1 **A quantitative three-dimensional (3D) image analysis tool for maximal acquisition of spatial**

2 **heterogeneity data**

3

4 Mark C. Allenby, MSc[1], Ruth Misener, PhD[2], Nicki Panoskaltsis, MD PhD [1,3], Athanasios Mantalaris,

5 PhD[1,+]

6 [1]Biological Systems Engineering Laboratory, Department of Chemical Engineering, [2]Department of

7 Computing, [3]Department of Hematology, Imperial College London, London, UK

8

9

10

11

12 Mark C. Allenby, MSc

13 Biological Systems Engineering Laboratory

14 Department of Chemical Engineering, South Kensington Campus

15 Imperial College London, London, UK, SW7 2AZ

16 Telephone: +44 (0)20 7594 6407

17 Email: mark.allenby@imperial.ac.uk

18

19 Ruth Misener, PhD

20 Huxley Building

21 Department of Computing, South Kensington Campus

22 Imperial College London, London, UK, SW7 2AZ

23 Telephone: +44 (0)20 7594 8315

24 Email: r.misener@imperial.ac.uk

25

26

1   Nicki Panoskaltsis, MD, PhD:

2   Department of Hematology

3   Imperial College London, Northwick Park & St. Mark's campus, London, UK, HA1 3UJ

4   Telephone: +44 (0)20 8869 2742

5   Fax: +44 (0) 208 869 2745

6   Email: n.panoskaltsis@imperial.ac.uk

7

8

9   [+]Corresponding Author:

10   Athanasios Mantalaris, PhD

11   ACEX 511, Biological Systems Engineering Laboratory

12   Department of Chemical Engineering, South Kensington Campus

13   Imperial College London, London, UK, SW7 2AZ

14   Telephone: +44 (0)20 7594 5601

15   Fax: +44 (0)20 75945638

16   Email: a.mantalaris@imperial.ac.uk

1 **Abstract**

2   Three-dimensional imaging techniques provide spatial insight to environmental and cellular

3 interactions and are implemented in various fields, including tissue engineering, but have been

4 restricted by limited quantification tools which misrepresent or underutilize the cellular phenomena

5 captured. This study develops image post-processing algorithms pairing complex Euclidean metrics

6 with Monte Carlo simulations in order to quantitatively assess cell and microenvironment spatial

7 distributions while utilizing, for the first time, the entire 3D image captured. While current methods

8 only analyze a central fraction of presented confocal microscopy images, the proposed algorithms

9 can utilize 210% more cells to calculate 3D spatial distributions which can span a 23-fold longer

10 distance. These algorithms seek to leverage the high sample cost of 3D tissue imaging techniques by

11 extracting maximal quantitative data throughout the captured image.

1 **Introduction**

2   Native and engineered tissues maintain spatial architecture of cells, extracellular matrix

3 proteins and humoral factors via interactive adhesive motifs, paracrine niche interaction, and

4 soluble factor gradients. The interplay of these factors promotes *de novo* cell and tissue formation,

5 pathogen response, cellular chemotaxis, and tissue homeostasis [1-5]. Developments for the *in situ*

6 capture of these cell-environment processes are focused on designing instruments and probes to

7 enhance imaging depth, spatial resolution, speed, and specificity [6-10]. However, the quantitative

8 characterization of imaged tissue environments is often constrained by spatially selective or

9 underutilized image analyses that discard data that are expensive to acquire, which inherently limits

10 the applicability of these algorithms [11-16]. Therefore, image analysis tools that enable quantitating

11 spatial heterogeneity by unbiasedly utilizing the complete tissue sample image thus maximizing

12 information output are required.

13   Quantitative image analysis has been applied in various systems to determine how spatial

14 tissue organization influences cell activity and fate. Studies include *in vivo* animal models, centered

15 around intravital imaging at continuous but short (<24 hours) durations limited to the few animals or

16 tissues (zebrafish, murine calvarial marrow) able to be imaged while alive [17-21], or *in vitro* 2D cell

17 cultures appropriate for studies of spatial organization and cell-cell association [22-24]. Most *in vivo*

18 and *in vitro* 3D cultures are not able to implement intravital imaging throughout growth area of

19 interest and require invasive techniques or culture termination to characterize the formation of cell-

20 environment tissue networks *in situ* [15-16, 25-29]. Maximizing extraction of data from these rare,

21 expensive 3D analyses has been a major driver to develop new imaging instruments and techniques

22 rather than the postimaging algorithms that quantitate their output [6-10].

23   Herein, quantitative tools are developed to analyze 3D spatial heterogeneity throughout the

24 entire captured image in order to maximize data output for expensive, destructive imaging analyses.

25 They work by dividing the full image space into regions, counting the number of cells within each

26 region, and calculating the volume of each region so that regional cell density can be determined.

1   The novel algorithms enable the assessment of complex volumes in regions towards the periphery of

2   the image thus increasing analysis distance (range) as well information content (number of cells

3   surveyed). Nonetheless, the proposed 3D analysis represents an estimation whose accuracy is

4   dependent on user-defined inputs. Therefore, results must be validated against controls, such as

5   regions of known volume and data of known density.

6   Two algorithms to quantify full-image spatial heterogeneity are presented and applied to

7   immunofluorescent (IF) images of a 3D culture system consisting of a synthetic scaffold impregnated

8   with hollow fibres [30]. The first algorithm quantifies the distribution of cellular density from a

9   region of interest and is applicable to recent *in vivo* studies, including the measurement of oxygen

10 and cellular phenotype, stress, and depiction of regional niches of cell activity and migration [14-19].

11 The second algorithm quantifies cellular association by calculating the distribution of cellular density

12 from each cell imaged. Cell association metrics have been useful to assess roles of cell-cell

13 communication, cell attraction and repulsion, as well as cell co-localization [11-13, 23, 29]. While

14 current cellular distribution algorithms only analyze a central fraction of imaged data, the method

15 presented herein allows for the utilization of all cells and environmental space imaged [11-13].

16 Specifically, the presented algorithms utilize 3.6-fold more 2D area or $2.5 \cdot 10^{4}$-fold more 3D volume

17 and survey 1116 cells instead of 235 cells using the 2D analysis or 524 cells using the 3D analysis to

18 help leverage the high cost of tissue culture.

1 **Materials and Methods**

2 *3D Cultures*

3 Mononuclear cells (MNCs) from human umbilical cord blood were cultured in a hollow fiber

4 bioreactor as previously described [30]. In brief, a polyurethane scaffold was formed around ceramic

5 hollow fibers and coated with collagen. Cord blood MNCs were inoculated into the scaffold while

6 hollow fibers were perfused with serum-free medium for 28 days, when a representative SEM to

7 demonstrate bioreactor topology was taken (Figure 1A).

8 *Immunofluorescent Sample Preparation*

9 For analyses, the 3D cultures were snap-frozen in liquid nitrogen and preserved at $-80^{o}C$

10 until sectioning. Sectioning occurred at $-20^{o}C$ into thick histological sections of approximately 1mm.

11 The frozen sections were directly fixed in ice cold 4% paraformaldehyde (BDH Laboratory Sciences,

12 Poole, Dorset, UK) in PBS (Life Technologies, Paisley, UK) overnight followed by a 2-hour

13 permeabilization with 0.1% Triton X-100 (Sigma-Aldrich, Poole, Dorset, UK) in staining buffer [1%

14 BSA (Sigma-Aldrich), 0.5% Tween-20 (Promega, Southampton, Hampshire, UK), 0.01% NaN3 (Sigma-

15 Aldrich) in PBS]. This was followed by a 4-hour blocking step with 10% donkey serum (AbCam,

16 Cambridge, UK) in staining buffer and overnight incubation with primary antibodies and isotype

17 controls (Table 1; AbCam) in staining buffer at $4^{o}C$. Thereafter, 6-hour staining was undertaken with

18 secondary antibodies: donkey anti-rat Alexa Fluor 488, donkey anti-sheep Alexa Fluor 488 (AF488),

19 donkey anti-rabbit Alexa Fluor 555 (AF555), and donkey anti-mouse Alexa Fluor 647 (AF647) all at

20 1:500 dilution in staining buffer at $4^{o}C$ (Life Technologies). These steps were followed by a DAPI

21 counterstain (Fisher Scientific, Loughborough, Leics, UK) overnight at 50 µM in PBS at $4^{o}C$, and

22 samples were stored in 0.01% NaN_3 in PBS at $4^{o}C$. Each step was separated by single or multiple 15

23 minute washing steps in appropriate buffer.

1 **Table 1:** Confocal 3D immunofluorescent microscopy primary antibody staining protocol. All
2 products were obtained from AbCam with product number stated. The two samples presented in
3 Figures 1-5 were prepared alongside appropriate isotype and unstained controls.

Fig. 1B	Fig. 1B Isotype	Fig. 1C	Fig. 1C Isotype	Unstained
1.7 µg/mL rat anti-human glycophorin-A (ab33386)	1.7 µg/mL rat IgG2b (ab18437)	6.7 µg/mL sheep anti-human Stro-1 (ab192766)	6.7 µg/mL sheep IgG (ab37385)	None
8.3 µg/mL rabbit anti-human C-Kit (ab111033)	8.3 µg/mL rabbit IgG (ab172730)	8.3 µg/mL rabbit anti-human osterix (ab94744)	8.3 µg/mL rabbit IgG (ab172730)	None
33.3 µg/mL mouse anti-human erythropoietin-receptor (ab56310)	33.3 µg/mL mouse IgG2a (ab91361)	3.33 µg/mL mouse anti-human osteopontin (ab69498)	33.3 µg/mL mouse IgG2a (ab91361)	none

4

5
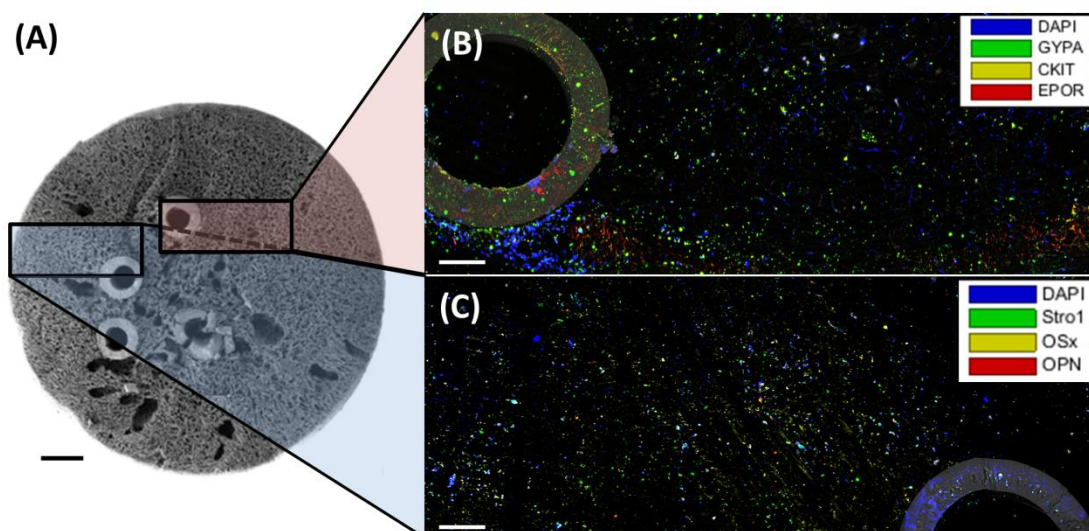
6

7 **Figure 1:** Imaging regions of the 3D culture device. (A) The topology of the hollow fiber bioreactor as

8 a representative SEM cross-section, and (B, C) surveyed 5-color confocal images. Scale bar in (A) is 1

9 mm and in (B) and (C) are 100 µm, and bioreactor material reflection is represented in greyscale.

10 Markers analyzed were DAPI for cell nuclei, and cell phenotypes: glycophorin-A (GYPA),

11 erythropoietin receptor (EPO-R) and stem cell growth receptor (c-kit) in Fig. 1B, and Stro-1 (Stro1),

12 osterix (OSx), and osteopontin (OPN) in Fig. 1C.

1 *Confocal Microscopy Image Acquisition*

2 The fixed, stained microscope sections were imaged on a Leica SP5 inverted confocal

3 microscope with Leica LAS AF software (Leica, Milton Keynes, UK) using 405, 453, 488, 543, and 633

4 nm lasers and filters for DAPI, reflectance, AF488, AF555, and AF647 stains in 2 sequences of

5 excitation for 2- then 3-color collection without any detectable spill-over. Images were taken at

6 512x512 pixel resolution using a 10x dry microscope lens for a resolution of approximately 3.03 μm

7 per pixel with Z-stacks acquired in 5 μm slices, selected as the minimum lens magnification to ensure

8 single-cell stain recognition [19]. All laser voltage, filter wavelength, and capture settings were

9 identical for corresponding samples and controls. Each sample was captured as three adjacent 3D

10 images with 7% overlap to allow collation of three individual 1551 x 1551 μm^2 square images into a

11 4437 x 1551 μm^2 rectangular image with a depth of 350 μm for Figure 1B, and 230 μm for Figure 1C.

12 Unstained and isotype controls were captured under identical conditions. Images were not

13 manipulated.

14 *Cell Localization*

15 Once captured, images were analyzed using the Fiji image processing package of ImageJ: (1)

16 the *grid/collection* package collated together adjacent 3D images, (2) the *subtract background*

17 package decreased auto-fluorescence and self-absorbance, and, finally (3) the *3D object counter*

18 package located the center of each individual fluorescence stain, *$(X,Y,Z)_c$,* which was exported into

19 MATLAB for further analysis (The MathWorks, Inc., Natick, USA) [31-33]. Representative confocal

20 images are shown in Figure 1B and 1C in Table 2, with stain localization for Figure 1B in Figure 2A

21 and 2B. The presence of each fluorescent antibody stain, *$(X,Y,Z)_c$,* was biologically validated in

22 comparison to isotype controls, which contained less than 1% of detected sample stains. All

23 computational post processing was run on a 3.4 GHz machine with 8 GB RAM and an Intel® Core™ i7-

24 4770 CPU. Processing times are provided and correspond as 1 second to $1.2 \cdot 10^8$ random points

25 sampled from a uniform distribution in MATLAB on this machine.

26 *2D Distance Distribution*

1   Assessing cell density as a function of distance from an environmental region of interest is

2   useful in examining cellular gradients across tissue sections [14-19]. In order to determine cellular

3   distribution in 2D, 3D confocal image z-stacks (Figure 1B) were compressed into 2D so that a given

4   line of interest running perpendicular to the image plane becomes a point $(X,Y)_a$. Cellular

5   heterogeneity was studied as a function of XY distance from this point of interest (Figure 2C and 2D),

6   where individual cell position $(X,Y)_c$ and its distance from the point of interest $(X,Y)_a$ was defined as

7   2D Euclidean distance:

8    $$R_c = \sqrt{(X_c - X_a)^2 + (Y_c - Y_a)^2}$$ [1]

9   The full image field distance was defined as the distance from the point of interest to the furthest

10   image pixel, R_{max}, and was then discretized into bins of radii which form concentric circles of area

11   $\pi(R_i^2 - R_{i-1}^2)$. Whenever the circle of radius R_i centered at $(X,Y)_a$ encompassed space outside of

12   the rectangular image limits, the extraneous area, A_e, was subtracted off for each breached side. In

13   addition, if two of these breached sides overlapped to form a vertex of the rectangular limits, this

14   overlapping extraneous area, A_{ee}, must be added back. The resulting equation becomes:

15 $$A_i = \pi\left(R_i^2 - R_{i-1}^2\right) - A_e + A_{ee}$$ [2]

16   Areas A_e and A_{ee} were arc angles of off-center circle area segments that were analytically calculated

17   through a series of trigonometric identities (refer to Supplementary Information) and all terms

18   correspond to Figure 3A and are summarized in Table 2. An illustration comparing concentric circle

19   areas against concentric circle areas bounded to image limits (A_i) used in Figure 2 is demonstrated in

20   Figure 3B.

21   In order to assess complex 3D geometries that can also be validated against analytical 2D

22   solutions, Monte Carlo methods have been employed to estimate complex geometrical spaces by

23   generating points from a uniform distribution within a measureable region, such as the complete

24   image volume [34]. However, this method may be inaccurate when insufficient points are generated.

25   Therefore, in the 2D case, *estimation error* was defined as the percent difference of the Monte Carlo

26   estimation from the analytical solution. As there was no analytical solution for many 3D cases,
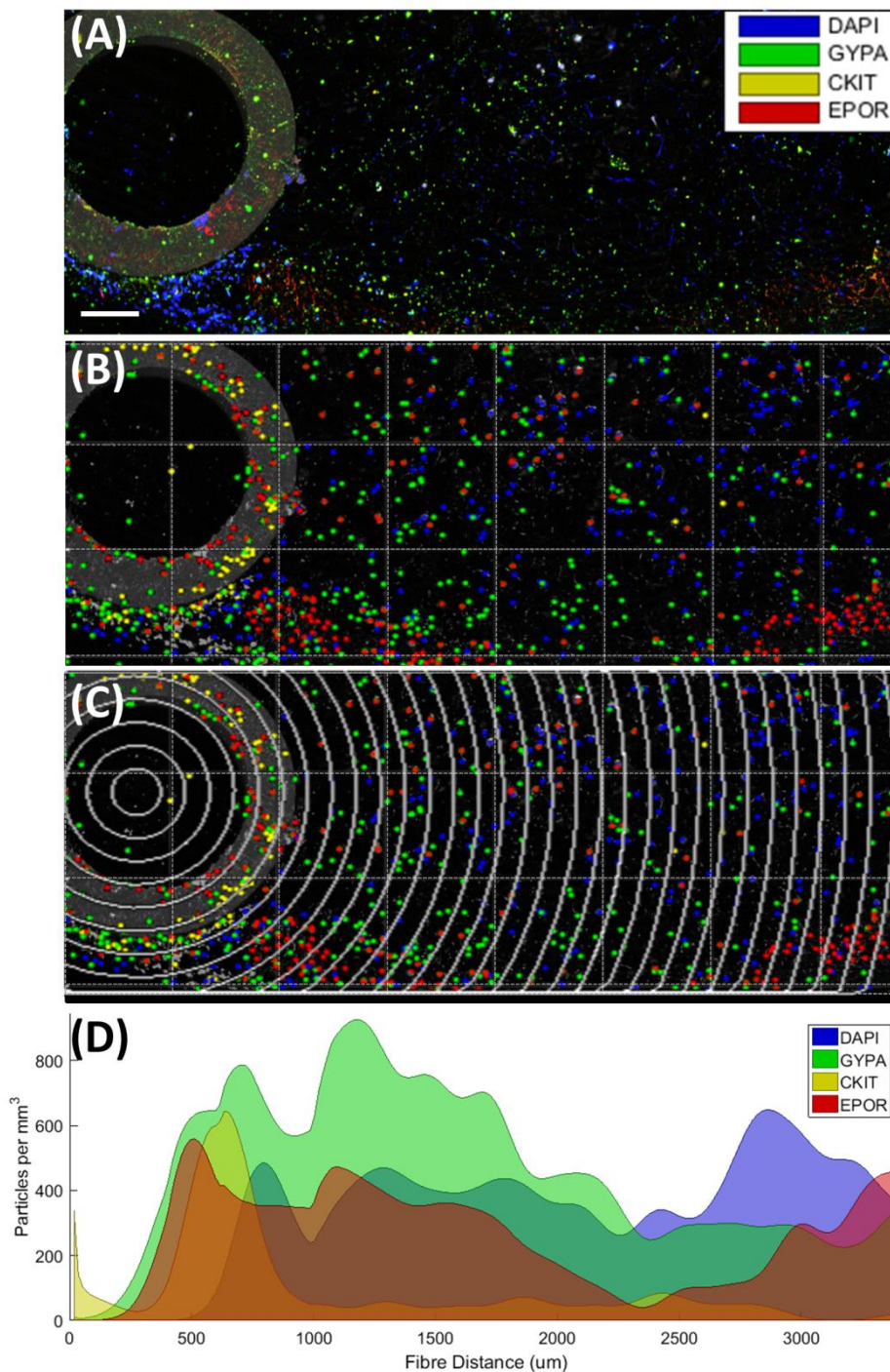
1

2 **Figure 2**: A visualization of the 2D distance cell density analysis. (A) Immunofluorescent images were

3 captured and (B) fluorescent markers were identified by object-based co-localization. The fiber

4 center was defined and (C) volumes of concentric spaces confined to image limits were calculated

5 explicitly or estimated probabilistically in order to determine (D) cell and phenotype density at all

6 distances captured by confocal microscopy. Scale bar in (A) is 100 µm.
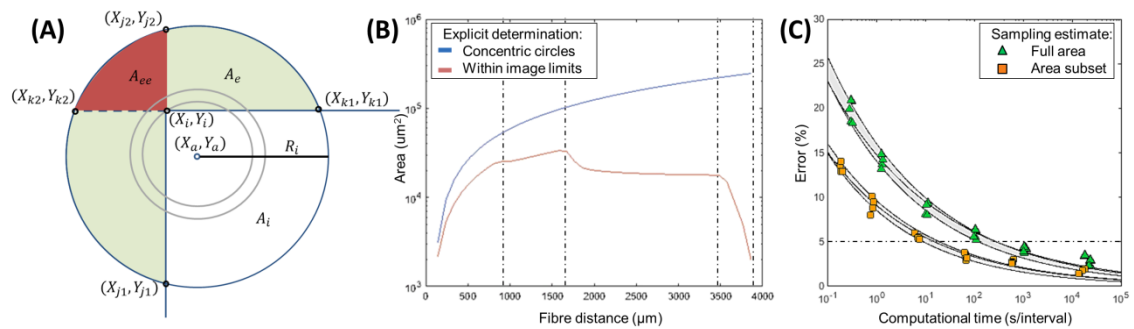
1

2 **Figure 3**: 2D tissue distance cell density analysis. (A) Mathematical notations used for 2D concentric

3 neighborhoods surrounding the point of interest. (B) A visualization of concentric circle area (blue)

4 and the area confined to discontinuous image limits (red) with iterative radii of 50 μm. (C) Monte

5 Carlo estimation error from explicit solutions for successively tighter iterative Monte Carlo

6 convergences (10^{-1},…,10^{-6}) averaged over 4 image geometries and 40 distance intervals per image

7 inside (green) the full image cuboid or (orange) the smallest subset cuboid where trend lines were

8 fitted to $E = k_1 e^{-k_2 \ln t}$ for the mean, m_E, and the 95% confidence interval, $m_E \pm 1.96 \frac{\sigma_E}{\sqrt{n}}$ for n=4,

9 shaded in grey.

---

11 *Monte Carlo convergence* was defined as the percent difference between estimated Monte Carlo

12 solutions when 10^4 additional uniformly generated points were appended [35].

13   The Monte Carlo method was validated for the typical 2D distance distribution (Figure 3C) by

14 calculating the average *estimation error* and *Monte Carlo convergence* across 4 image geometries,

15 each broken into 40 equidistant intervals from a point of interest. An estimation error of 2.75% from

16 the explicit calculation was reached for 10^{-6} % Monte Carlo convergence in iterations of 10^4 points,

17 but required 8.6 minutes. A 10^{-2} % Monte Carlo convergence, which corresponded to an average of

18 2.4 x 10^6 generated points or 0.02 seconds per distance interval with estimation error below 5% in

19 the 2D distance distribution, was utilized for all subsequent distance distribution analyses.

20 ***3D Distance Distribution***

1   The analysis of 3D cell distributions away from a specific point of interest is often utilized for

2   cell co-localization studies, but is currently limited to cells either near the center of the image or only

3   for very close cell proximities, creating regional bias in analysis [11-12, 23, 29]. To analyze 3D cell

4   distribution from a point of interest $(X, Y, Z)_a$ throughout the entire image, the distance from

5    $(X, Y, Z)_a$ to the furthest pixel imaged was divided into intervals, R_i. Cell $(X, Y, Z)_c$ exists within the

6   interval (R_{i-1}, R_i) if:

7    $$R_{i-1} < R_c = \sqrt{(X_c - X_a)^2 + (Y_c - Y_a)^2 + (Z_c - Z_a)^2} \leq R_i \qquad\qquad [3]$$
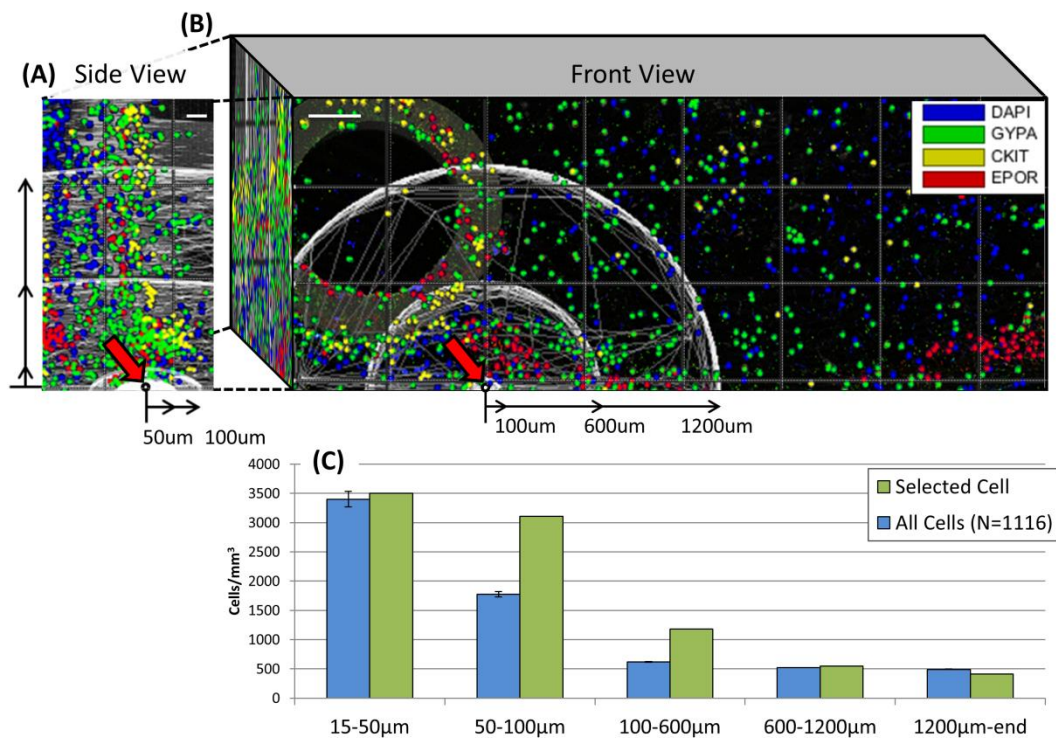
8
9   **Figure 4**: A visualization of the 3D cell clustering density analysis in confocal images. (A, B) A cell of

10   interest was defined and volumes of concentric spheres confined to image limits were estimated

11   through point sampling in order to determine (C) cell density at any imaged neighborhoods away

12   from a given cell (red arrow, green bars), which can then be iterated for all 1,116 (+/- SEM) cells

13   throughout the image (blue bars). Scale bars in (A, B) are 100 μm.

1   Therefore, cell density within neighborhood (R_{i-1}, R_i) can be calculated as the quantity of all cells

2   within (R_{i-1}, R_i) divided by the neighborhood's volume, V_i. If this neighborhood exists entirely

3   within image boundaries then the neighborhood volume is $V_i = \frac{4}{3}\pi(R_i{}^3 - R_{i-1}{}^3)$. However, if a portion of

4   this neighborhood exists outside of the image boundaries, then calculation of V_i may not hold an

5   explicit solution. In this case, Monte Carlo point sampling methods were implemented to estimate

6   volume V_i: random 3D points $P_N = (X, Y, Z)_N$ were generated from a uniform distribution (known as

7   Monte Carlo-sampling) across the entire imaged space, such that $P_N \in [(0,0,0), (X,Y,Z)_L]$. Then, P_n

8   is defined as the fraction of P_N within the sphere volume of interest, such that $P_n \in ((X,Y,Z)_a \pm$

9    $R_i)$, and an estimate of concentric neighborhood volume is provided as:

$$V_i = V_L \frac{\sum P_n}{\sum P_N} - V_{i-1} \qquad [4]$$

10

11   Where, V_L is the full image volume and the concentric neighborhood volume of the prior

12   radial interval R_{i-1} is given by V_{i-1}. This process of uniform point generation and volume calculation

13   was repeated and appended until iterations of calculated V_i vary less than a chosen Monte Carlo

14   convergence limit, as discussed above. The estimation of a 100 μm cell neighborhood (100 μm, 4.2 x

15   10^{-3} mm^3) inside a large image volume, V_L (2.4 mm^3), required sampling 5.8 x 10^7 uniformly

16   distributed random points and 0.49 seconds, or 9.1 minutes if repeated for the 1,116 cells within

17   Figure 1B.

18   Computational time decreased by 96% from 9.1 minutes to 22 seconds when Monte Carlo

19   point sampling methods were applied within small regions around each cell rather than across the

20   entire image space. A smaller cuboid of known volume, V_{LS}, was constructed to exactly encase the

21   cell neighborhood (volume $(2R_i)^3$ centered on cell $(X,Y,Z)_c$) that was then confined within the

22   image limits ($V_{LS} \leq 8$ x 10^{-3} mm^3 for R_i = 100 μm cell neighborhood). Points, P_{NS}, were sampled

23   within the cuboid sub-volume such that $P_{NS} \in V_{LS} \subset V_L$, and the fraction of P_{NS} within the sphere

1   volume of interest, $P_{nS} \in ((X, Y, Z)_a \pm R_i)$, provided a more efficient estimate of concentric

2   neighborhood volume than sampling points within the entire image limits as:

3
$$V_i = V_{LS} \frac{\sum P_{nS}}{\sum P_{NS}} - V_{i-1} \qquad [5]$$

4   Fewer sampled points were required within V_{LS} compared to V_L to reach the same Monte

5   Carlo convergence: 2.4 x 10^6 uniformly distributed random points required 0.02 seconds per cell

6   neighborhood, or 22 seconds for all 1,116 cells in Figure 1B. Using the proposed method, 3D cell

7   density distributions can be quickly analyzed at distances up to half the diagonal image distance, or

8      $0.5\sqrt{X_L^2 + Y_L^2 + Z_L^2}$, for every cell present in the image..

9   Alternatively, an equally spaced 3D grid of k points across the image volume, V_L, or a similar

10   image subvolume, V_{LS}, can be used to estimate cell neighborhoods. However, iterative grid

11   generations must be performed with $(2k)^3$ additional points, which greatly increases computational

12   time, and is discussed as a different method to random point generation from a uniform distribution

13   (refer to Supplementary Information, Figure S2).

14   ***Statistics***

15   Quantitative results are represented as mean ± standard error of 4 replicate experimental samples

16   prepared and assessed identically (Figure 1B), or 4 replicate computationally generated samples as

17   stated.

18   **Table 2:** Mathematical parameters.

Parameter	Description	Ref
$(X,Y)_c$; $(X,Y,Z)_c$	2D/3D cell position	[1], [3]
$(X,Y)_a$; $(X,Y,Z)_a$	Line or point of interest	[1], [3]
R_f	Cell to region of interest distance	[1], [3]
R_{max}	Maximum surveyable distance	-
A_i	Examined region area	[2]
R_i; R_{i-1}	Bin of examinable region	[2]
A_e; A_{ee}	Extraneous binned region area	[2]
V_L; $(X,Y,Z)_L$	Total image volume; limits	[4], [5]
V_i; V_{i-1}	Examined region volume	[4], [5]
V_{LS}; $(X,Y,Z)_{LS}$	Subset volume; limits	[5]

19

1   **Results**

2   ***Enhanced Utilization of Imaged Data***

3   The presented algorithms completely assess 2D and 3D cell density distributions of imaged

4   histological data whereas current methods disregard data near image boundaries [11-13]. Therefore,

5   the algorithms herein enhance utilization of image data in two ways: a) by surveying more points of

6   interest, and b) by probing further distances from each point of interest. With current approaches,

7   an ideally centered point of interest within Figure 1B could only be inspected up to a maximum 2D

8   distance of 775 μm (the closest image boundary); in contrast, using the presented algorithms, a 2D

9   distribution of at minimum 2300 μm (the furthest imaged pixel) can now be calculated from any

10   point within the image.

11   Cell density distribution algorithms are typically repeated for multiple points, or cells, of

12   interest. A conservative comparison with current methods was performed by analysing 3D cell

13   density distributions within a 100 μm neighborhood around every cell to demonstrate cell to cell

14   association in Figure 1B (similar to Bjornsson *et al* [11]). Current methods utilize only cells further

15   than 100 μm from an image boundary as points of interest (N=524) while with the presented

16   approach all imaged cells were utilised (N=1116) at a 23-fold greater distance. The inclusion of these

17   cells improved cell density estimations by 26% and 23% within 15-50 and 50-100 μm intervals,

18   respectively.

19   Accurate manual scoring allows for the exact measurement of cell positions and cell-to-cell

20   distances, but would be infeasible to manually pinpoint thousands of cells in a large image and

21   infeasible to measure the hundreds of thousands of intracellular distances and the volumetric

22   neighborhoods to calculate cellular density distributions within. The proposed algorithm was

23   compared to manual scoring of random fractions of the entire image in Figure 1B, which revealed

24   that partial manual scoring reduced accuracy due to the non-uniform distribution of cells and was

1 unable to analyze cell distributions at long distances as described in the Supplementary Information,

2 while still reducing analysis time from 30 min down to 1 min.

3 *Random Sampling Accuracy*

4   The point sampling algorithm was validated by assessing uniformly generated random cell

5 distributions, as shown in Figure 5A. These data sets were comprised of 4 types of 500

6 computationally generated cells which were uniformly distributed within a 4437 x 1551 x 230 μm^3

7 space in order to produce a known cell density while replicating the number of cells and image

8 geometry of experimental samples, as performed in current studies [14-15]. The sampling method

9 accurately analyzed the uniform distribution of cells, where 2,000 uniformly generated cell positions

10 were found to have an average of 1257.6 $\pm$ 44.8, 1244.8 $\pm$ 22.4, and 1297.2 $\pm$ 2.0 cells per mm^3 in

11 15-50, 50-100, and 100-600 μm neighborhoods of all other cells, and 1297.0 $\pm$ 0.2 cell per mm^3

12 throughout the remaining 2300 μm of the image space (n = 2,000 cells). Small 50 μm and 100 μm

13 neighborhoods contained significantly smaller numbers of cells, and may have accounted for their

14 higher variability, but even so, the density of cells was found to closely approximate a uniform

15 distribution of 2000 cells in the image volume of 1.54 mm^3.

16 **3D Distribution with Regional Asymmetries**

17   Many histological images contain gaps or spaces without cells, such as vessels, areas of

18 decalcified bone, or the edge of the sectioned sample. These voids may underestimate cell density

19 during computational analysis if they are averaged into dense cellular neighborhoods. However, if

20 significantly larger numbers of cells could be assessed with full image field analysis, regional

21 asymmetries might be more accurately determined. Cell clustering density analyses were run on an

22 identical image including or excluding regions with engineered void space (Figure 5B and 5C), and

23 showed the region containing the voidage played a relatively minor role in influencing paracrine cell-

24 cell niche communication, but a much greater regional effect correlating to its inclusion or exclusion

25 from the 2D distance cell density analysis. There was a 7.8%, 11.1%, 44.6% difference in cell density

26 in 15-50, 50-100, and 100-600 μm neighborhoods between the image with and without the fiber,

1 and a 46.2% difference in cell density throughout the rest of the image, demonstrating that this

2 regional phenomena altered cell localization at large 2D tissue scales, but cell association at close

3 intercellular distances remained consistent throughout the image.
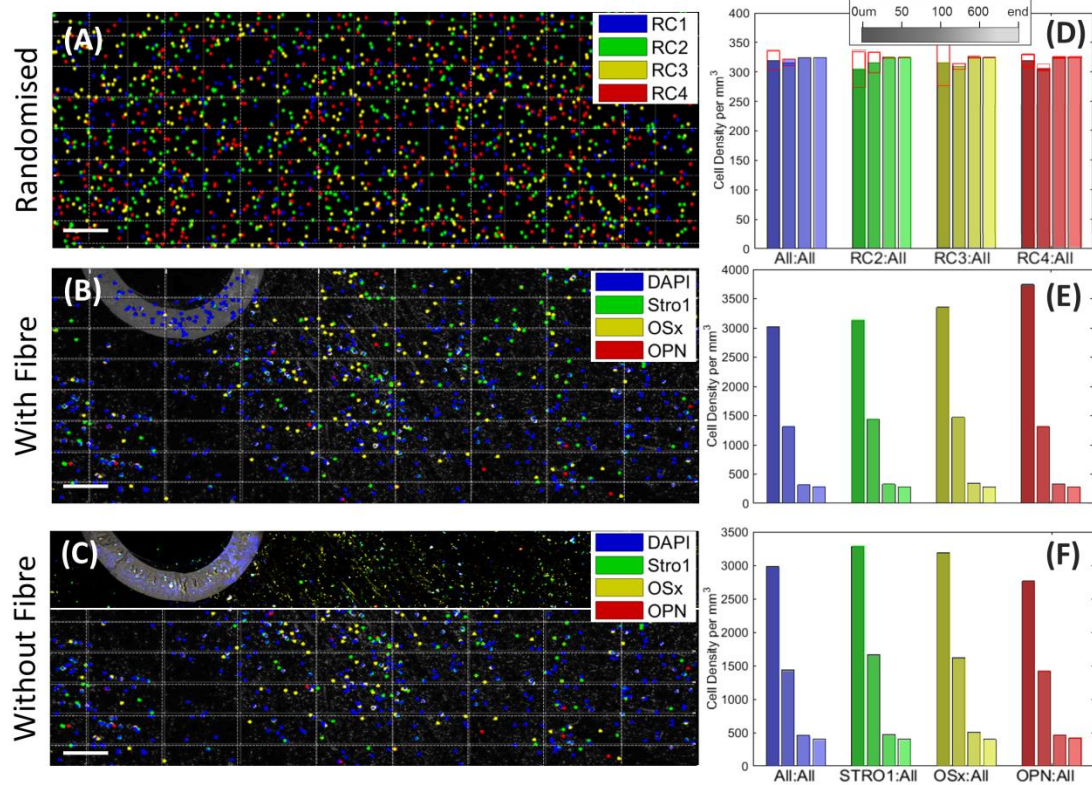
4

5 **Figure 5**: Assessment of methods and applications for calculating 3D cell clustering density.

6 (A) Cell clustering density estimation for 4 iterations (+/- SEM, red) of 4 types of 500 computationally

7 generated cells, or, for (B) cells imaged from a hollow fiber bioreactor with and (C) without the

8 hollow fiber region, representing a spatial asymmetry. Scale bars in (A, B, C) are 100 µm. Objects

9 analyzed include randomized control cells (RC1-RC4), as well as fluorescent markers.

1   **Discussion**

2   3D image analyses enable the acquisition of information on spatial cell and tissue

3   heterogeneity due to their quantitative algorithms [6-7]. However, current calculations of cell

4   distributions are often constrained to a central subset of the image provided, and are further limited

5   when phenomena of interest are off-center or replicate events per image are assessed. When the

6   entire tissue volume of interest is infeasible to image, recent studies have compensated by

7   quantitating images of incomplete tissue subsections either through normalization against uniformly

8   generated random data within a similar tissue topology [13-15], or by only analysing phenomena

9   that exist within the most central subsection of the image [11-12, 16]. The former method only

10   provides relative quantitation of images while the later discards nearly half of 3D imaged data [11-

11   12]. The presented analyses in this study quantitate 2D and 3D distributions of phenomena

12   throughout any full image, providing a minimum 50-fold increase in analytical range (increased

13   distance times increased cells analyzed) throughout coupling complex Euclidean metrics with Monte

14   Carlo point sampling estimations.

15   The analysis of cell distributions can be parametrically tuned for computational speed and

16   analytical resolution to avoid miscalculation. Analytical resolution is dependent on the quantity of

17   regions the image is partitioned into and, for 3D distributions, the Monte Carlo simulation is also

18   dependent on the quantity of uniformly generated random points. The partitioning of image regions

19   is a balance of precision where larger, coarser partitions may average local spatial phenomena into a

20   larger population of cells, while smaller, finer partitions could provide misleading cell densities due

21   to the sharp geometrical discontinuities that exist at image boundaries and the small number of cells

22   that exist within small partitions. The partition balance is observed as the measurement of small cell-

23   proximal 3D regions held larger error (4% standard error at 0.0042 mm^3) when compared to larger,

24   cell-distal 3D regions (0.015% standard error at 2.4 mm^3). The generation of sufficient quantities of

25   uniformly generated random points to estimate complex 3D volumes is dependent on system and

1   user-specific error tolerance; in the presented case an average of 2.4×10^6 points were generated,

2   which yielded a 5% error when compared against known controls, as discussed above.

3   Many studies of cell recognition and cell-cell or cell-tissue co-localization have avoided the

4   need for efficient image analysis algorithms by capturing multiple or larger images at lower

5   magnification, increasing the cost and time of these already demanding techniques. Morales-

6   Navarrete *et al* [36] proposed imaging at a moderate resolution (25x, 1 μm^3 voxel), then re-tiling

7   parts of this low-resolution image with high resolution (63x, 0.03 μm^3 voxel) to capture both

8   vasculature of approximately 100-300 μm diameters for murine liver reconstruction followed by

9   cellular and subcellular reconstruction of hepatocytes. Studies by Khorshed *et al* [19] and

10 Sundaramurthy *et al* [37] enumerated intravital murine hematopoietic cells at moderate resolutions

11 (20-40x, 0.2-7.3 μm^3 voxel) in order to inspect intra- and inter-cellular activity for small populations

12 of approximately 25 cells. The presented study depicts cells at significantly lower resolution (10x, 46

13 μm^3 voxel), as an optimal balance between image volume analyzed at a tissue scale while still

14 capturing many hematopoietic cells (1000's of 500-33,500 μm^3 cells) at high enough resolution to

15 pinpoint cell nuclei and associated fluorescent markers amongst their environment. This lower

16 magnification and resolution has also been employed in semiautomatic cell and fluorescent marker

17 identification in 3D tissue cultures [38], to assess megakaryocyte distance from vascularization in

18 murine bone marrow [39], and used as standard images to compare cell motion tracking algorithms

19 across 14 laboratories [40].

20 The quantitative evaluation of cellular density and heterogeneity is frequently studied and

21 used *in vitro* and *in vivo* 3D applications inspecting roles of cellular and environmental interactions.

22 Many 3D culture systems employ surfaces, scaffolds, carriers, or fibers as an aim to impart spatial

23 heterogeneity to mimic and, upon implantation, integrate with many tissue asymmetries present

24 physiologically [1-4, 25-28]. Pinpointing cell niche placement and cell association as a tool towards

25 developmental understanding, tracking disease progression, and mammalian cell culture bioprocess

26 design has become increasingly studied in *ex vivo* analogues, however, few intravital analyses have

1   been performed until recently, and none in the adult human [14-29, 36-40]. Cell co-localization

2   algorithms frequently inspect small distances from cells or tissues of interest such as the effect of

3   high oxygen tension on cells within 150 µm of a vascular sinus and the migration of cells within 100

4   µm of bone surfaces [16, 19]. The assessment of spatial heterogeneity away from more complex

5   regions of interest, such as uneven bone surfaces or noncircular vessel walls, could utilize these

6   proposed algorithms to capture the full image-field in one of two ways: in an identical method if

7   these phenomena can be described as a function, or through an iterative process once the complex

8   phenomena of interest have been meshed or mapped similar to the detection of the murine marrow

9   bone surface by Khorshed *et al* [19].

10   These algorithms are capable of being applied to any quantitative 3D imaging platform,

11   including magnetic resonance imaging (MRI), computed tomography (CT), and optical imaging (UV,

12   Visible, IR) due to their use of spatially "tagged" contrast agents or probes (e.g. nanomaterials,

13   labelled small and large molecules, fluorescent proteins), where discrete cells, phenomena, or

14   regions of interest can be spatially identified [6-7]. MRI and CT imaging allow for full specimen

15   penetration but lack multiplexed probe detection, have poorer sensitivity, and lower resolution; only

16   recent advancements have allowed imaging at single-cell resolution [6]. Optical imaging allows for

17   the quantitative analysis of a wide variety of cell-environment interactions and can characterize

18   multiple phenotypes, organelle, and matrix proteins simultaneously by implementing light

19   microscopy (LM) with immunohistochemistry (IHC) or immunofluorescence (IF) imaging of probes

20   and chemicals delineating cell or tissue type, status, and their interactions to quantify the effects of

21   the tissue environment on cell differentiation, proliferation, and stimulation spatially at multiple

22   aspect sizes [8-9]. The versatility of optical imaging is frequently implemented to determine

23   quantitative co-localization from the subcellular protein scale up to multicellular and tissue scales

24   [10-28].

25   The complete quantitative utilization of 3D imaged space is required to leverage the high

26   cost of native or engineered tissue histology. The study of spatial heterogeneity in tissue constructs

1   has become increasingly popular both due to advancements in 3D culture systems to mimic

2   physiological structures and functions, as well as advancements in imaging technologies able to

3   better define the performance of natural and engineered cellular environments. The presented

4   image analyses, for the first time, pair Euclidean metrics with Monte Carlo estimations to determine

5   co-localization across the entire imaged data set, assessing 2.2-fold more area per frame in 2D

6   distance metrics to depict cell populations in more distant, peripheral imaged areas from

7   phenomena of interest. In 3D distance metrics, all cell neighborhoods can be analyzed to half the

8   maximum image length, whereas basic algorithms may only assess one, perfectly central cell per

9   image; hence, these tools provide 10^2 to 10^4 more replicate cell measurements per image at a

10   similar, if not greater, scope. The fabrication, culture, and analysis of tissue engineered and clinical

11   samples for image analysis represents a time-, labor-, and cost-intensive process, and efficient post-

12   analysis tools to quantitatively assess the full image field, as presented here, are needed to detail

13   cellular interactions amongst their environment.

1   **Acknowledgements**

6   **Disclosure Statement**

7   No competing financial interests exist.

References

[1] Cha, C., Liechty, W.B., Khademhosseini, A., Peppas, N.A. Designing biomaterials to direct stem cell fate. ACS Nano **6**, 11, 9353, 2012.

[2] Pampaloni, F., Reynaud, E.G., Stelzer, E.H.K., The third dimension bridges the gap between cell culture and live tissue. Nat Rev Mol Cell Biol **8**, 11, 839, 2007.

[3] Baker B.M., Chen, C.S., Deconstructing the third dimension: how 3D culture microenvironments alter cellular cues. J Cell Sci **125**, 13, 3015, 2012.

[4] Khademhosseini, A., Langer, R., Borenstein, J., Vacanti, J.P., Microscale technologies for tissue engineering and biology. Proc Natl Acad Sci U S A **103**, 8, 2480, 2006.

[5] Niklason, L.E., Yeh, A.T., Calle, E.A., Bai, Y., Valentín, A. and Humphrey, J.D., Enabling tools for engineering collagenous tissues integrating bioreactors, intravital imaging, and biomechanical modeling. Proc Natl Acad Sci U S A **107**, 8, 3335, 2010.

[6] Appel, A.A., Anastasio, M.A., Larson, J.C. and Brey, E.M., Imaging challenges in biomaterials and tissue engineering. Biomaterials, **34**, 28, 6615, 2013.

[7] Weissleder, R., Nahrendorf, M., Advancing biomedical Imaging. Proc Natl Acad Sci U S A **112**, 47, 14424, 2015.

[8] Ntziachristos, V., Going deeper than microscopy: the optical imaging frontier in biology. Nat Methods **7**, 8, 603, 2010.

[9] Flower, Maggie A., ed. Webb's physics of medical imaging. Boca Raton: CRC Press, 2012.

[10] Ntziachristos, V., Floresence molecular imaging. Annu Rev Biomed Eng, **8**, 1, 2006.

[11] Bjornsson, C.S., Lin, G., Al-Kofahi, Y., Narayanaswamy, A., Smith, K.L., Shain, W. and Roysam, B., Associative image analysis: a method for automated quantification of 3D multi-parameter images of brain tissue. J Neurosci Methods, **170**, 1, 165, 2008.

[12] Delarue, M., Joanny, J. F., Jülicher, F., & Prost, J., Stress distributions and cell flows in a growing cell aggregate. Interface focus, **4**, 6, 20140033, 2014.

[13] Meddens, M. B., Rieger, B., Figdor, C. G., Cambi, A., & Van Den Dries, K., Automated podosome identification and characterization in fluorescence microscopy images. Microscopy and microanalysis, **19**, 1, 180, 2013.

[14] Acar, M., Kocherlakota, K.S., Murphy, M.M., Peyer, J.G., Oguro, H., Inra, C.N., Jaiyeola, C., Zhao, Z., Luby-Phelps, K. and Morrison, S.J., Deep imaging of bone marrow shows non-dividing stem cells are mainly perisinusoidal. Nature, **526**, 7571, 126, 2015.

[15] Takaku, T., Malide, D., Chen, J., Calado, R.T., Kajigaya, S. and Young, N.S., Hematopoiesis in 3 dimensions: human and murine bone marrow architecture visualized by confocal microscopy. Blood, **116**, 15, e41, 2010.

[16] Nombela-Arrieta, C., Pivarnik, G., Winkel, B., Canty, K.J., Harley, B., Mahoney, J.E., Park, S.Y., Lu, J., Protopopov, A. and Silberstein, L.E., Quantitative imaging of haematopoietic stem and progenitor

cell localization and hypoxic status in the bone marrow microenvironment. Nat Cell Biol, **15**, 5, 533, 2013.

[17] Celso, C.L., Fleming, H.E., Wu, J.W., Zhao, C.X., Miake-Lye, S., Fujisaki, J., Côté, D., Rowe, D.W., Lin, C.P. and Scadden, D.T., Live-animal tracking of individual haematopoietic stem/progenitor cells in their niche. Nature **457**, 7225, 92, 2009.

[18] Spencer, J.A., Ferraro, F., Roussakis, E., Klein, A., Wu, J., Runnels, J.M., Zaher, W., Mortensen, L.J., Alt, C., Turcotte, R. and Yusuf, R., Direct measurement of local oxygen concentration in the bone marrow of live animals. Nature, **508**, 7495, 269, 2014

[19] Khorshed, R.A., Hawkins, E.D., Duarte, D., Scott, M.K., Akinduro, O.A., Rashidi, N.M., Spitaler, M. and Celso, C.L., Automated identification and localization of hematopoietic stem cells in 3D intravital microscopy data. Stem cell reports, **5**, 1, 139, 2015.

[20] Bertrand, J.Y., Kim, A.D., Violette, E.P., Stachura, D.L., Cisson, J.L. and Traver, D., Definitive hematopoiesis initiates through a committed erythromyeloid progenitor in the zebrafish embryo. Development **134**, 23, 4147, 2007.

[21] Vera, B., Zon, L.I., High throughput in vivo phenotyping: the zebrafish as a tool for drug discovery for hematopoietic stem cells and cancer. Drug Discov Today Dis Models **10**, 1, e17, 2013.

[22] Thon, J.N., Mazutis, L., Wu, S., Sylman, J.L., Ehrlicher, A., Machlus, K.R., Feng, Q., Lu, S., Lanza, R., Neeves, K.B. and Weitz, D.A., Platelet bioreactor-on-a-chip. Blood, **124**, 12, 1857, 2014.

[23] Ma, Z., Wang, J., Loskill, P., Huebsch, N., Koo, S., Svedlund, F.L., Marks, N.C., Hua, E.W., Grigoropoulos, C.P., Conklin, B.R. and Healy, K.E. Self-organizing human cardiac microchambers mediated by geometric confinement. Nat Commun, **6,** 7413, 2015.

[24] Toh, Y.C., Xing, J., Yu, H., Modulation of integrin and E-cadherin-mediated adhesions to spatially control heterogeneity in human pluripotent stem cell differentiation. Biomaterials **50**, 87, 2015.

[25] Di Buduo, C.A., Wray, L.S., Tozzi, L., Malara, A., Chen, Y., Ghezzi, C.E., Smoot, D., Sfara, C., Antonelli, A., Spedden, E. and Bruni, G., Programmable 3D silk bone marrow niche for platelet generation ex vivo and modeling of megakaryopoiesis pathologies. Blood, **125**, 14, 2254, 2015.

[26] Frezza C., Wouters, B.G. McGuigan, A.P., A three-dimensional engineered tumour for spatial snapshot analysis of cell metabolism and phenotype in hypoxic gradients. Nat Mater, **15**, 2, 227, 2016.

[27] Eng, G., Lee, B.W., Parsa, H., Chin, C.D., Schneider, J., Linkov, G., Sia, S.K. and Vunjak-Novakovic, G., Assembly of complex cell microenvironments using geometrically docked hydrogel shapes. Proc Natl Acad Sci U S A, **110**, 12, 4551, 2013.

[28] Rnjak-Kovacina, J., Wray, L.S., Golinski, J.M. and Kaplan, D.L., Arrayed Hollow Channels in Silk-Based Scaffolds Provide Functional Outcomes for Engineering Critically Sized Tissue Constructs. Adv Funct Mater, **24**, 15, 2188, 2014.

[29] Liu, Z., Gerner, M.Y., Van Panhuys, N., Levine, A.G., Rudensky, A.Y. and Germain, R.N., Immune homeostasis enforced by co-localized effector and regulatory T cells. Nature, **528**, 7581, 225, 2015

[30] Panoskaltsis, N., Macedo, H.M.M., Blanco, M.T.M., Mantalaris, A., Livingston, A.G., 3d hollow fibre bioreactor systems for the maintenance, expansion, differentiation, and harvesting of human stem cells and their progeny [Patent], WO 2012/069841 A1, 2012.

1 [31] Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S.,
2 Rueden, C., Saalfeld, S., Schmid, B. and Tinevez, J.Y., Fiji: an open-source platform for biological-
3 image analysis. Nat Methods, **9**, 7, 676, 2012.

4 [32] Preibisch, S., Saalfeld, S. and Tomancak, P., Globally optimal stitching of tiled 3D microscopic
5 image acquisitions. Bioinformatics, **25**, 11, 1463, 2009

6 [33] Bolte, S. and Cordelieres, F.P., A guided tour into subcellular colocalization analysis in light
7 microscopy. J Microsc, **224**, 3, 213, 2006

8 [34] Smith, R.L., Efficient Monte Carlo procedures for generating points uniformly distributed over
9 bounded regions. Oper. Res., **32**, 6, 1296, 1984

10 [35] Cowles, M.K. and Carlin, B.P., Markov chain Monte Carlo convergence diagnostics: a
11 comparative review. Amer. Statist. Assoc., **91**, 434, 883, 1994.

12 [36] Morales-Navarrete, H., Segovia-Miranda, F., Klukowski, P., Meyer, K., Nonaka, H., Marsico, G.,
13 Chernykh, M., Kalaidzidis, A., Zerial, M. and Kalaidzidis, Y., A versatile pipeline for the multi-scale
14 digital reconstruction and quantitative analysis of 3D tissue architecture. eLife, **4**, e11214, 2015

15 [37] Sundaramurthy, V., Barsacchi, R., Chernykh, M., Stöter, M., Tomschke, N., Bickle, M., Kalaidzidis,
16 Y. and Zerial, M., Deducing the mechanism of action of compounds identified in phenotypic screens
17 by integrating their multiparametric profiles with a reference genetic screen. Nat Protoc, **9**, 2, 474,
18 2014.

19 [38] De Boodt, S., Poursaberi, A., Schrooten, J., Berckmans, D. and Aerts, J.M., A semiautomatic cell
20 counting tool for quantitative imaging of tissue engineering scaffolds. Tissue Eng Part C Methods, **19**,
21 9, 697, 2013.

22 [39] Malara, A., Currao, M., Gruppi, C., Celesti, G., Viarengo, G., Buracchi, C., Laghi, L., Kaplan, D.L.
23 and Balduini, A., Megakaryocytes Contribute to the Bone Marrow-Matrix Environment by Expressing
24 Fibronectin, Type IV Collagen, and Laminin. Stem Cells, **32**, 4, pp.926-937, 2014.

25 [40] Chenouard, N., Smal, I., De Chaumont, F., Maška, M., Sbalzarini, I.F., Gong, Y., Cardinale, J.,
26 Carthel, C., Coraluppi, S., Winter, M. and Cohen, A.R., Objective comparison of particle tracking
27 methods. Nat Methods, **11**, 3, 281, 2014.

1   **Supplementary Information**

2   **Method: 2D Tissue Distance Explicit Derivation**

3   Cell distance from hollow fiber centers, $(X, Y)_a$, was defined as 2D Euclidean distance:

4    $$R_c = \sqrt{(X_c - X_a)^2 + (Y_c - Y_a)^2}$$ [1]

5   The full image field distance was defined as the distance from the fiber to the furthest pixel of the

6   image, R_{max}. The full distance was then discretized into bins of equally spaced radii, R_i, which form

7   concentric circles of area $\pi(R_i^2 - R_{i-1}^2)$. However, if the circle of radius R_i centered at $(X, Y)_a$

8   encompassed space outside of the cubic image limits, the extraneous volume, A_e, must be

9   subtracted off:

10   $$A_e = \frac{R_i^2}{2}\left(\frac{\pi S_i}{180} - \sin(S_i)\right)$$ [S.1]

11   The arc angle of the circle area segment protruding out of the image, S_i, is calculated through

12   trigonometric identities:

13   $$S_i = \cos^{-1}\left(\frac{D_{a \to k1}^2 + D_{a \to k2}^2 - (Y_{k1} - Y_{k2})}{2 D_{a \to k1}^2 D_{a \to k2}^2}\right)$$ [S.2]

14   Where $D_{a \to k1}$ represents the 2D Euclidean distance between point $(X, Y)_a$ and $(X, Y)_{k1}$, see **Figure**

15   **3A** for terminology.

16   In cases where the extraneous circle area included a vertex of the rectangular image limits,

17   two overlapping extra-image volumes have been subtracted from the circular area which must be

18   corrected by adding back the overlapping extraneous area A_{ee}:

$$A_{ee} = S_i - \sqrt{P_A(P_A - D_{a \to k2})(P_A - D_{a \to i})(P_A - D_{i \to k2})} -$$

19   $$\sqrt{P_B(P_B - D_{a \to j2})(P_B - D_{a \to i})(P_B - D_{i \to j2})}$$ [S.3]

20   Where S_i, from **Equation S.2**, here describes the arc angle of the circle slice inscribing A_{ee}, while P_A

21   and P_B are triangle perimeters which exclude the overlapped circle section A_{ee}, described as:

22   $$P_A = 0.5 D_{a \to k2} + D_{a \to i} + D_{i \to k2}$$ [S.4]

23 $$P_B = 0.5 D_{a \to j2} + D_{a \to i} + D_{i \to j2}$$ [S.5]

24   So that the final volume of the concentric circles constrained to rectangular image limits is:

1
$$A_i = \pi\left(R_i^2 - R_{i-1}^2\right) - A_e + A_{ee} \tag{2}$$

2   With all terms corresponding to the diagram in Figure 3a.

3   **Results: Comparison of 3D Distance Analysis against Manual Scoring**

4   The calculation of 3D cell density distribution using the proposed algorithm greatly saved

5   time in comparison to a manual scoring method. For instance, the algorithm was able to calculate

6   the distribution of cells in 4 concentric regional neighborhoods (0-15 μm, 15-50 μm, 50-100 μm,

7   >100 μm) around each of the 1116 cells in Figure 1B within 62 seconds. This analysis was compared

8   against the manual scoring of 3D cell density within random fractions of the image, another method

9   to hasten analysis. To do this, Figure 1B was divided into 100 equal parts, and 5 image parts of this

10   100 were randomly selected (using MATLAB's random number generator). The image parts were

11   then manually scored for cell position, then cell to cell distances were measured and sorted into 2
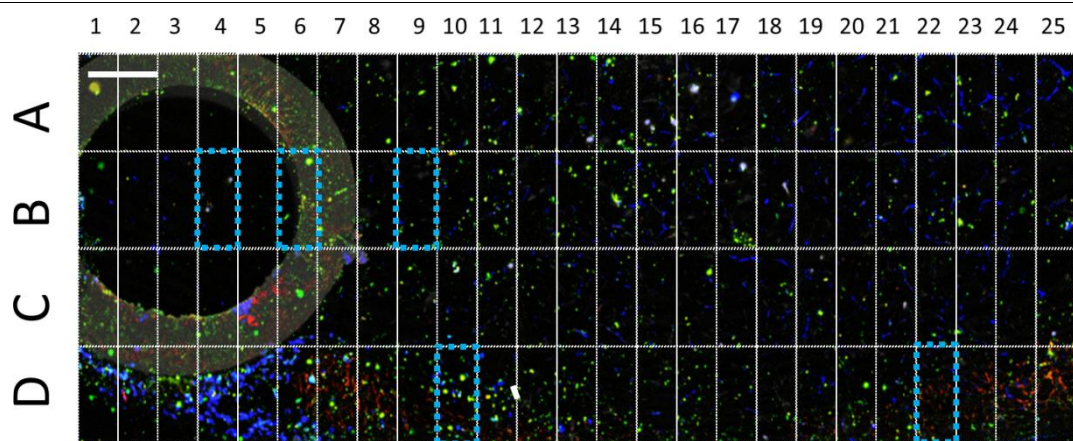
12

13   **Figure S1:** Comparison of algorithm efficiency in comparison with manual scoring. Figure 1B was

14   partitioned into 100 equal regions, and 5 parts were selected by random number generation

15   (MATLAB), highlighted in blue as partitions B4, B6, B9, D10 and D22 which represented 5% of the full

16   image area. Cells were manually identified within each partition, 3D cell to cell distances were

17   identified if within 15-50 or 50-100 μm, and the average 3D volume of concentric spheres 15-50 or

18   50-100 μm away from cells of interest were manually estimated to calculate cell densities, and

19   reported in Table S1.

1 **Table S1**: Comparison of calculating 3D cell density distributions by manually scoring 5 random
2 image parts against utilising the proposed algorithm across the entire image.

	Manual Scoring of 5x 1% image parts	Automated estimation of the entire image
Time of Analysis	30 minutes	1 minute
Cells Utilized	67	1116
15-50 µm neighborhood density (cells/mm^3)	25600 +/- 11300	3400 +/- 131
50-100 µm neighborhood density (cells/mm^3)	2220 +/- 832	1776 +/- 46

3

4 partitions (15-50 µm, 50-100 µm), and the average volume of these partitions within the partial, 1%,

5 image boundaries was approximated (437,500 µm^3, 2,700,000 µm^3) in order to calculate density.

6 Surveying 5x 1% image parts required 30 minutes for manual scoring (N=1), or 10.2 seconds for the

7 algorithm (N=20), illustrated in Figure S1.

8 Surveying 5 image parts comprising 5% of the entire image was an imprecise method to

9 calculate 3D cell density distribution. Manual scoring provided inconsistent results, with standard

10 errors of 44% and 38% about a mean which was 750% and 25% different than when using the

11 proposed algorithm across the complete image (Table S1). These differences were likely due to both

12 a large variability between image parts, and the lower number of cells utilized. When manually

13 scoring, there existed between 2 and 28 cells per image (average of 13.4), and when applying the

14 algorithm to 20 replicates of this 5 image parts method, there existed between 0 and 24 cells per

15 image (average of 13.3 +/- 3.2), demonstrating the 5 parts only had 5.9% of the total 1116 cells in

16 the complete image. The lack of cells and inconsistency of regions surveyed increased standard error

17 significantly, from 2.9% to 29% and from 1.3% to 20.3% for the 15-50 µm and 50-100 µm regions

18 when limiting the algorithm to 5x 1% image parts.

19 Analysing segmented image fractions reduces time while greatly reducing precision and also

20 limiting features of the presented algorithm. This algorithm is unique in its ability to determine cell

21 density at the furthest distance possible, and when Figure 1B is segmented into 100 parts of 388 x

22 177 x 350 µm^3, the furthest distance between two points that the algorithm could analyze cell

1   density would be 550 μm, instead of 4700 μm when implementing the full image. Furthermore,

2   implementing a manual scoring method to measure complex image partition volumes is tedious and

3   difficult to approximate. Therefore, the proposed algorithm reduces calculation time while

4   maintaining a higher precision and functionality in comparison to methods of random image scoring.

5   **Discussion: Grid Sampling Comparison to Random Sampling**

6   Volume estimation by sampling random points from a uniform distribution within a known

7   area converged much more quickly than dividing the image space or subset space into a grid of

8   evenly-spaced points. Uniform random point generation required 1.3 ± 0.2 milliseconds per cell to

9   reach 1% estimated volume convergence (n = 1,000 cells) while grid point generation required $7.9 \pm$

10   10.8 minutes per cell to reach the same convergence (n = 10 cells). The large time and variability for

11   grid point generation may be attributed to the $(2k)^3$ additional points that must be generated for

12   each subsequent iteration of k points so that an evenly-spaced grid would be generated at identical

13   limits for unbiased point generation. A visual comparison of volume estimations by their convex hulls

14   can be found in Figure S2.
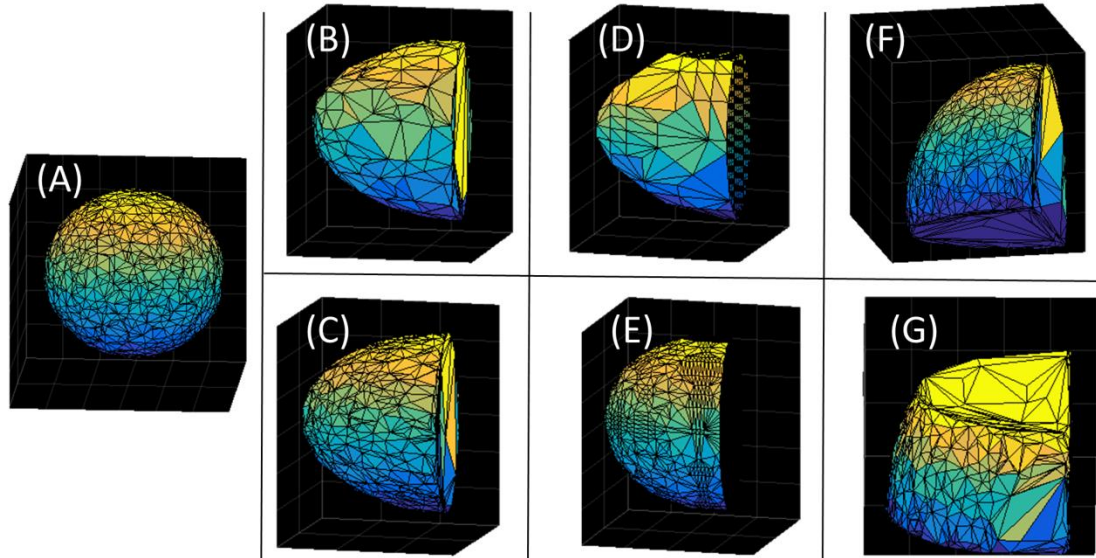
1

2   **Figure S2**: Convex hulls illustrating the cell clustering 50 μm neighborhood estimation for (A) a full

3   sphere by 17 generations of 10,000 uniform random points, (B) a sphere's intersection with the

4   image limits with 10,000 uniform and (C) 100,000 uniform points or (D) 10,000 evenly-spaced points

5   and (E) 100,000 evenly-spaced points. Larger, 600 μm cell neighborhoods tend to intersect 2 or 3

6   planes of image limits, defined by (F) 17 and (G) 13 generations of 10,000 uniform points.