence diagrams of control (K): 4h (left) and 24h (right)

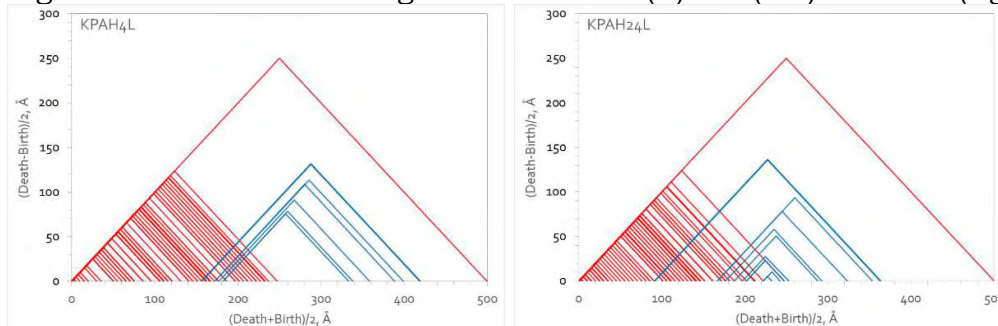


Figure 293 Persistence diagrams of low exposure (L): 4h (left) and 24h (right)

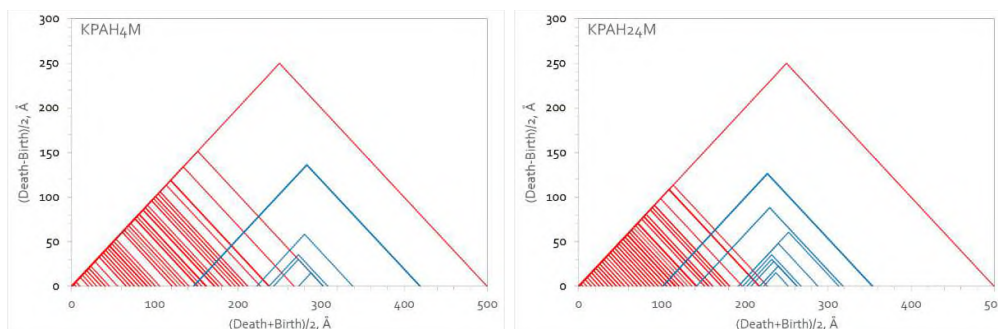


Figure 294 Persistence diagrams of middle exposure (M): 4h (left) and 24h (right)

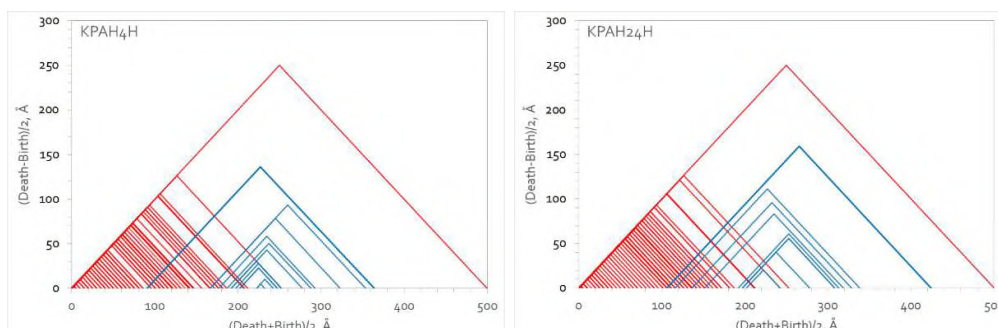


Figure 295 Persistence diagrams of high exposure (H): 4h (left) and 24h (right)

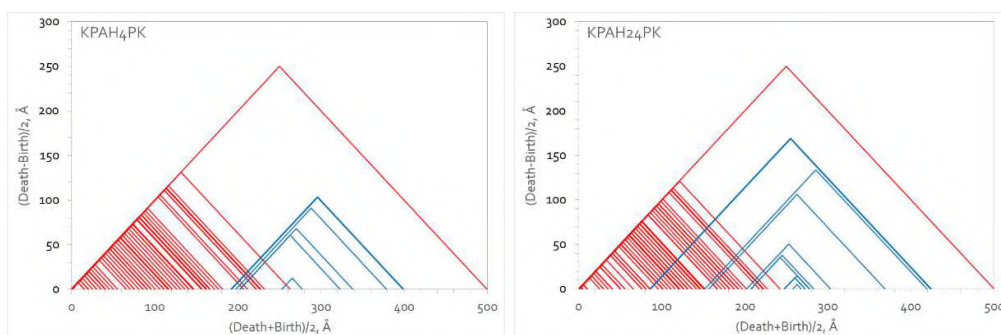


Figure 296 Persistence diagrams of positive control (PK): 4h (left) and 24h (right)

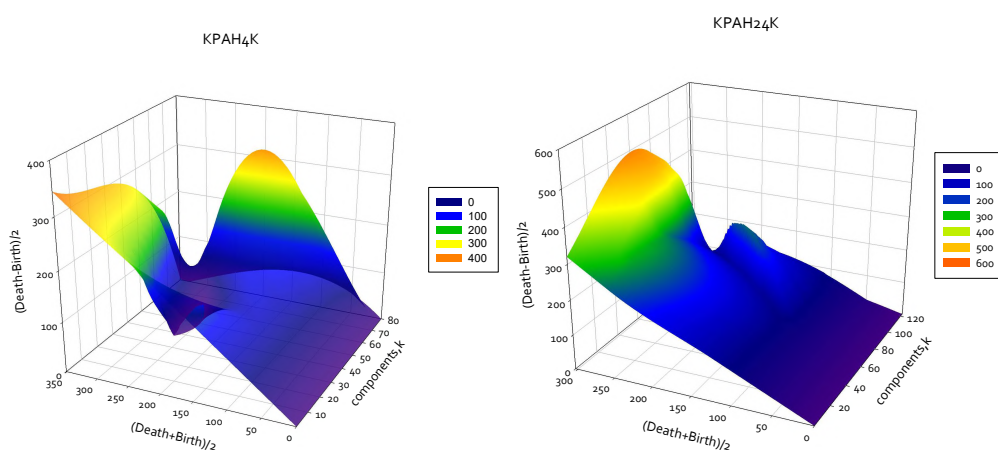


Figure 297 3D-Persistence diagrams of control (K): 4h (left) and 24h (right)

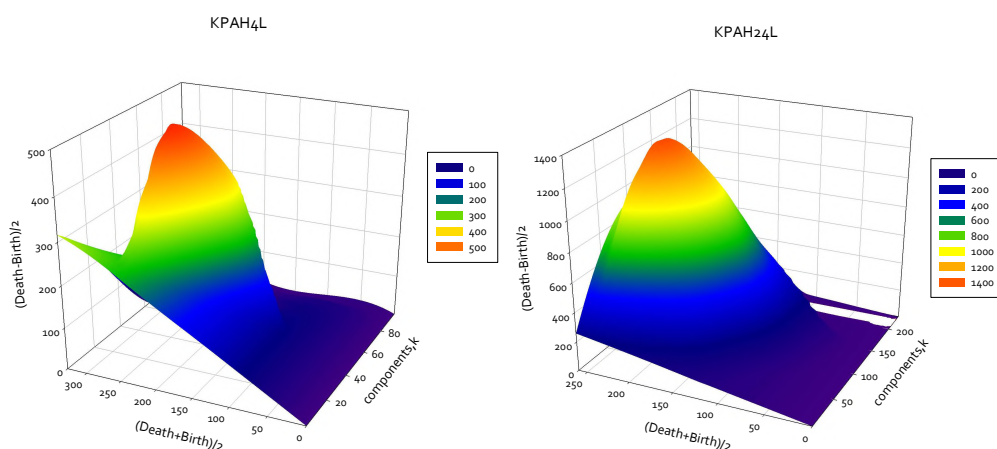


Figure 298 3D-Persistence diagrams of low exposure (L): 4h (left) and 24h (right)

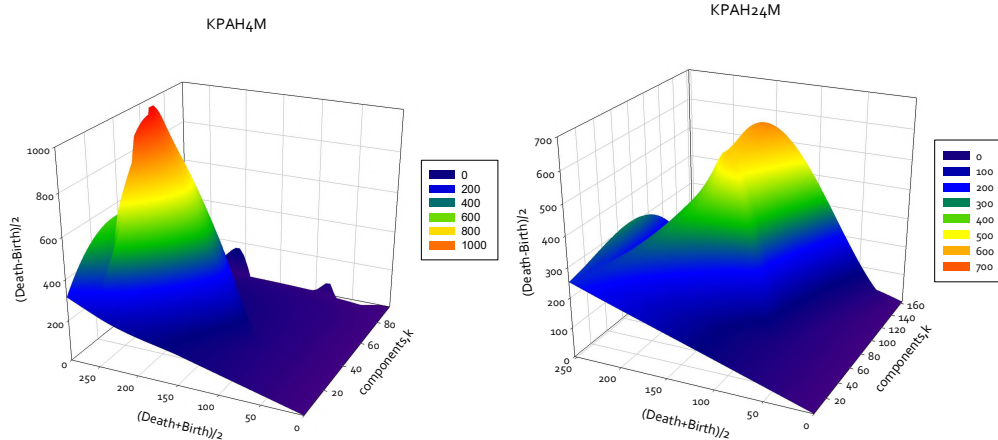


Figure 2993D- Persistence diagrams of middle exposure (M): 4h (left) and 24h (right)

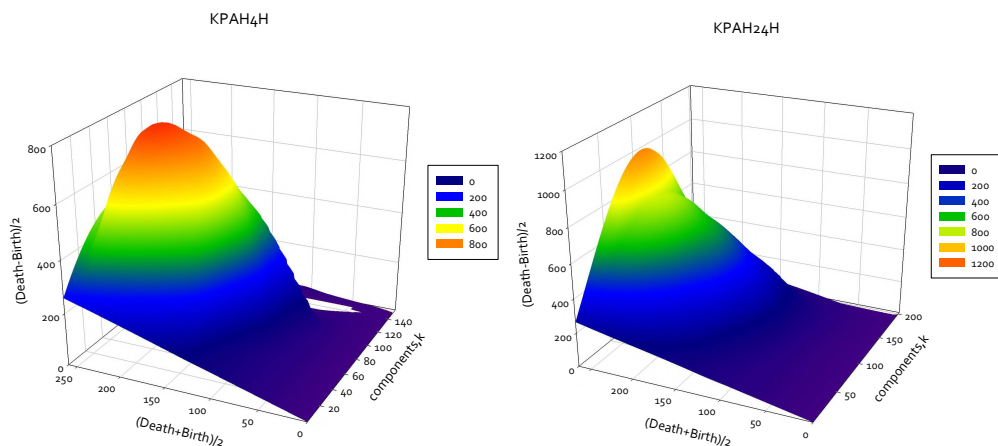


Figure 300 3D-Persistence diagrams of high exposure (H): 4h (left) and 24h (right)

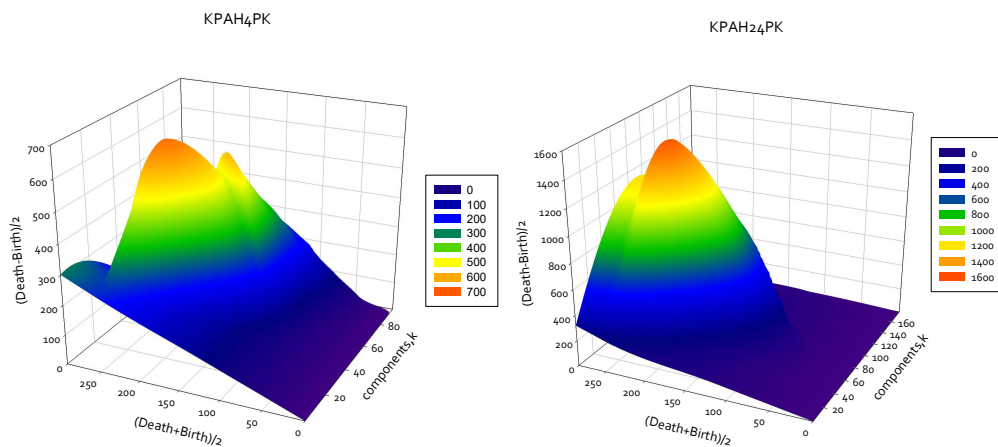


Figure 301 3D-Persistence diagrams of positive control (PK): 4h (left) and 24h (right)

Barcodes

Betti intervals or barcodes¹⁶⁴ help describe how the homology of $X(t)$ changes with t . A k -dimensional Betti interval, with endpoints $[t_{\text{start}}, t_{\text{end}})$, corresponds roughly to a k -dimensional hole that appears at filtration time t_{start} , remains open for $t_{\text{start}} \leq t < t_{\text{end}}$, and closes at time t_{end} . We are often interested in Betti intervals that persist for a long filtration range.

The most important features that Persistent Homology returns from the object is its Barcode. The axis indicates the radii of the balls grown around our data points and the bars themselves show when a generator is formed from this process (its birth) and when it gets filled in and vanishes (its death). Naturally, the red H_0 generators are all formed simultaneously and individually vanish as points get connected in the filtration.

Note: The length of these bars describes the "persistence" of our generators and the longer the bar, the more significant the corresponding generator.

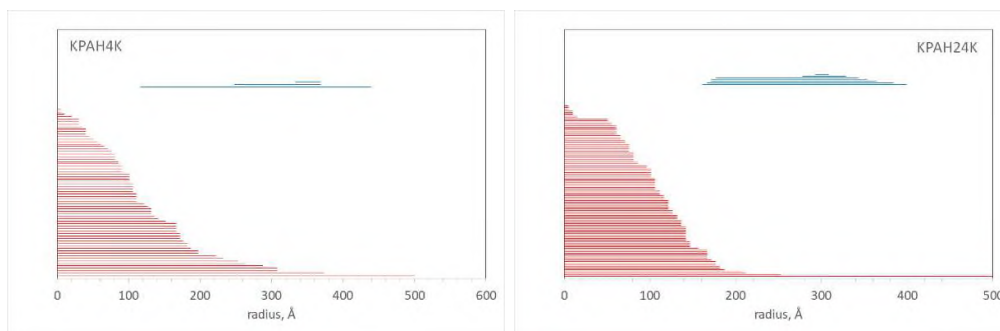


Figure 302 Barcodes of control (K): 4h (left) and 24h (right)

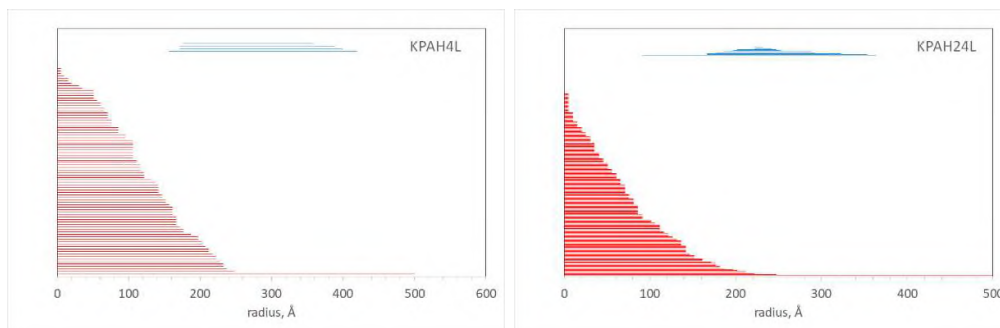


Figure 303 Barcodes of low exposure (L): 4h (left) and 24h (right)

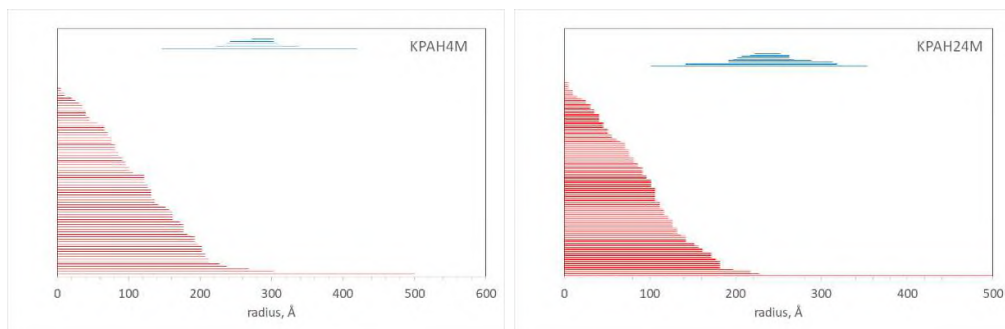


Figure 304 Barcodes of middle exposure (M): 4h (left) and 24h (right)

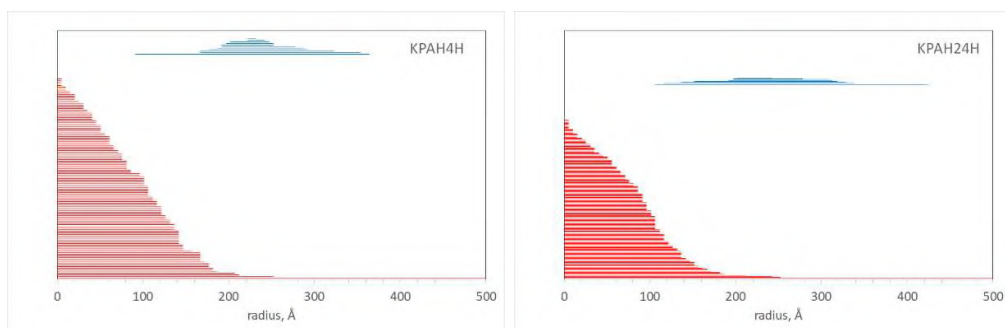


Figure 305 Barcodes of high exposure (H): 4h (left) and 24h (right)

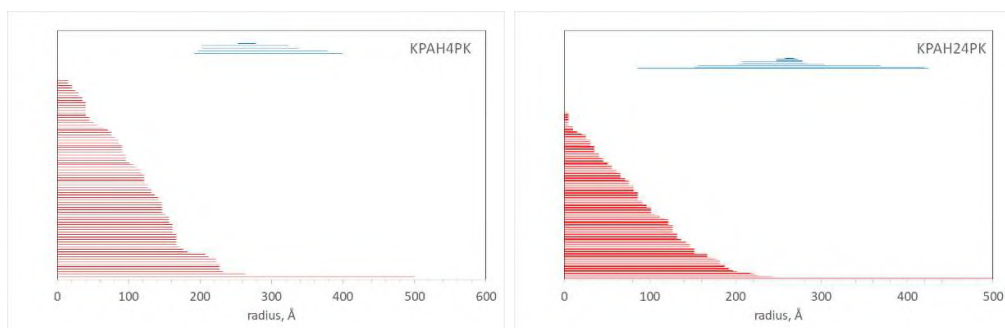


Figure 306 Barcodes of positive control (PK): 4h (left) and 24h (right)

Shannon persistent entropy

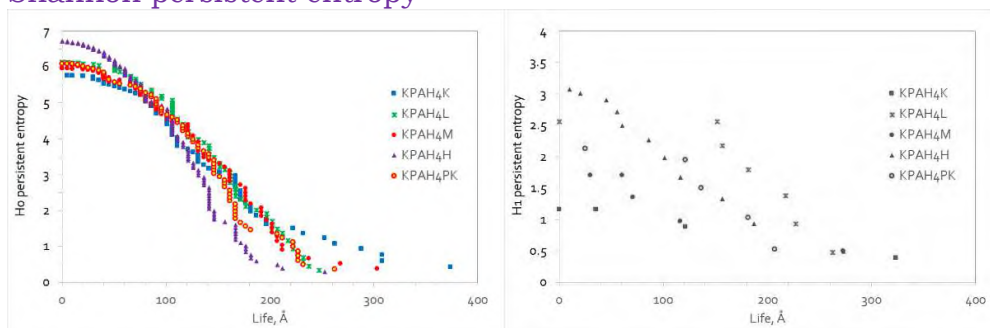


Figure 307 Shannon H_0 persistent entropy (left) and Shannon H_1 persistent entropy **versus** Life (=Death-Birth) (Å) at 4h

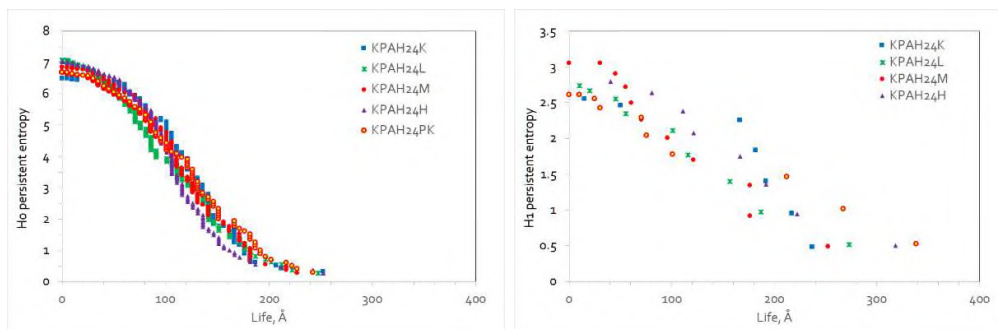


Figure 308 Shannon H_0 persistent entropy (left) and Shannon H_1 persistent entropy **versus** Life (=Death-Birth) (Å) at 24h

Intuitively, entropy measure different barcodes length. A barcode with uniform lengths has small entropy.

Table 27 Shannon H_0 persistent Entropy. H_1 persistent Entropy and Total persistent Entropy

	exposure	H_0 persistent entropy	H_1 persistent entropy	Total persistent entropy
KPAH4K	0	5.7665	1.1626	5.8221
KPAH4L	1	6.1358	2.5563	6.2466
KPAH4M	8	5.9625	1.7064	6.0232
KPAH4H	16	6.7123	3.0693	6.8229
KPAH4PK		6.0912	2.1244	6.1803
KPAH24K	0	6.4928	2.5506	6.5783
KPAH24L	1	7.0682	2.7389	7.1344
KPAH24M	8	6.8503	3.0486	6.9497
KPAH24H	16	7.0203	2.7977	7.0846
KPAH24PK		6.6844	2.6112	6.7591

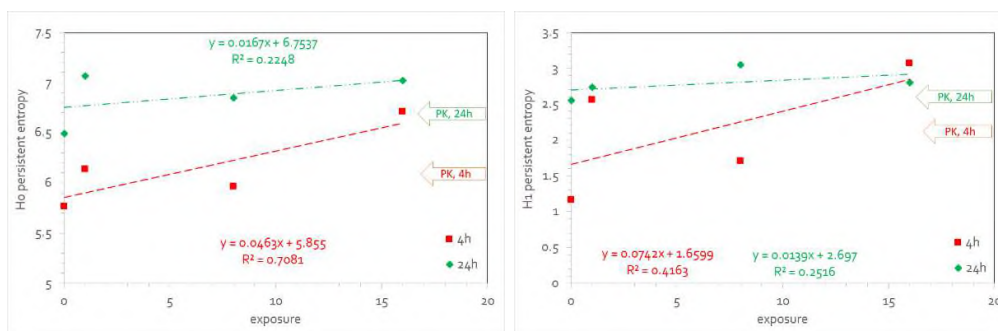


Figure 309 Shannon H_0 persistent Entropy (left) and H_1 persistent Entropy (right) **versus** exposure (h)

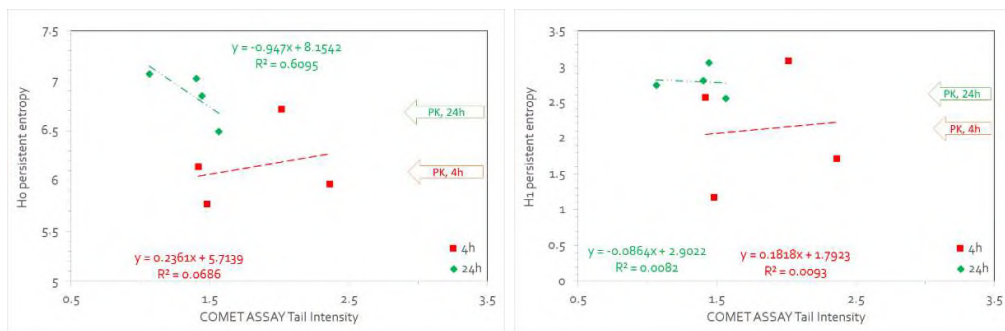


Figure 310 Shannon H0 persistent Entropy (left) and H1 persistent Entropy (right) versus COMET ASSAY Tail Intensity

Discussion

The initial question posed in designing this study was whether treatment is related to the morphology in comet assay images and their ability to fill the surrounding space. To investigate this, we checked for correlations between all parameters and treatment, and divided the sample into three treatment groups (L, M and H). Few variables showed any significant correlation with treatment, some *not even by chance* (?). This finding was further corroborated by the treatment group separation, which also yielded no significant differences.

This result leads to several possible conclusions.

First, the chosen morphology of comet assay images might be too morphologically uniform to detect significant differences. This contrasts somewhat with the second part of the study, where clear morphological differences based on the spatial origin of treatments were observed. However, this does not negate the possibility that morphology of comet assay images is uniform within individual spatial groups.

Another possible explanation is that multifractal spectra are not suitable for this analysis. Studies using monofractal analysis have shown significant differences in the box-count dimension. The central point of the generalized dimensions spectrum $DQ(Q)$ at $Q = 0$ represents the Minkowski–Bouligand (*i.e.*, box-count) dimension, so we could have expected statistically significant differences at least in this part of the spectrum, but they were absent.

The third conclusion is that the morphology of individual morphology of in comet assay images in this structure does not change significantly with treatments. To confirm this conclusion, further studies on larger samples are necessary. The differences could also suggest a reduction in the space-filling complexity and an increase in local heterogeneity of finer structures with treatments. However, the distributions of the variables were not statistically significantly different among treatments, which does not support these notions. A larger sample could reveal more noticeable differences or confirm the findings of this study, showing no significant morphological differences with treatments.

Conclusions

Image analysis for the exposures (L, M, H) and 4h time did not produce significant variations in the parameters.

Image analysis for the exposures (L, M, H) and 24h time did not produce significant variations in the parameters, except for H exposure and 24h time: Area, Perimeter Length, Length, Width, Convex Area, Convex Perimeter Length, Caliper Length, Caliper Width!

Multi-scale focal fractal dimension may only reveal effect of time (4h, 24h) in separate H exposure level and control.

References

- ¹ Collins, A. R., Azqueta Oscoz, A, Brunborg, G., Gaivão, I., Giovannelli, L., Kruszewski, M., Smith, C.C., Štětina, R. The comet assay: topical issues, *Mutagenesis* 23(3) (2008) 143-151, <https://doi.org/10.1093/mutage/gem051>
- ² Gagné, F. Chapter 10 - Genotoxicity. *Biochemical Ecotoxicology Principles and Methods* (2014) 171-196. <https://doi.org/10.1016/B978-0-12-411604-7.00010-6>
- ³ Afiahayati, Anarossi, E., Yanuarieska, R. D., Nuha, F. U., Mulyana, S. (2020). Comet Assay Classification for Buccal Mucosa's DNA Damage Measurement with Super Tiny Dataset Using Transfer Learning. *In: Huk, M., Maleszka, M., Szczerbicki, E. (eds) Intelligent Information and Database Systems: Recent Developments. ACIIDS 2019. Studies in Computational Intelligence*, vol 830. Springer, Cham. https://doi.org/10.1007/978-3-030-14132-5_22
- ⁴ Jia, X., Garboczi, E. Advances in shape measurement in the digital world. *Particuology* 26 (2016) 19-31. <https://doi.org/10.1016/j.partic/2015.12.005>
- ⁵ Hoeijmakers, J.H. DNA damage, aging, and cancer. *N. Engl. J. Med.* 2009;361(15):1475-1485. <https://doi.org/10.1056/NEJMra0804615>
- ⁶ Ostling, O., Johanson, K. J. Microelectrophoretic study of radiation-induced DNA damages in individual mammalian cells. *Biochem. Biophys. Res Commun.* 123 (1984) 291-298. [https://doi.org/10.1016/0006-291X\(84\)90411-X](https://doi.org/10.1016/0006-291X(84)90411-X)
- ⁷ Fairbairn, D. W., Olive, P. L., O'Neill, K. L. The comet assay: a comprehensive review. *Mutat. Res.* 1995;339(1):37-59. [https://doi.org/10.1016/0165-1110\(94\)00013-3](https://doi.org/10.1016/0165-1110(94)00013-3)
- ⁸ Clementi, E., Garajova, Z., Markkanen, E. Measuring DNA Damage Using the Alkaline Comet Assay in Cultured Cells. *Bio-protocol* 11(16) (2021) e4119. <https://doi.org/10.21769/BioProtoc.4119>
- ⁹ Seal, S., Trapotsi, M. A., Spjuth, O., Singh, S., Carreras-Puigvert, J., Greene, N., Bender, A., Carpenter, A. E. Cell Painting: a decade of discovery and innovation in cellular imaging. *Nat Methods.* 22(2) (2025) 254-268. <https://doi.org/10.1038/s41592-024-02528-8>. Erratum in: *Nat Methods.* 22(2) (2025) 447, <https://doi.org/10.1038/s41592-024-02578-y>
- ¹⁰ F. Pagliari, M.-F. Spadea, P. Montay-Gruel, A. Puspitasari-Kokko, J. Seco, L. Tirinato, A. Accardo, F. De Angelis, F. Gentile. Nano-Topography Enhanced Topological-Cell-Analysis in Radiation-Therapy. *Adv. Healthcare Mater.* 14 (2025) 2405187. <https://doi.org/10.1002/adhm.202405187>
- ¹¹ R. M. Haralick, K. Shanmugam and I. Dinstein. Textural Features for Image Classification. *IEEE Trans. Syst. Man Cyber.* SMC-3(6) (1973) 610-621. Nov., <https://doi.org/10.1109/TSMC.1973.4309314>
- ¹² Walker, R. F., Jackway, P., Longstaff, I. D. Improving Co-occurrence Matrix Feature Discrimination. *DICTA '95, 3rd Conference on Digital Image Computing: Techniques and Application*, 6 - 8 December, 1995.

https://espace.library.uq.edu.au/view/UQ:11262/walker95improvin.pdf?dsi_version=5836535c8f4487e98e57afeda7ca3197

- ¹³ Haralick, R. M. Statistical and structural approaches to texture. *Proc. IEEE* 67(5) (1979) 786-804. <https://doi.org/10.1109/PROC.1979.11328>
- ¹⁴ Hoffman, J., Zheng, S., Zhang, H., Murphy, R.F., Dahl, K. N. Image-based discrimination of the early stages of mesenchymal stem cell differentiation. *Mol. Biol. Cell* 35(8) (2024) 103. <https://doi.org/10.1091/mbc.E24-02-0095>
- ¹⁵ Nanni, L., Brahnam, S., Ghidoni, S., Menegatti, E., Barrier, T. Different approaches for extracting information from the co-occurrence matrix. *PLOS ONE* 8(12) (2013) e83554. <https://doi.org/10.1371/journal.pone.0083554>
- ¹⁶ Murata, S., Herman, P., Lakowicz, J. R. Texture analysis of fluorescence lifetime images of AT- and GC-rich regions in nuclei. *J. Histochem. Cytochem.* 49(11) (2001) 1443-1451. <https://doi.org/10.1177/002215540104901112>
- ¹⁷ Gipp, M., Marcus, G., Harder, N., Suratanee, A., Rohr, K., König, R., Männer, R.. Haralick's Texture Features Computation Accelerated by GPUs for Biological Applications. *In: Bock, H., Hoang, X., Rannacher, R., Schlöder, J. (eds) Modeling, Simulation and Optimization of Complex Processes.* Springer, Berlin, Heidelberg (2012). https://doi.org/10.1007/978-3-642-25707-0_11
- ¹⁸ Sutcliffe, M. D., Tan, P. M., Fernandez-Perez, A., Nam, Y.-J., Munshi, N. V., Saucerman, J. J. High content analysis identifies unique morphological features of reprogrammed cardiomyocytes. *Sci. Rep.* 8 (2018) 1258. <https://doi.org/10.1038/s41598-018-19539-z>
- ¹⁹ Paduano, V., Tagliaferri, D., Falco, G., Ceccarelli, M. Automated Identification and Location Analysis of Marked Stem Cells Colonies in Optical Microscopy Images. *PLOS ONE* 8(12) (2013) e80776. <https://doi.org/10.1371/journal.pone.0080776>
- ²⁰ King, R. D., Brown, B., Hwang, M., Jeon, T., George, A. T. Fractal dimension analysis of the cortical ribbon in mild Alzheimer's disease. *Neuroimage* 53 (2010) 471-479. <https://doi.org/10.1016/j.neuroimage.2010.06.050>
- ²¹ Dioguardi, N., Grizzi, F., Franceschini, B., Bossi, P., Russo, C. Liver fibrosis and tissue architectural change measurement using fractal-rectified metrics and Hurst's exponent. *World J. Gastroenterol.* 12 (2006) 2187-2194. <https://doi.org/10.3748/wjg.v12.i14.2187>
- ²² Canals, M., Olivares, R., Labra, F., Novoa, F.F. Ontogenetic changes in the fractal geometry of the bronchial tree in *Rattus norvegicus*. *Biol. Res.*;33(1) (2000) 31-35. <https://doi.org/10.4067/s0716-97602000000100010>
- ²³ Castelli, C., Losa, G.A. Ultrastructural complexity of nuclear components during early apoptotic phases in breast cancer cells. *Anal. Cell. Pathol.* 23 (2001) 1-9. <https://doi.org/10.1155/2001/828309>
- ²⁴ Lopes, R., Betrouni, N. Fractal and Multifractal Analysis: A Review. *Med. Image Anal.* 13(4) (2009) 634-649. <https://doi.org/10.1016/j.media.2009.05.003>

-
- ²⁵ Gentile, F., Tirinato, L., Battista, E., Causa, F., Liberale, C., di Fabrizio, E. M., Decuzzi, P. *Biomaterials* 31 (28) (2010) 7205-7212. <https://doi.org/10.1016/j.biomaterials.2010.06.016>
- ²⁶ Datseris, G., Kottlarz, I., Braun, A. P., Parlitz, U. Estimating fractal dimensions: A comparative review and open source implementations. *Chaos* 33 (10) (2023) 102101. <https://doi.org/10.1063/5.0160394>
- ²⁷ Mark, D. M., Aronson, P. B. Scale-Dependent fractal dimensions of topographic surfaces: An empirical investigation with applications in geomorphology and computer mapping. *Mathematical Geology*, 16 (7) (1984) 671-683. <https://doi.org/10.1007/BF01033029>
- ²⁸ Kimler, V. A., Tracy-Bee, M., Ollie, C. D., Langer, R. M., Montante, J. M., Marks C. R. C., Freeman, D. C. R., Hough, R. A., Taylor, J. D. Characterization of melanophore morphology by fractal dimension analysis. *Pigment Cell Res.* 17(2) (2004) 165-172. <https://doi.org/10.1046/j.1600-0749.2003.00125.x>
- ²⁹ Gilmore, S., Hofmann-Wellenhof, R., Muir, J., Soyer, H.P., Lacunarity analysis: a promising method for the automated assessment of melanocytic naevi and melanoma. *PLoS One* 4(10) (2009) e7449. <https://doi.org/10.1371/journal.pone.0007449>
- ³⁰ Gould, D. J., Vadakkan, T. J., Poche, R. A., Dickinson, M. E., Multifractal and lacunarity analysis of microvascular morphology and remodeling. *Microcirculation* 18 (2011) 136-151. <https://doi.org/10.1111/j.1549-8719.2010.00075.x>
- ³¹ Tambasco, M., Costello, B. M., Kouznetsov, A., Yau, A., Magliocco, A. M. Quantifying the architectural complexity of microscopic images of histology specimens. *Micron* 40(4) (2009) 486-494. <https://doi.org/10.1016/j.micron.2008.12.004>
- ³² Smith, T. G., Marks, W. B., Lange, G. D., Sheriff Jr., W.H., Neale, E.A. A fractal analysis of cell images. *J. Neurosci. Methods* 27(2) (1989) 173-180. [https://doi.org/10.1016/0165-0270\(89\)90100-3](https://doi.org/10.1016/0165-0270(89)90100-3)
- ³³ Smith, T. G., Lange, G. D., Marks, W. B. Fractal Methods and Results in Cellular Morphology-Dimensions, Lacunarity and Multifractals. *J. Neurosci. Methods* 69(2) (1996) 123-136. [https://doi.org/10.1016/S0165-0270\(96\)00080-5](https://doi.org/10.1016/S0165-0270(96)00080-5)
- ³⁴ Rosenberg, E. Generalized Dimensions and Multifractals. In: *Fractal Dimensions of Networks*. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-43169-3_16
- ³⁵ Salat, H., Murcio, R., Arcaute, E. Multifractal Methodology. *Phys. A Stat. Mech. Its Appl.* 473 (2017) 467-487. <https://doi.org/10.1016/j.physa.2017.01.041>
- ³⁶ Telesca, L., Lapenna, V., Macchiato, M. Mono- and multifractal investigation of scaling properties in temporal patterns of seismic sequences. *Chaos, Solitons Fract.* 19(1) (2004) 1-15. [https://doi.org/10.1016/S0960-0779\(03\)00188-7](https://doi.org/10.1016/S0960-0779(03)00188-7)
- ³⁷ Feder, J. *Fractals*, 1st ed.; Springer: New York, NY, USA, 1998.
- ³⁸ Ignatova, A.M., Zemlyanova, M.A., Naimark, O.B., Zaitseva, N.V. Practical Aspects of Using Multifractal Formalism to Assess the Morphology of Biological Tissues. *Sovrem. Tekhnologii Med.* 15(3) (2023) 35-40. <https://doi.org/10.17691/stm2023.15.3.04>
- ³⁹ Karperien, A. *FracLac* for ImageJ; JavaDoc, source code, and jar, (Version 2.5) (2001). https://imagej.net/ij/plugins/fraclac/Frac_Lac.jar

-
- ⁴⁰ Karperien, A. FracLac User's Guide (2001).
<http://rsbweb.nih.gov/ij/plugins/fractal/FLHelp/Introduction.htm>
- ⁴¹ Karperien, A., Ahammer, H., Jelinek, H. F. Quantitating the subtleties of microglial morphology with fractal analysis. *Front Cell Neurosci.* Jan 30 (2013) 7:3. doi:
<https://doi.org/10.3389/fncel.2013.00003>
- ⁴² Karperien, A., Jelinek, H. Fractal, multifractal, and lacunarity analysis of microglia in tissue engineering. *Front. Bioeng. Biotechnol., Sec. Biomaterials*, 3 (2015), Article 51, 1-4
<https://doi.org/10.3389/fbioe.2015.00051>
- ⁴³ Yao, B., He, H., Kang, S., Chao, Y., He, L. A Review for the Euler Number Computing Problem. *Electronics*, 12(21) (2023) 4406. <https://doi.org/10.3390/electronics12214406>
- ⁴⁴ Yang, H., Sengupta, S. Intelligent shape recognition for complex industrial tasks. *IEEE Control Syst. Mag.* 8 (1988) 23-30. <https://doi.org/10.1109/37.473>
- ⁴⁵ Nayar, S., Bolle, R. Reflectance-based object recognition. *Int. J. Comput. Vis.* 17 (1996) 219-240. <https://doi.org/10.1007/BF00128232>
- ⁴⁶ Hashizume, A., Suzuki, R., Yokouchi, H., Horiuchi, H., Yamamoto, S. An algorithm of automated RBC classification and its evaluation. *Bio. Med. Eng.* 28(1) (1990) 25-32.
<https://doi.org/10.11239/jsmbe1963.28.25>
- ⁴⁷ Snidaro, L., Foresti, G. L. Real-time thresholding with Euler numbers. *Pattern Recognition Lett.* 24 (9-10) 1533 - 1544. [https://doi.org/10.1016/S0167-8655\(02\)00392-6](https://doi.org/10.1016/S0167-8655(02)00392-6)
- ⁴⁸ Ernesti, F., Schneider, M., Winter, S., Hug, D., Last, G., Böhlke, T. Characterizing digital microstructures by the Minkowski-based quadratic normal tensor. *Math. Meth. Appl. Sci.* 46(1) (2023) 961-985. <https://doi.org/10.1002/mma.8560>,
<https://doi.org/10.48550/arXiv.2007.15490>
- ⁴⁹ Chiu, S. N., Stoyan, D., Kendall, W. S., Mecke, J. *Stochastic geometry and its applications* (John Wiley & Sons, 2013).
- ⁵⁰ Schneider, R., Weil, W. *Stochastic and Integral Geometry*, Springer (2008).
- ⁵¹ Huber, M. B., Nagarajan, M., Leinsinger, G., Lawrence, A. Ray, L. A., Wismüller, A. Classification of interstitial lung disease patterns with topological texture features. *Proc. SPIE* 7624, *Medical Imaging 2010: Computer-Aided Diagnosis*, 762410 (9 March 2010);
<https://doi.org/10.1117/12.844318>
- ⁵² Zhao, F., Mendonça, P., Kaucic, R. Image-based automated defect recognition via statistical learning of Minkowski functionals. *Eur. Conf. NonDestr. Test.* (2010) 1-10.
- ⁵³ Minkowski, H. Allgemeine Lehrsätze über die konvexen Polyeder. In: *Ausgewählte Arbeiten zur Zahlentheorie und zur Geometrie*. Teubner-Archiv zur Mathematik 12 (1897) 198-220.
https://doi.org/10.1007/978-3-7091-9536-9_5
- ⁵⁴ R  th, C., Monetti, R., Bauer, J., Sidorenko, I., M  ller, D., Matsuura, M., Lochm  ller, E-M., Zysset, P., Eckstein, F. Strength through structure: visualization and local assessment of the trabecular bone structure. *New J. Phys.* 10 (2008) 125010. <https://doi.org/10.1088/1367-2630/10/12/125010>

-
- ⁵⁵ Attneave, F., Arnoult, M. D. The quantitative study of shape and pattern perception. Psychol. Bull. 53 (1956) 452-471. <https://psycnet.apa.org/doi/10.1037/h0044049>
- ⁵⁶ Silva, L. E. V., Senra Filho, A. C. S., Fazan, V. P. S., Felipe, J. C., Murta, L. O. Junior. Two-dimensional sample entropy: assessing image texture through irregularity. Biomed. Phys. Eng. Express 2(4) (2016) 045002
<https://doi.org/10.1088/2057-1976/2/4/045002>
- ⁵⁷ Pal, N. R., Pal, S. K. Entropy: a new definition and its applications. IEEE Trans. Syst. Man Cyber. 21(5) (1991) 1260-1270. <https://doi.org/10.1109/21.120079>
- ⁵⁸ Thum, Ch. Measurement of the Entropy of an Image with Application to Image Focusing. Optica Acta 31(2) (1984) 203-211. <https://doi.org/10.1080/713821475>
- ⁵⁹ Arbeláez, P., Cohen, L. Energy partitions and image segmentation. J. Math. Imag. Vision 20(1-2) (2004) 43-57. <https://doi.org/10.1023/B:JMIV.0000011318.77653.44>
- ⁶⁰ Harshita, Sh., Zerbe N., Lohmann, S., Kayser, K., Hellwich, O., Hufnagl, P. A review of graph-based methods for image analysis in digital histopathology. Diagn. Pathol. (2015) 1-61. <https://doi.org/10.17629/www.diagnosticpathology.eu-2015-1:61>
- ⁶¹ Godrèche, C., Kostov, I. I., Yekutieli, I. I. Topological correlations in cellular structures and planar graph theory. Phys. Rev. Lett. 69(18) (1992) 2674-2677. <https://doi.org/10.1103/PhysRevLett.69.2674>
- ⁶² Vincent, L., Heijmans, H. Graph Morphology in Image Analysis. *In: Mathematical Morphology in Image Processing*, CRC Press, 1992
- ⁶³ Zampirolli, F. dA. Neighborhood graphs built with morphological operators. Revista de Informatica Aplicada/J. Appl. Comput. 4(2) (2010) 5-12. https://seer.uscs.edu.br/index.php/revista_informatica_aplicada/article/view/309/287
- ⁶⁴ de Assis Zampirolli, F., Stransky, B., Lorena, A.C., de Melo Paulon, F.L. Segmentation and classification of histological images-application of graph analysis and machine learning methods. *In Graphics, Patterns and Images (SIBGRAPI)*, 2010 23rd SIBGRAPI Conference on, pages 331-338. IEEE, 2010. <https://doi.org/10.1109/SIBGRAPI.2010.51>
- ⁶⁵ Lund, A.W., Bilgin, C.C., Hasan, M.A., McKeen, L.M., Stegemann, J.P., Yener, B., Zaki, M.J., Plopper, G.E. Quantification of spatial parameters in 3D cellular constructs using graph theory. J. Biomed. Biotechnol. 2009 (2009) 928286. <https://doi.org/10.1155/2009/928286>
- ⁶⁶ Serra J., Image analysis and mathematical morphology, Academic Press (1982).
- ⁶⁷ Haas, A., Serra, J., Matheron, G., Morphologie mathématique et granulométries en place. Ann. Mines 11 (1967) 736-753, 12 (1967) 767-782.
- ⁶⁸ Raymond, E., Raphael, M., Grimaud, M., Vincent, L., Binet, J.L., Meyer, F. Germinal center analysis with the tools of mathematical morphology on graphs. Cytometry 14(8) (1993) 848-861. <https://doi.org/10.1002/cyto.990140803>
- ⁶⁹ Ali, S., Veltri, R., Epstein, J.A., Christudass, C., Madabhushi, A. Cell cluster graph for prediction of biochemical recurrence in prostate cancer patients from tissue microarrays. *In:*

SPIE Medical Imaging, pages 86760H-86760H. International Society for Optics and Photonics, 2013. <https://doi.org/10.1117/12.2008695>

⁷⁰ Demir, C., Gultekin, S.H., Yener, B. Augmented cell-graphs for automated cancer diagnosis. *Bioinformatics* 21 (2) (2005) 7-12. <https://doi.org/10.1093/bioinformatics/bti1100>

⁷¹ Bilgin, C.C., Bullough, P., Plopper, G.E., Yener, B. ECM-Aware Cell-Graph Mining for Bone Tissue Modeling and Classification. *Data Min. Knowl. Discov.* 20(3) (2009) 416-438. <https://doi.org/10.1007/s10618-009-0153-2>

⁷² Sudbø, J., Marcelpoil, R., Reith, A. New algorithms based on the Voronoi Diagram applied in a pilot study on normal mucosa and carcinomas. *Anal. Cell Pathol.* 21(2) (2000) 71-86. <https://doi.org/10.1155/2000/389361>

⁷³ Altunbay, D., Cigir, C., Sokmensuer, C., Gunduz-Demir, C. Color graphs for automated cancer diagnosis and grading. *IEEE Trans. Biomed. Eng.* 57(3) (2010) 665-674. <https://doi.org/10.1109/TBME.2009.2033804>

⁷⁴ Lee, G., Sparks, R., Ali, S., Shih, N. N., Feldman, M.D., Spangler, E., Rebbeck, T., Tomaszewski, J. E., Madabhushi, A. Co-occurring gland angularity in localized subgraphs: predicting biochemical recurrence in intermediate-risk prostate cancer patients. *PLoS One* 9(5) (2014) e97954. <https://doi.org/10.1371/journal.pone.0097954>

⁷⁵ Huang, C.H., Veillard, A., Roux, L., Loménie, N., Racocceanu, D. Time-efficient sparse analysis of histopathological whole slide images. *Comput. Med. Imaging Graph.* 35(7-8) (2011) 579-591. <https://doi.org/10.1016/j.compmedimag.2010.11.009>

⁷⁶ Hua, J., Huang, M. L., Huang, W., Zhao, C. Applying Graph Centrality Metrics in Visual Analytics of Scientific Standard Datasets. *Symmetry*, 11(1) (2019) 30 <https://doi.org/10.3390/sym11010030>

⁷⁷ Chen, W. K. *Graph theory and its engineering applications*, World Scientific, 1997.

⁷⁸ Chartrand G, Lesniak L, Zhang P *Graphs and digraphs*, 39. CRC Press (2010).

⁷⁹ Estrada E, Bonchev D. Section 13.1. chemical graph theory. In: *Handbook of graph theory* (2013), pp. 1538-1558.

⁸⁰ Heijmans, H. J. A. M., Nacken, P. Toet, A., Vincent, L. Graph morphology. *J. Visual Commun. Image Representation* 3(1) (1990) 24-38. [https://doi.org/10.1016/1047-3203\(92\)90028-R](https://doi.org/10.1016/1047-3203(92)90028-R)

⁸¹ Toet, A. Hierarchical clustering through morphological graph transformation. *Pattern Recognition Letters*, 12(7) (1991) 391 - 399. [https://doi.org/10.1016/0167-8655\(91\)90280-Y](https://doi.org/10.1016/0167-8655(91)90280-Y)

⁸² Theiler, J. Efficient algorithm for estimating the correlation dimension from a set of discrete points. *Phys. Rev. A* 36 (9) (1987) 44-56. <https://doi.org/10.1103/PhysRevA.36.4456>

⁸³ Simpelaere, D. Correlation Dimension. *J. Stat. Phys.* 90 (1998) 491-509. <https://doi.org/10.1023/A:1023232624745>

⁸⁴ Graham, R. L., Pollak, H. O. On the addressing problem for loop switching. *The Bell Syst. Techn. J.* 50(8) (1971) 2495-2519. <https://doi.org/10.1002/j.1538-7305.1971.tb02618.x>

-
- ⁸⁵ Buckley, F., Harary F. Distance in Graphs. Addison-Wesley Pub. Co., New York (1990)
- ⁸⁶ Aouchiche, M. Hansen, P. Distance spectra of graphs: a survey. Linear algebra and its Appl. 458 (2014) 301-386. <https://doi.org/10.1016/j.laa.2014.06.010>
- ⁸⁷ Young, G., Householder, A. Discussion of a set of points in terms of their mutual distances. Psychometrika 3 (1938) 19-22. <https://doi.org/10.1007/BF02287916>
- ⁸⁸ Harriman, J. E. Distance and entropy for density matrices. J. Chem. Phys. 115(20) (2001) 9223-9232. <https://doi.org/10.1063/1.1412877>
- ⁸⁹ Chen, Z., Dehmer, M., Shi, Y. A Note on Distance-based Graph Entropies. Entropy, 16(10) (2014) 5416-5427. <https://doi.org/10.3390/e16105416>
- ⁹⁰ Omar, Y. M., Plapper, P. A Survey of Information Entropy Metrics for Complex Graphs. Entropy, 22(12) (2020) 1417. <https://doi.org/10.3390/e22121417>
- ⁹¹ Mowshowitz, A., Dehmer, M. Entropy and the Complexity of Graphs Revisited. Entropy, 14(3) (2012) 559-570. <https://doi.org/10.3390/e14030559>
- ⁹² Bonchev, D., Trinajstić, N. Information theory, distance matrix, and molecular branching, J. Chem. Phys. 67 (1977) 4517-4533. <https://doi.org/10.1063/1.434593>
- ⁹³ Lü, L., Chen, D., Ren, X., Zhang, Q., Zhang, Y., Zhou, T. Vital vertices identification in complex graphs. Phys. Rep. 650 (2016) 1-63. <http://dx.doi.org/10.1016/j.physrep.2016.06.007>
- ⁹⁴ Muhiuddin, G., Samanta, S., Aljohani, A. F., Alkhaibari, A. M. A Study on Graph Centrality Measures of Different Diseases Due to DNA Sequencing. Mathematics, 11(14) (2023) 3166. <https://doi.org/10.3390/math11143166>
- ⁹⁵ Khodaei, A., Eskandari, S., Sharifi, H., Mozaffari-Tazehkand, B. PRCFX-DT: a new graph-based approach for feature selection and classification of genomic sequences. BMC Bioinformatics 26, 159 (2025). <https://doi.org/10.1186/s12859-025-06183-4>
- ⁹⁶ Samanta, S., Dubey, V.K., Sarkar, B. Measure of influences in social graphs. Appl. Soft Comput. 99 (2021) 106858. <https://doi.org/10.1016/j.asoc.2020.106858>
- ⁹⁷ Das, K., Samanta, S., Pal, M. Study on centrality measures in social graphs: A survey. Soc. Netw. Anal. Min. 8 (2018) 13. <https://doi.org/10.1007/s13278-018-0493-2>
- ⁹⁸ Pandey, S.D., Ranadive, A.S., Samanta, S., Sarkar, B. Bipolar-valued fuzzy social graph and centrality measures. Discret. Dyn. Nat. Soc. 2022 (2022) 9713575. <https://doi.org/10.1155/2022/9713575>
- ⁹⁹ Lokesha, V., Usha, A., Ranjini, P.S., Devendraiah, K.M. Topological indices on model graph structure of Alveoli in human lungs, Proc. Jangjeon Math. Soc. 18(4) (2015) 435-453. <https://doi.org/10.17777/pjms.2015.18.4.002>
- ¹⁰⁰ Lokesha, V., Shruti, R., Deepika, T. Reckoning of the Dissimilar Topological indices of Human Liver. Appl. Math. Nonlin. Sci. 3(1) (2018) 265-276. <https://doi.org/10.21042/AMNS.2018.1.00020>
- ¹⁰¹ Jeong, H., Mason, S. P., Barabási, A. L., Oltvai, Z. N. Lethality and centrality in protein graphs. Nature 411(6833) (2001) 41-42. <https://doi.org/10.1038/35075138>

-
- ¹⁰² Freeman, L. C. Centrality in Social Graphs I: Conceptual Clarification. *Soc. Netw.* 1 (1979) 215-239. [https://doi.org/10.1016/0378-8733\(78\)90021-7](https://doi.org/10.1016/0378-8733(78)90021-7)
- ¹⁰³ Rout, T., Mohapatra, A., Kar, M., Sabyasachi, Patra, S., Muduly, D. Centrality Measures and Their Applications in Graph Analysis: Unveiling Important Elements and Their Impact. *Procedia Comput. Sci.* 235 (2024) 2756-2765. <https://doi.org/10.1016/j.procs.2024.04.260>
- ¹⁰⁴ Albert, R., Jeong, H., Barabasi, A. L. Diameter of the world wide web. *Nature* 401 (1999) 130-131. <https://doi.org/10.1038/43601>
- ¹⁰⁵ Blue, M., Bush, B., Puckett, J. Unified approach to fuzzy graph problems. *Fuzzy Sets Syst.* 125(3) (2002) 355-368. [https://doi.org/10.1016/S0165-0114\(01\)00011-2](https://doi.org/10.1016/S0165-0114(01)00011-2)
- ¹⁰⁶ Freeman, L. C. A set of measures of centrality based on betweenness. *Sociometry*, 40(1) (1977) 35-41. <https://psycnet.apa.org/doi/10.2307/3033543>
- ¹⁰⁷ Sabidussi, G. The centrality index of a graph. *Psychometrika* 31(4) (1966) 581-603. <https://doi.org/10.1007/BF02289527>
- ¹⁰⁸ Bavelas, A. Communication patterns in task-oriented groups. *J. Acoust. Soc. Am.* 22(6) (1950) 725-730. <https://psycnet.apa.org/doi/10.1121/1.1906679>
- ¹⁰⁹ Luce, R. D., Perry, A. D. A method of matrix analysis of group structure. *Psychometrika*, 14 (1949) 95-116. <https://doi.org/10.1007/BF02289146>
- ¹¹⁰ Newman, M. E. J. Random Graphs with Clustering. *Phys. Rev. Lett.* 103 (2009) 058701. <https://doi.org/10.1103/PhysRevLett.103.058701>
- ¹¹¹ Fatima, U., Hina, S., Wasif, M. A novel global clustering coefficient-dependent degree centrality (GCCDC) metric for large graph analysis using real-world datasets. *J. Comput. Sci.* 70 (2023) 102008. <https://doi.org/10.1016/j.jocs.2023.102008>
- ¹¹² Watts, D., Strogatz, S. Collective dynamics of 'small-world' graphs. *Nature*, 393 (1998) 440-442. <https://doi.org/10.1038/30918>
- ¹¹³ Yousuf, M.I., Anwer, I., Anwar, R. Empirical characterization of graph sampling algorithms. *Soc. Netw. Anal. Min.* 13 (2023) 66. <https://doi.org/10.1007/s13278-023-01060-5>
- ¹¹⁴ Bonacich, P. Some unique properties of eigenvector centrality. *Soc. Netw.*, 29(4) (2007) 555-564. <https://doi.org/https://doi.org/10.1016/j.socnet.2007.04.002>
- ¹¹⁵ Bonacich, P., Lloyd, P. Eigenvector-like measures of centrality for asymmetric relations. *Soc. Netw.* 23(3) (2001) 191-201. [https://doi.org/https://doi.org/10.1016/S0378-8733\(01\)00038-7](https://doi.org/https://doi.org/10.1016/S0378-8733(01)00038-7)
- ¹¹⁶ Katz, L. A new status index derived from sociometric analysis. *Psychometrika* 18(1) (1953) 39-43. <https://doi.org/10.1007/BF02289026>
- ¹¹⁷ Guetzkow, H., Simon, H. A. The Impact of Certain Communication Nets Upon Organization and Performance in Task-Oriented Groups. *Manag. Sci.* 1(3-4) (1955) 233-250. <https://doi.org/10.1287/mnsc.1.3-4.233>
- ¹¹⁸ Zhan, J., Gurung, S., Parsa, S. P. K. Identification of top-K vertices in large graphs using Katz centrality. *J. Big Data* 4 (2017) 16. <https://doi.org/10.1186/s40537-017-0076-5>

-
- ¹¹⁹ Stephenson, K., Zelen, M. Rethinking centrality: Methods and examples. *Soc. Netw.* 11(1) (1989) 1-37. [https://doi.org/10.1016/0378-8733\(89\)90016-6](https://doi.org/10.1016/0378-8733(89)90016-6)
- ¹²⁰ Brandes, U., Fleischer, D. Centrality Measures Based on Current Flow. *Proc. 22nd Symp. Theoretical Aspects of Computer Science (STACS '05)*. LNCS 3404, pp. 533-544. Springer-Verlag, 2005. https://doi.org/10.1007/978-3-540-31856-9_44
- ¹²¹ Latora, V., Marchiori, M. *New J. Phys.* 9 (2007) 188. <https://doi.org/10.1088/1367-2630/9/6/188>
- ¹²² Poulin, R., Boily, M.-C., Mâsse, B. R. Dynamical systems to define centrality in social graphs, *Soc. Netw.* 22(3) (2000) 187-220. [http://refhub.elsevier.com/S0370-1573\(16\)30157-0](http://refhub.elsevier.com/S0370-1573(16)30157-0)
- ¹²³ Altmann, M. Reinterpreting graph measures for models of disease transmission, *Soc. Netw.* 15(1) (1993) 1-17. [http://refhub.elsevier.com/S0370-1573\(16\)30157-0](http://refhub.elsevier.com/S0370-1573(16)30157-0)
- ¹²⁴ Hage, P., Harary, F. Eccentricity and centrality in graphs, *Soc. Netw.* 17 (1) (1995) 57-63. [http://refhub.elsevier.com/S0370-1573\(16\)30157-0](http://refhub.elsevier.com/S0370-1573(16)30157-0)
- ¹²⁵ Ilić, A., Gutman, I. Eccentric connectivity index of chemical trees, *MATCH Commun. Math. Comput. Chem.* 65 (2011) 731-744. <https://doi.org/10.48550/arXiv.1104.3206>
- ¹²⁶ Pavlopoulos, G. A., Secrier, M., Moschopoulos, C. N., Soldatos, T.G., Kossida, S., Aerts, J., Schneider, R., Bagos, P. G. Using graph theory to analyze biological graphs. *BioData Mining* 4 (2011) 10. <https://doi.org/10.1186/1756-0381-4-10>
- ¹²⁷ Jamil M. K., Distance-Based Topological Indices and Double Graph. *Iran. J. Math. Chem.* 8(1) (2017) 83-91. <https://doi.org/10.22052/ijmc.2017.43073>
- ¹²⁸ Newman, M. E. J. Assortative Mixing in Graphs. *Phys Rev Lett.* 89 (2002) 208701. <https://doi.org/10.1103/PhysRevLett.89.208701>
- ¹²⁹ Newman, M. Mixing patterns in graphs. *Phys. Rev. E*, 67 (2003) 026126. <https://doi.org/10.1103/PhysRevE.67.026126>
- ¹³⁰ Noldus, R., Van Mieghem, P. Assortativity in complex graphs. *J Complex Netw* 3(4) (2015) 507-542. <https://doi.org/10.1093/comnet/cnv005>
- ¹³¹ Quayle, A., Siddiqui, A., Jones, A. Modeling graph growth with assortative mixing. *Eur. Phys. J. B* 50 (2006) 617-630. <https://doi.org/10.1140/epjb/e2006-00170-5>
- ¹³² Das, K., Gutman, I., Milovanović, I., Milovanović, E., Furtula, B. Degree-Based Energies of Graphs. *Linear Algebra Appl.* 554 (2018) 185-204. <https://doi.org/10.1016/j.laa.2018.05.027>
- ¹³³ Zheng, R., Su, P., Jin. S. Arithmetic-Geometric Matrix of Graphs and Its Applications." *Appl. Math. Comput.* 42 (2023) 127764 1-11. <https://doi.org/10.1016/j.amc.2022.127764>
- ¹³⁴ Kirchhoff, G. Über die Auflösung der Gleichungen, auf welche man bei der untersuchung der linearen verteilung galvanischer Ströme geführt wird. *Ann. Phys. Chem.* 148 (1847) 497-508, <https://doi.org/10.1002/andp.18471481202>
- ¹³⁵ Y. Amanatidis, B. Green, M. Mihail, Graphic realizations of joint-degree matrices. (2008). <https://doi.org/10.48550/arXiv.1509.07076>

-
- ¹³⁶Czabarka, E., Dutle, A., Erdos, P. L., Miklos, I. On realizations of a joint degree matrix. *Discrete Appl. Math.* 181 (2015) 283 - 288. <https://doi.org/10.1016/j.dam.2014.10.012>
- ¹³⁷ Merris, R. Laplacian matrices of graphs: A survey. *Linear Algebra Appl.* 197-198 (1994) 143-176. [https://doi.org/10.1016/0024-3795\(94\)90486-3](https://doi.org/10.1016/0024-3795(94)90486-3)
- ¹³⁸ Estrada, E. *The Structure of Complex Graphs: Theory and Applications*. Oxford University Press, New York (2011).
- ¹³⁹ Anderson, W.N., Morley, T.D. Eigenvalues of the Laplacian of a graph. *Linear and Multilinear Algebra*, 18 (1985) 141-145. <https://doi.org/10.1080/03081088508817681>
- ¹⁴⁰ Lu, W. E, J., Yao, Y. The landscape of complex networks. *arXiv e-prints*, arXiv:1204.6376, 2012.
- ¹⁴¹ Horak, D., Maletić, S., Rajković, M. Persistent homology of complex networks. *J. Stat. Mech.* (2009) P03034. <https://doi.org/10.1088/1742-5468/2009/03/P03034>, <https://doi.org/10.48550/arXiv.0811.2203>
- ¹⁴² Petri, G., Sciamiero, M., Donato, I., Vaccarino, F. Topological strata of weighted complex networks. *PLOS One*, 8(6) (2013) e66506. <https://doi.org/10.1371/journal.pone.0066506>
- ¹⁴³ Rieck, B., Fugacci, U., Lukasczyk, J., Leitte, H. Clique community persistence: A topological visual analysis approach for complex networks. *IEEE Trans. Vis. Comput. Graph.* 24(1) (2018) 822-831. <https://doi.org/10.1109/TVCG.2017.2744321>
- ¹⁴⁴ Munkres J. R. *Elements of algebraic topology*. CRC Press (2018).
- ¹⁴⁵ Zomorodian, A., Carlsson, G. Computing persistent homology. *Discrete Comput. Geom.* 33, 249-274 (2005). <https://doi.org/10.1007/s00454-004-1146-y>
- ¹⁴⁶ Edelsbrunner, H., Letscher, D., Zomorodian, A. Topological persistence and simplification. *Discrete Comput. Geom.* 28, (2002) 511-533. <https://doi.org/10.1007/s00454-002-2885-2>
- ¹⁴⁷ Carlsson, G., Zomorodian, A., Collins, A., Guibas, L. J. Persistence barcodes for shapes. *International Journal of Shape Modeling* 11(2) (2005) 149-187. <https://doi.org/10.1142/S0218654305000761>
- ¹⁴⁸ Edelsbrunner, H., Harer, J. Persistent homology-a survey. *Contemp. Math.* 453 (2008) 257-282. <https://doi.org/10.1090/conm/453/08802>
- ¹⁴⁹ Robins, V., Saadatfar, M., Delgado-Friedrichs, O., Sheppard, A. P. Percolating length scales from topological persistence analysis of micro-CT images of porous materials. *Water Resour. Res.* 52(1) (2016) 315-329. <https://doi.org/10.1002/2015WR017937>
- ¹⁵⁰ Lee, H., Kang, H., Chung, M. K., Kim, B.-N., Lee, D. S. Persistent brain network homology from the perspective of dendrogram. *IEEE Trans. Med. Imag.* 31(12) (2012) 2267-2277. <https://doi.org/10.1109/tmi.2012.2219590>
- ¹⁵¹ Chung, M. K., Bubenik, P., Kim, P. T. Persistence diagrams of cortical surface data. In *International Conference on Information Processing in Medical Imaging*. Springer, 2009, pp. 386-397. https://doi.org/10.1007/978-3-642-02498-6_32

-
- ¹⁵² Sizemore, A. E., Giusti, C., Kahn, A., Vettel, J. M., Betzel, R. F., Bassett, D. S. Cliques and cavities in the human connectome. *J Comput Neurosci* 44 (2018) 115-145 <https://doi.org/10.1007/s10827-017-0672-6>
- ¹⁵³ Bendich, P., Marron, J., Miller, E., Pieloch, A., Skwerer, S. Persistent homology analysis of brain artery trees. *Ann. Appl. Stat.* 10(1) (2016) 198-218. <https://doi.org/10.1214/15-AOAS886>
- ¹⁵⁴ Gamble J., Heo, G. Exploring uses of persistent homology for statistical analysis of landmark-based shape data. *J. Multivar. Anal.* 101(9) (2010) 2184-2199. <https://doi.org/10.1016/j.jmva.2010.04.01>
- ¹⁵⁵ Mamuye, A. L., Rucco, M., Tesei, L., Merelli, E. Persistent Homology Analysis of RNA. *Comput. Math. Biophys.* 4(1) (2016) 14-25. <https://doi.org/10.1515/mlbmb-2016-0002>
- ¹⁵⁶ Zomorodian, A., Carlsson, G. Computing persistent homology. *Discrete Comput. Geom.* 33(2) (2005) 249-274. <https://doi.org/10.1007/s00454-004-1146-y>
- ¹⁵⁷ Ghrist, Robert. "Barcodes: the persistent topology of data." *Bulletin of the American Mathematical Society* 45(1) (2008): 61-75.
- ¹⁵⁸ Zomorodian, A. Fast construction of the Vietoris-Rips complex. *Comput. Graph.* 34(3) (2010): 263-271. <https://doi.org/10.1016/j.cag.2010.03.007>
- ¹⁵⁹ Otter, N., Porter, M. A., Tillmann, U., Grindrod, P., Harrington, H. A. A roadmap for the computation of persistent homology. *EPJ Data Science* 6(1) (2017) 17. <https://doi.org/10.1140/epjds/s13688-017-0109-5>
- ¹⁶⁰ Chintakunta, H., Gentimis, T., Gonzalez-Diaz, R., Jimenez, M. J., Krim, H. An entropy-based persistence barcode. *Pattern Recognition* 48(2), 391-401 (2015). <https://doi.org/10.1016/j.patcog.2014.06.023>
- ¹⁶¹ Merelli, E., Rucco, M., Sloot, P., Tesei, L. Topological characterization of complex systems: Using persistent entropy. *Entropy* 17(10), (2015) 6872-6892. <https://doi.org/10.3390/e17106872>
- ¹⁶² Rucco, M., Castiglione, F., Merelli, E., Pettini, M. (2016). Characterisation of the Idiotypic Immune Network Through Persistent Entropy. In: Battiston, S., De Pellegrini, F., Caldarelli, G., Merelli, E. (eds) *Proceedings of ECCS 2014. Springer Proceedings in Complexity*. Springer, pp. 117-128. Springer (2016). https://doi.org/10.1007/978-3-319-29228-1_11
- ¹⁶³ Xia, K. L., Li, Z. M., Mu, L. Multiscale persistent functions for biomolecular structure characterization. *Bull. Math. Biol.* 80(1) (2018) 1-31. <https://doi.org/10.1007/s11538-017-0362-6>
- ¹⁶⁴ Ghrist, R. Barcodes: the persistent topology of data. *Bull. Amer. Math. Soc.* 45(1) (2007) 61-75. <https://doi.org/10.1090/S0273-0979-07-01191-3>