Abstract

Image analysis without segmentation: A new method to measure cytoplasm to nucleus translocation

Ilya Ravkin and Vladimir Temov, Vitra Bioscience, Inc.

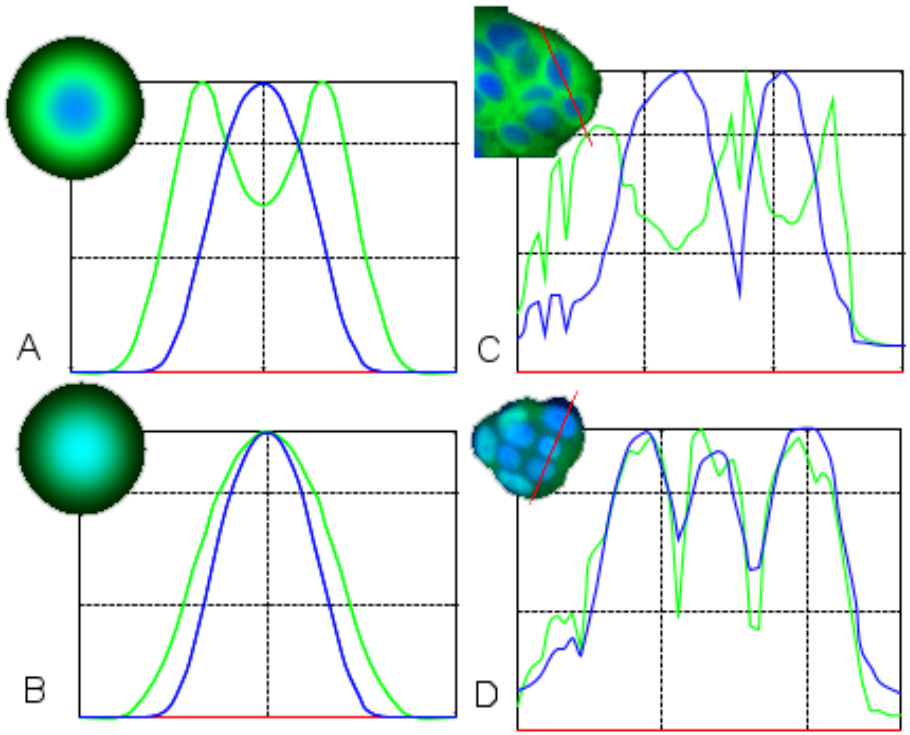
A common approach to measuring many cellular processes by image analysis is to start with segmenting the image into compartments of interest. Vitra Bioscience has developed a method of analysis of cytoplasm to nucleus translocation (CNT) that does not require subcellular segmentation. It can be performed individually on a cell-by-cell basis or globally on the whole image. The method is based on modeling and analysis of the 2D distribution of stains: signal (i.e. protein stain) and nuclear counterstain. The method does not have any user parameters and is very tolerant to variation in image acquisition. We analyzed the performance of this new method with respect to cell type, magnification, number of analyzed cells, accuracy of focusing, depth of field, and plate flatness. Using statistical methods to measure the quality of data produced by this new algorithm, we confirmed that it performs better than our implementation of algorithms based on segmentation. The performance assessment methodology will be discussed in the context of Vitra's application of this algorithm to measure translocation events at varying magnifications on the CellCard™ System and in the future on CellPlex™ assays. Furthermore, the presented performance assessment methodology can be used in the future to design the best screening strategy and to compare different cell analysis algorithms.

VITRA BIOSCIENCE

Images and profiles through model and real cells

To find a robust measure of nuclear translocation we have defined a model of spatial distribution of the nuclear counterstain and of the signal stain as it moves from the cytoplasm to the nucleus. The model was studied under some perturbations in order to find measures that are robust.

The model of cell staining comprises a bell-shaped intensity distribution of counterstain, which is shown in blue, and a bell-shaped distribution of signal stain, which is shown in green. For the negative case the distribution of signal stain is wider and has a bell-shaped crater. Profiles through the real cells show substantial similarity to the model profiles. (Profile C is plotted through two cells, profile D – through three cells. All profiles are independently normalized to their intensity maxima.).



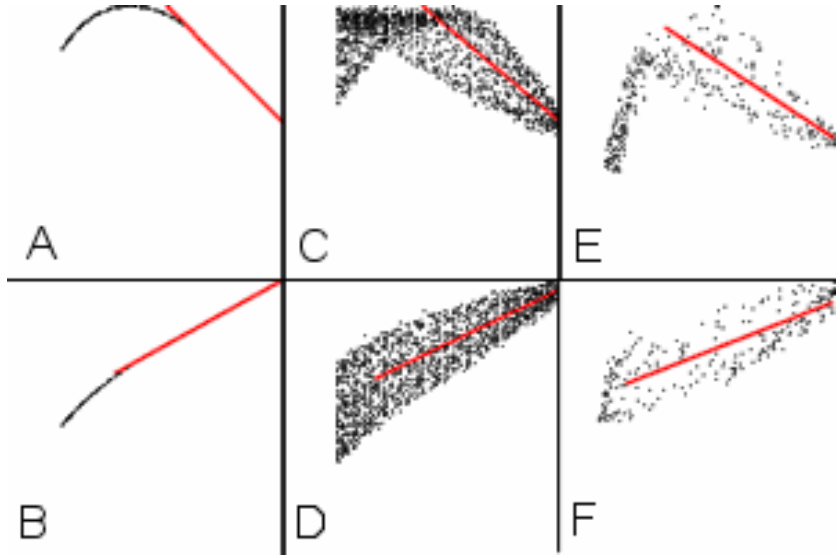
A,B – profiles through model distributions; C,D – profiles through real cells. A,C – negative, B,D – positive. Blue – counterstain, green – signal stain.

VITRA BIOSCIENCE

Model and experimental 2D distributions of signal stain and counterstain

To derive stable measures that characterize transitions from the negative to the positive case, we analyzed joint distributions of the stains on the model and on real cells. In the ideal case, the model spatial stain distributions are circularly symmetrical and aligned. The cross-histogram for this case is shown in panels A and B. If the model is perturbed by offsetting the centers of the two stains, by changing shape from circular to oval, or by adding noise, the distributions become fuzzy as shown in panels C and D. Typical negative and positive real cells have cross-histogram as shown in panels E and F.

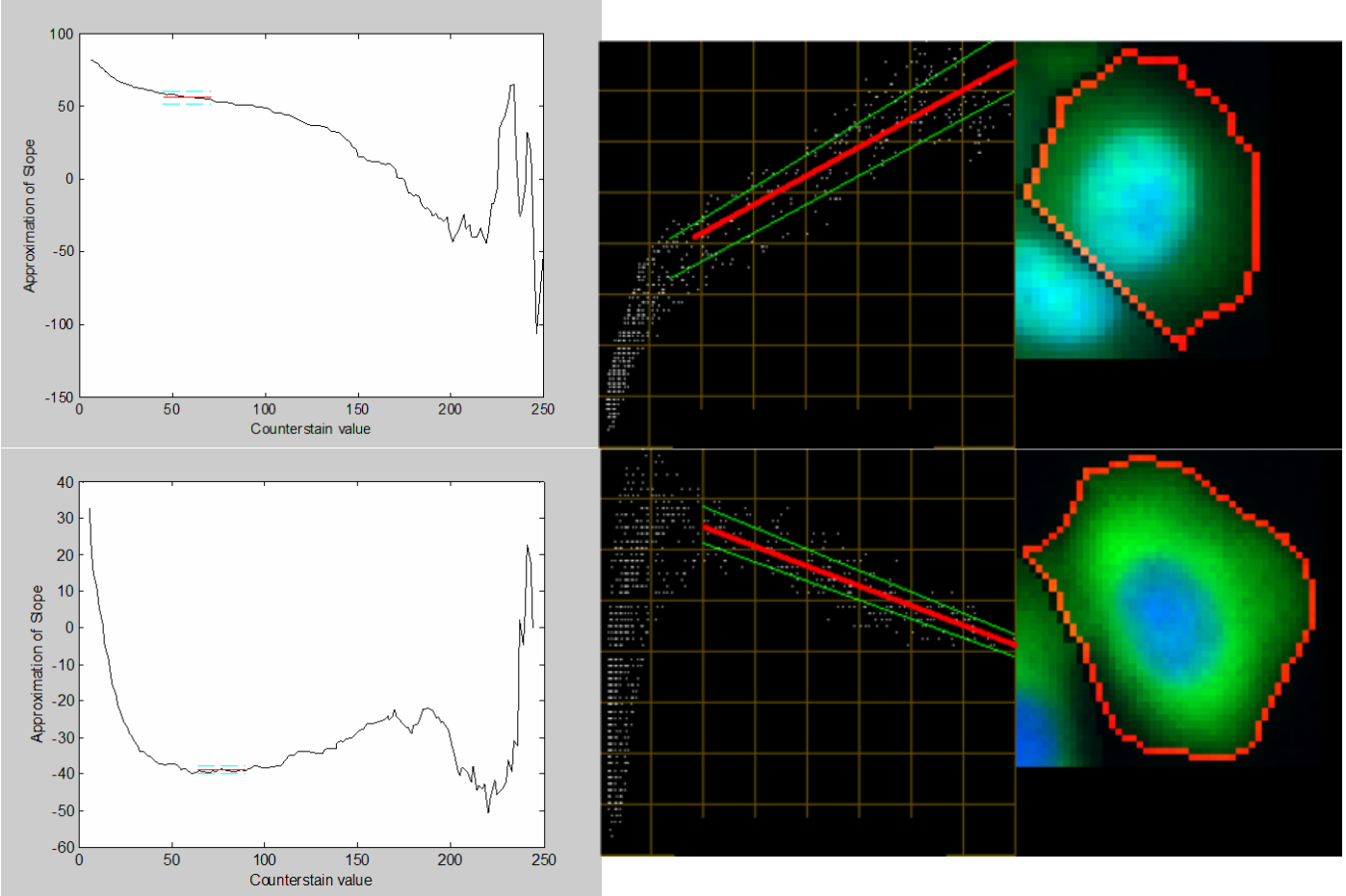
These distributions suggest that a translocation measure can be defined as the slope of a straight-line segment approximating the right side of the cross-histogram. This portion of the distribution corresponds to the more intense nuclear staining and is also close to the center of the nucleus. The farther from the center, the more diffuse the distribution, and the less reliable the approximation becomes.



Cross-histograms of counterstain (X-axis) and signal stain (Y-axis) in ideal model (A,B), perturbed model (C,D) and a real cell (E,F). Scale on both axes is 0-255. Red lines show the calculated approximation.

VITRA BIOSCIENCE

Parameters derived from 2D distributions

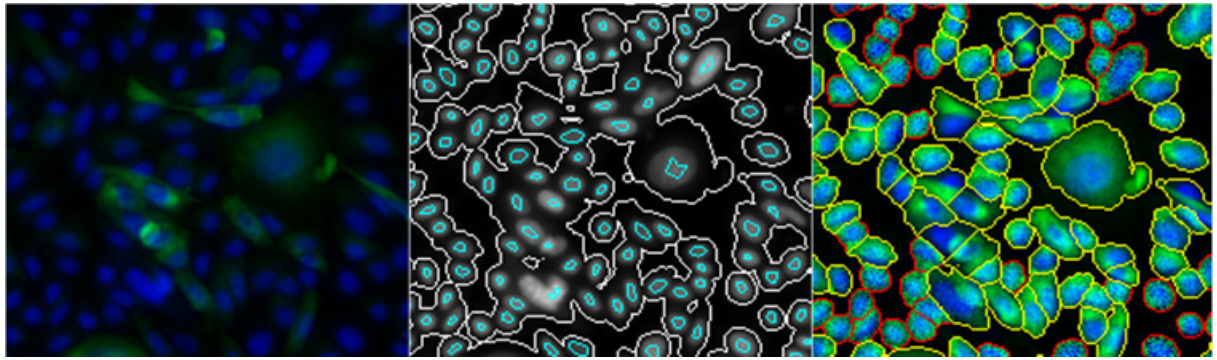


The portion of the distribution that is used for approximation with the straight line is found by plotting the approximated slope going from right to left and selecting the range where this approximation is the most stable. Color lines show the resulting value of slope. Top panel – nuclear localization (positive), bottom panel – cytoplasm localization (negative) of protein.

A variation on this method is to calculate two more slopes as shown in panels on the right. The top line is the regression line calculated on all points above the original slope segment (which we will refer to as Slope1); the bottom line is the regression line calculated on all points below the original slope segment. If all three slopes have the same sign, the result is the one with the greatest absolute value. If they have different signs, then the original slope is chosen. We call this measure Slope3.

VITRA BIOSCIENCE

Partitioning into components - Markers



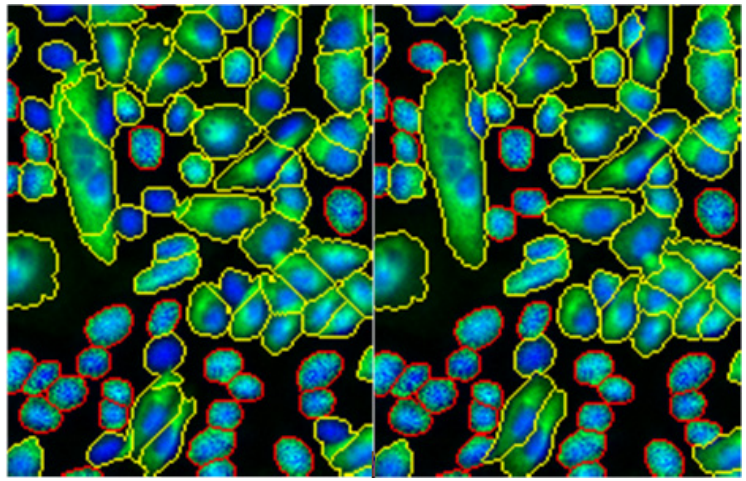
Left - original image of cells expressing GFP-fused p65 protein. Middle - smoothed nuclear counterstain image shown in grayscale with overlaid contours of markers (cyan) and contours of background areas (white). Right - partitioning of the image induced by the markers.

Partitioning into components serves a dual purpose. The first is for individual cell analysis; the second – as a step in optional intensity equalization. There are two main steps in the partitioning method: finding of markers and finding of separation lines.

Markers are found by the following algorithm. A fixed value (marker contrast) is subtracted with saturation from the image of nuclear counterstain and the resulting image is reconstructed[7] within the image of nuclear counterstain. This image is then subtracted from the counterstain image and the result is converted to binary image. The components of this image are the markers. A further restriction may be imposed on markers – only markers that have at least one pixel above a given threshold (marker brightness) are retained for the second step. Depending on magnification and noise level the image of nuclear counterstain may be smoothed prior to this algorithm. This method of determining markers can handle cells of different size and shape. Other methods, e.g. based on top-hat transform[7] may be used too.

Partitioning into components - watersheds of combined intensity images

Separation lines are defined as the watershed[4,5,6] of the inverted image of the linear combination of the counterstain image and the signal stain image. The reason to use linear combination rather than just the nuclear counterstain image is that cells are often nonsymmetrical and unevenly spaced. Separation lines from nuclear stain image may cut through the middle of cells. The use of signal stain produces more accurate separation lines. Coefficients of the linear combination may be varied depending on the peculiarities of staining and of image acquisition.



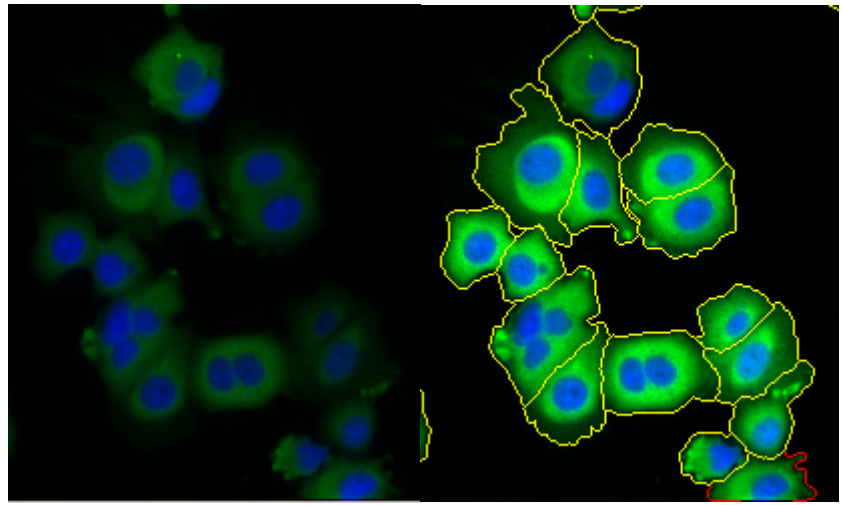
Left - image partitioning using only nuclear stain image (blue). Right – image partitioning using linear combination of nuclear and signal stains. Coefficient of linear combination for the nuclear image = 0.3, for the signal stain = 0.7

Normalization of intensities

The joint distributions of counterstain and signal stain are normalized to their respective maxima. This can be done for the distribution or for the image. The result is the same but normalizing the image gives additional feedback to the user and may reveal features that were not seen before normalization.

Normalization can be done in components described above. In this case all pixels from a component are multiplied by the same number, separately for signal stain and for counterstain.

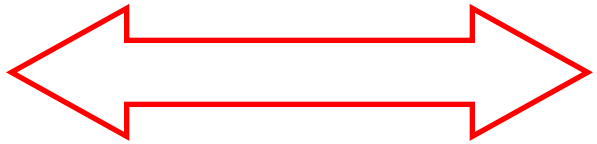
Alternatively, normalization can be done without partitioning the image by fitting a smooth surface to the images of signal stain and counterstain.



Left – original cells, right - cells separated by watershed lines with intensities of the counterstain (blue) and signal stain (green) normalized to the maximum in each compartment.

Quality assessment for cellular imaging assays now

Compare to visual assessment by a human



Use existing HTS quality measures

- *Very laborious*
- *Hard to quantify*
- *Subjective*

May not capture specific effects introduced by image analysis

Desired algorithm:

Sensitive to the variable of interest (e.g., concentration), but insensitive to all other variables (e.g., artifacts)

VITRA BIOSCIENCE

Methodology of the study

1. Establish a measure of quality

2. Identify specific sources of variability:

- Magnification,
- Image size (number of cells)
- Focusing
- Depth of field (e.g., confocal vs. widefield)
- Substrate (plate) properties
- Data extraction algorithm
- Parameters of the algorithm

3. Analyze how the measure of quality depends on these variables

V-factor – a generalization of Z-factor

(1) $$Z = 1 - 3\left(\frac{SD_{pos} + SD_{neg}}{\left|M_{pos} - M_{neg}\right|}\right)$$

-inf 0 0.5 1

If the values of the assay for its positive and negative states do not overlap (and if they do it is not a very useful assay), the z-factor can be manipulated intentionally, by applying a mathematical transformation that maps all positive values into a single value and all negative values into another single value.

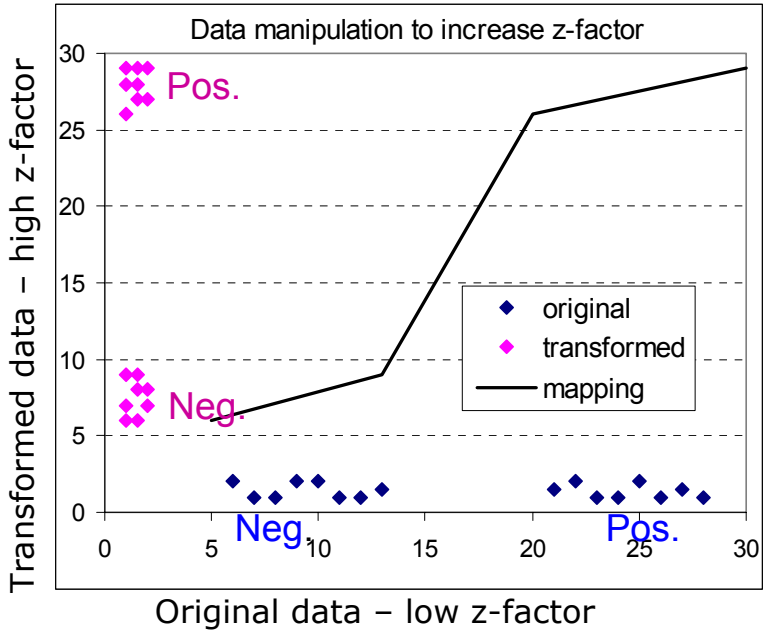
(2) $$V = 1 - 6\left(\frac{SD_{of\_fit}}{\left|M_{pos} - M_{neg}\right|}\right)$$

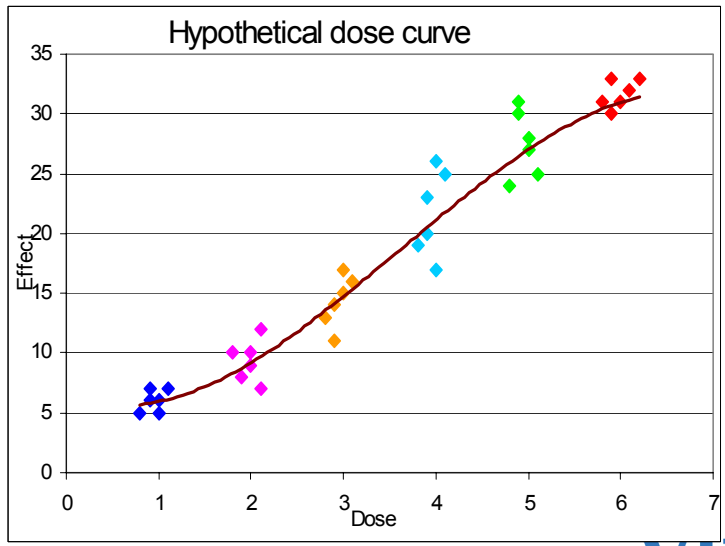
-inf 0 0.5 1

(3) $$SD_{of\_fit} = \sqrt{\frac{1}{n}\sum_{1}^{n}\left(f_{\text{model}} - f_{\text{experiment}}\right)^2}$$

(4) $$V = 1 - 6\left(\frac{Average\_SD}{\left|M_{pos} - M_{neg}\right|}\right)$$

(Alternative definition without a model)





VITRA BIOSCIENCE

Plate Layout

	1	2	3	4	5	6	7	8	9	10	11	12
A	-7	-7.523	-8	-8.523	-9	9.523	-10	-10.523	-11	-11.523	-12	-13
B	-7	-7.523	-8	-8.523	-9	9.523	-10	-10.523	-11	-11.523	-12	-13
C	-7	-7.523	-8	-8.523	-9	9.523	-10	-10.523	-11	-11.523	-12	-13
D	-7	-7.523	-8	-8.523	-9	9.523	-10	-10.523	-11	-11.523	-12	-13
E	-7	-7.523	-8	-8.523	-9	9.523	-10	-10.523	-11	-11.523	-12	-13
F	-7	-7.523	-8	-8.523	-9	9.523	-10	-10.523	-11	-11.523	-12	-13
G	-7	-7.523	-8	-8.523	-9	9.523	-10	-10.523	-11	-11.523	-12	-13
H	-7	-7.523	-8	-8.523	-9	9.523	-10	-10.523	-11	-11.523	-12	-13

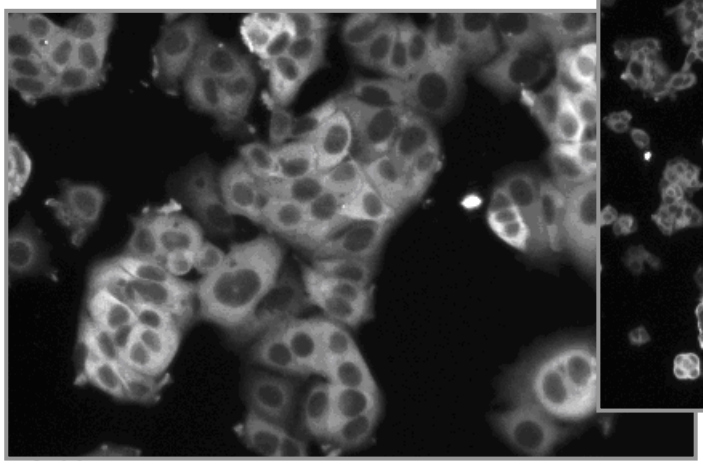
	MCF7		A549	TNF-a concentration in log g/mL

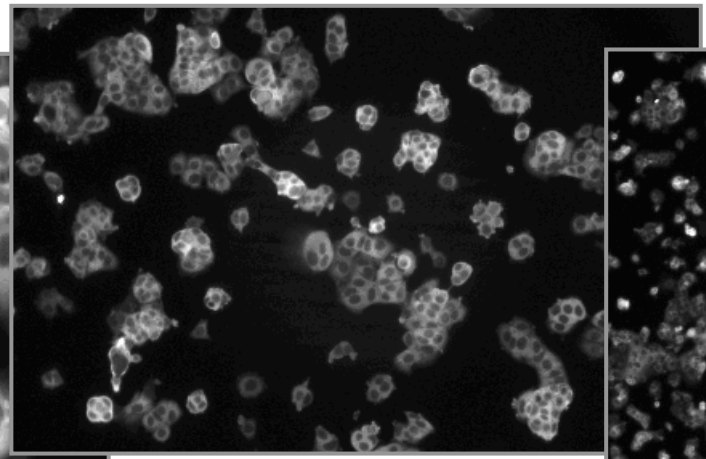
Cells were plated at about 10,000 cells per well in a 96 well microtiter plate (Packard ViewPlate) and incubated overnight. The cells were then treated with varying doses of TNFa, up to 100nm, for 30 minutes. This treatment results in the activation of NFKB and the translocation of the p65 subunit from the cytoplasm to the nucleus. The cells were subsequently fixed and immuno-stained for p65 and counterstained with the Hoechst nuclear dye.

The plate was scanned three times at 10X, 4X, and 2X magnifications acquiring one image per well and analyzed with the Vitra Bioscience algorithm. The data was then plotted and analyzed for quality using the V^2 and Z^1 factor calculations.

The data set is available for viewing and download at ftp://vitra.vitrabio.com/download/Images/NFkB-MCF7&A549.

10X





4X

2X

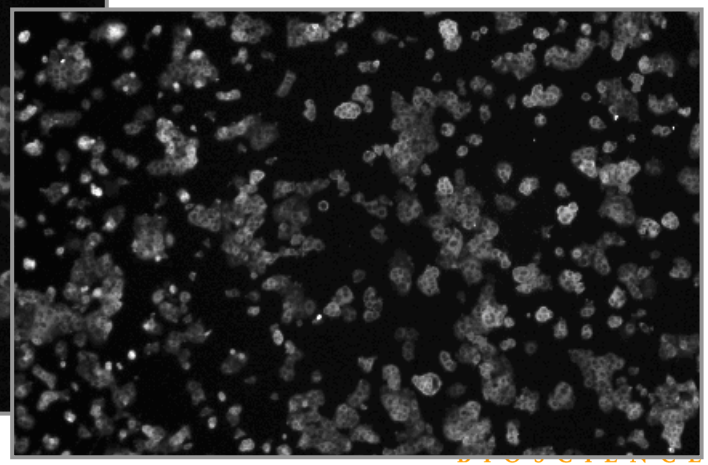
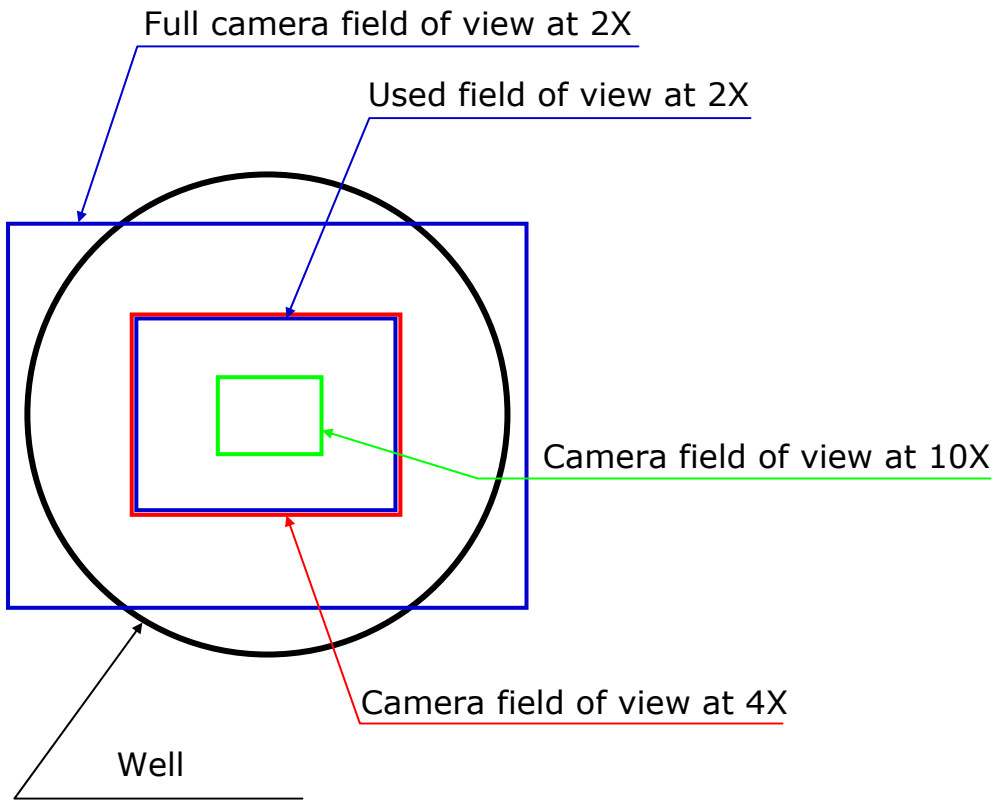
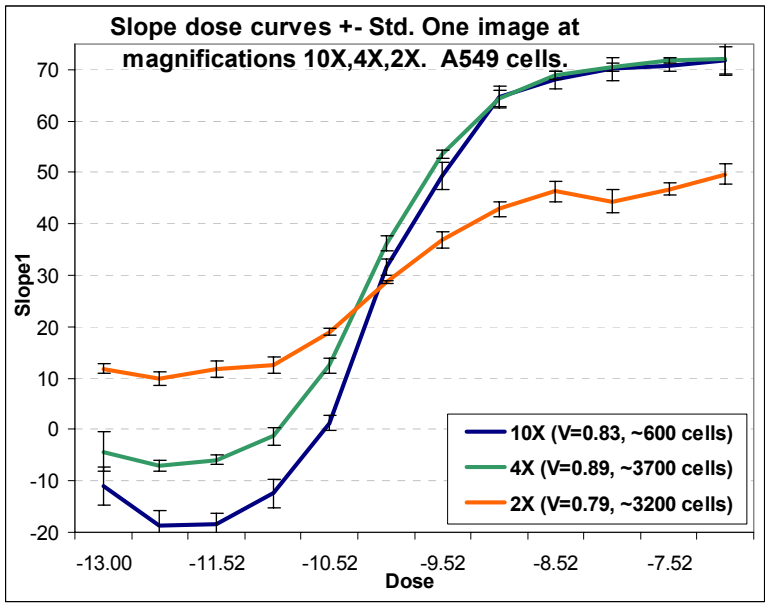


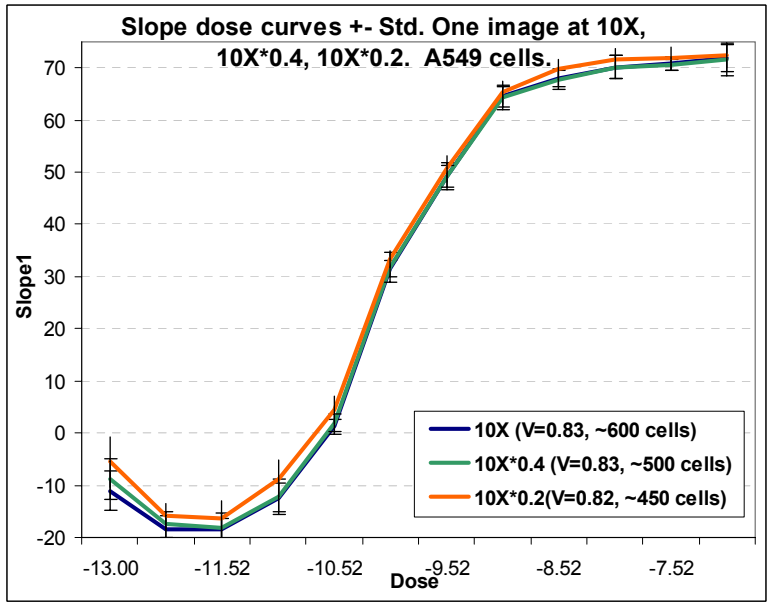
Image acquisition - fields of view at different magnifications

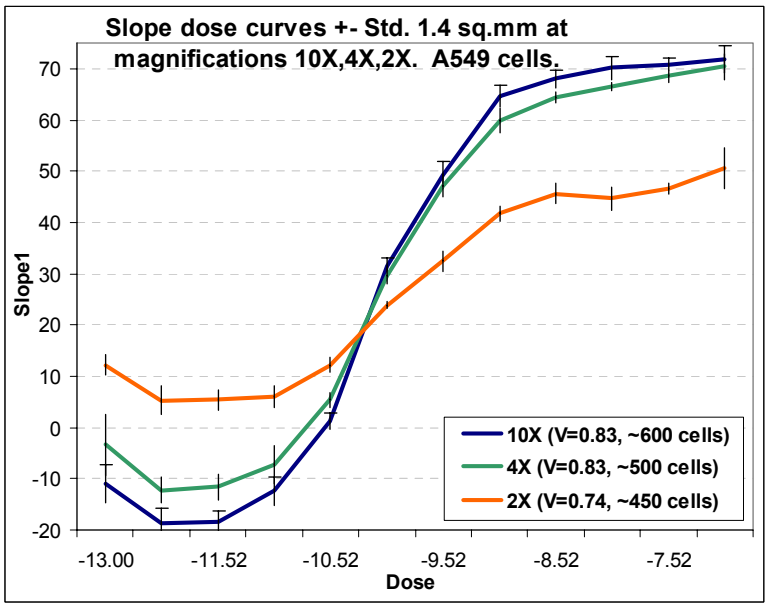


Note: At 2X we use ¼ of the image area to avoid cell plating artifacts at the edges of wells

Dynamic range and V-factor. A549 cells. Dependency on magnification, image size and subsampling.





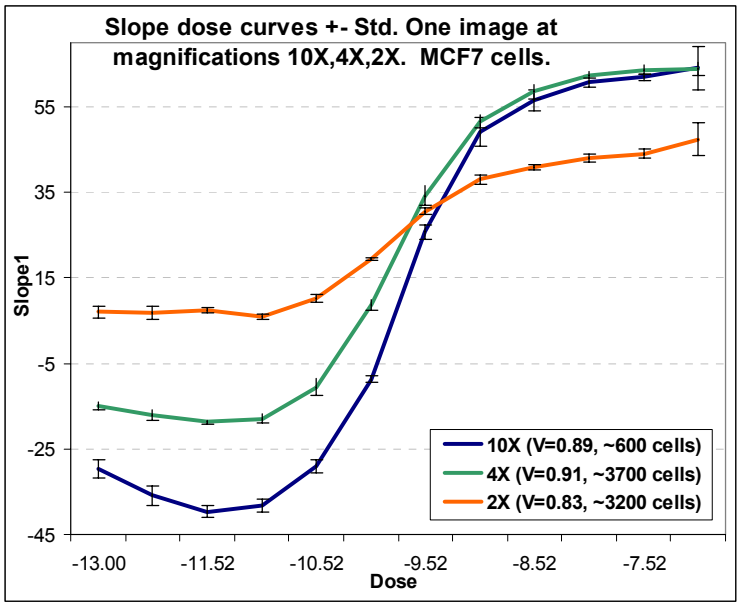


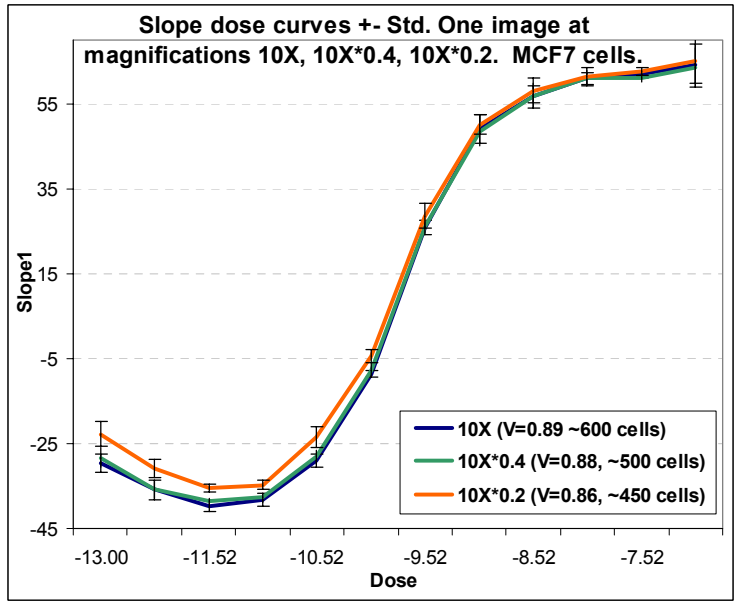
At lower magnifications the dynamic range is reduced, but the variability is also reduced, so the quality (V-factor) remains about the same. This may be due to the greater number of cells in one image at lower magnifications.

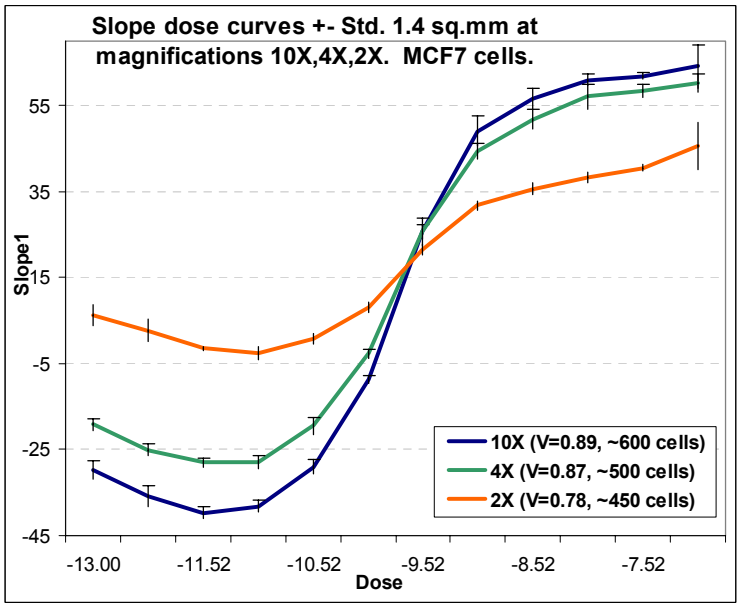
Subsampling does not cause reduction of dynamic range.

VITRA BIOSCIENCE

Dynamic range and V-factor. MCF7 cells. Dependency on magnification, image size and subsampling.







VITRA BIOSCIENCE

Comparison of Slope and Mask (N2R) algorithms by V-factor. Parameter dependency of the Mask algorithm.

A549

Slope	0.84	Slope global	0.83

N2R

Nuclear mask erosion=0

Gap \ Ring	1	2	3	4	5
0	0.77	0.81	0.81	0.80	0.80
1	0.81	0.80	0.80	0.80	0.80
2	0.79	0.79	0.79	0.79	0.80
3	0.78	0.78	0.79	0.80	0.81
4	0.79	0.79	0.80	0.81	0.82

Nuclear mask erosion=1

Gap \ Ring	1	2	3	4	5
0	0.67	0.73	0.79	0.80	0.80
1	0.78	0.80	0.80	0.80	0.80
2	0.80	0.80	0.79	0.79	0.80
3	0.79	0.78	0.79	0.79	0.80
4	0.78	0.78	0.79	0.80	0.81

Nuclear mask erosion=2

Gap \ Ring	1	2	3	4	5
0	0.47	0.57	0.68	0.75	0.79
1	0.64	0.70	0.77	0.79	0.79
2	0.73	0.77	0.79	0.79	0.79
3	0.79	0.79	0.78	0.78	0.79
4	0.78	0.78	0.78	0.78	0.79

MCF7

Slope	0.89	Slope global	0.88

N2R

Nuclear mask erosion=0

Gap \ Ring	1	2	3	4	5
0	0.85	0.86	0.85	0.84	0.84
1	0.85	0.84	0.84	0.83	0.83
2	0.83	0.83	0.83	0.82	0.82
3	0.82	0.82	0.82	0.82	0.82
4	0.82	0.82	0.82	0.81	0.81

Nuclear mask erosion=1

Gap \ Ring	1	2	3	4	5
0	0.81	0.84	0.87	0.86	0.85
1	0.86	0.87	0.86	0.85	0.84
2	0.86	0.85	0.84	0.84	0.84
3	0.83	0.83	0.83	0.83	0.83
4	0.83	0.82	0.82	0.82	0.82

Nuclear mask erosion=2

Gap \ Ring	1	2	3	4	5
0	0.66	0.71	0.79	0.83	0.86
1	0.75	0.8	0.85	0.86	0.86
2	0.82	0.85	0.86	0.86	0.85
3	0.86	0.86	0.85	0.85	0.85
4	0.85	0.84	0.84	0.84	0.84

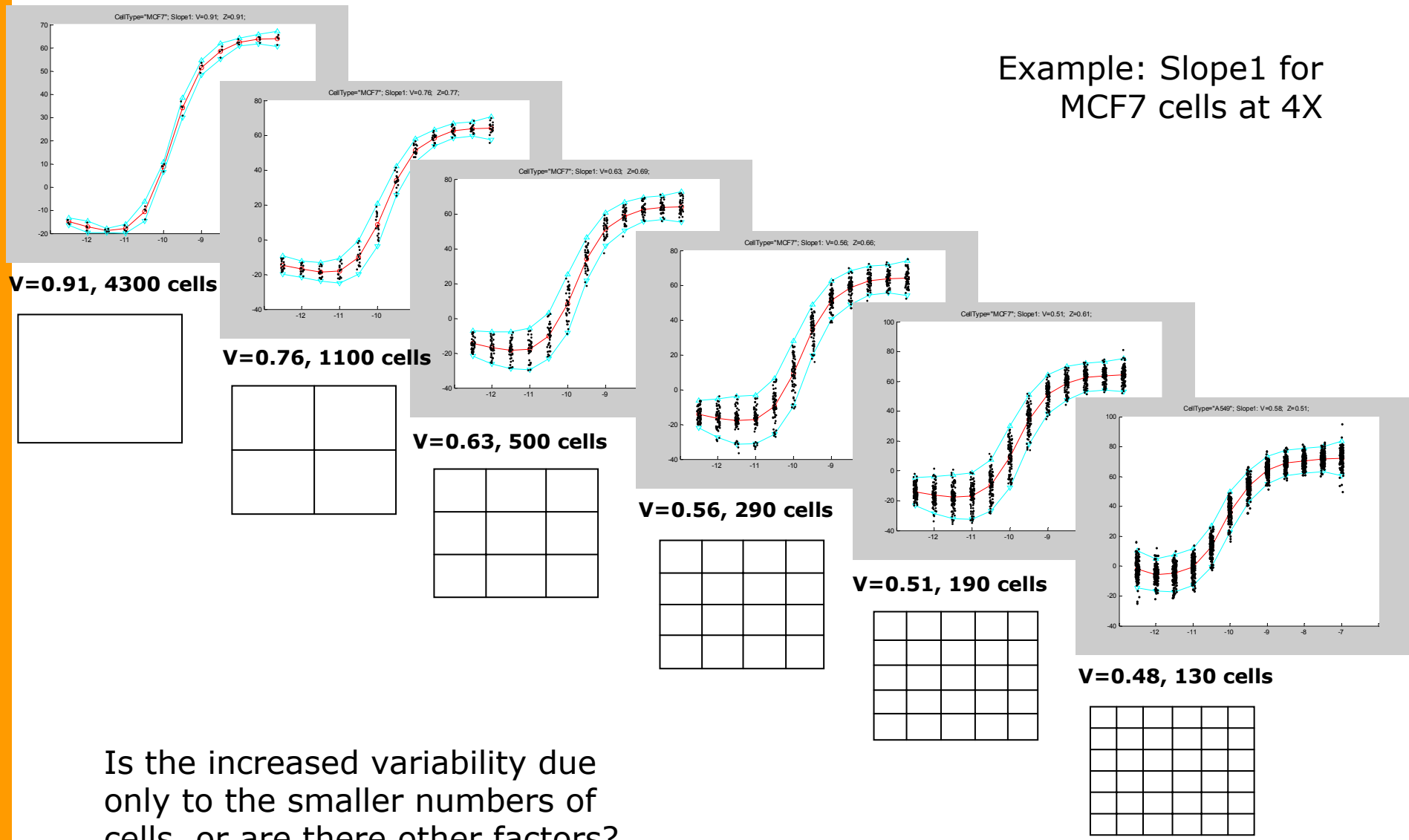
V-factor color coding: [0.86-0.90] [0.81-0.85] [0.76-0.80] [0.60-0.75] [0.40-0.59]

Images at 10X mag.

- The mask-based algorithm shows significant dependency of quality on parameters. The slope algorithm has no parameters.
- The slope algorithm gives better quality than mask-based algorithm at any parameter setting.
- The slope algorithm applied globally is almost as good as when applied on a cell-by-cell basis and is faster to compute.

VITRA BIOSCIENCE

Dependency of quality on the number of cells. Take 1.



Example: Slope1 for MCF7 cells at 4X

Is the increased variability due only to the smaller numbers of cells, or are there other factors?

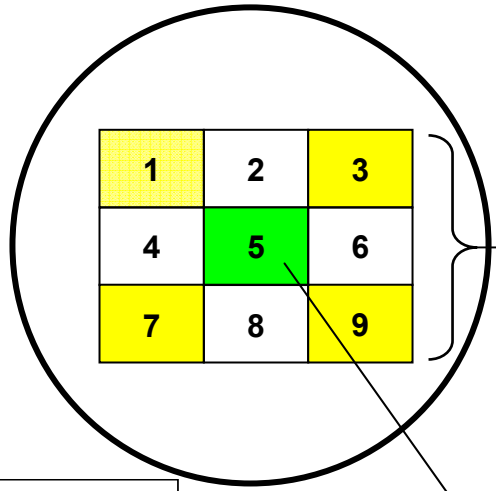
VITRA BIOSCIENCE

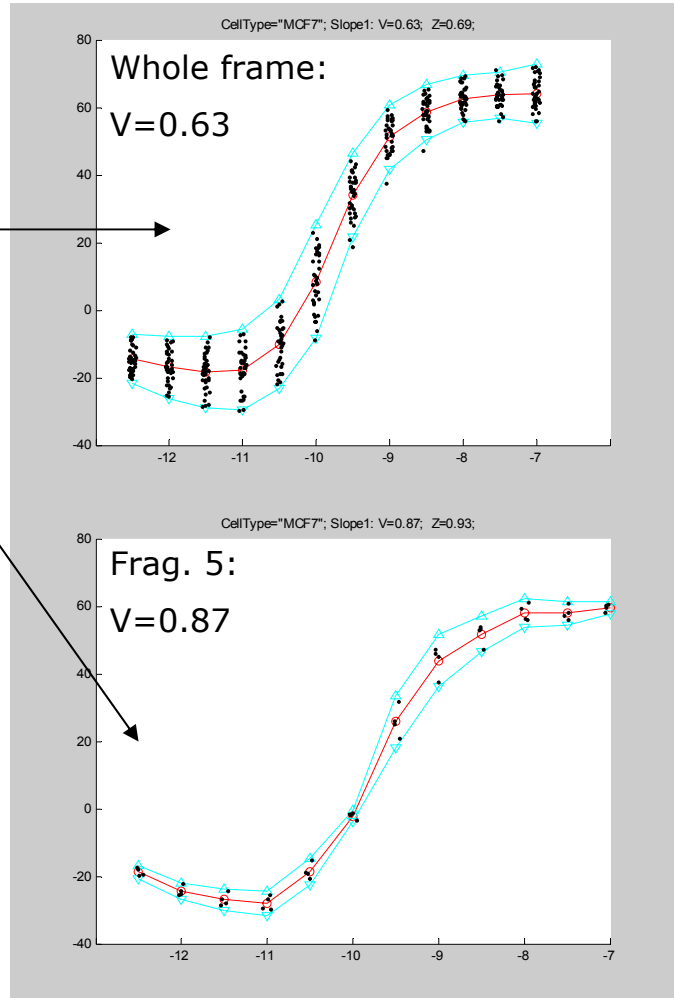
Dependency of CNT measures on well flatness

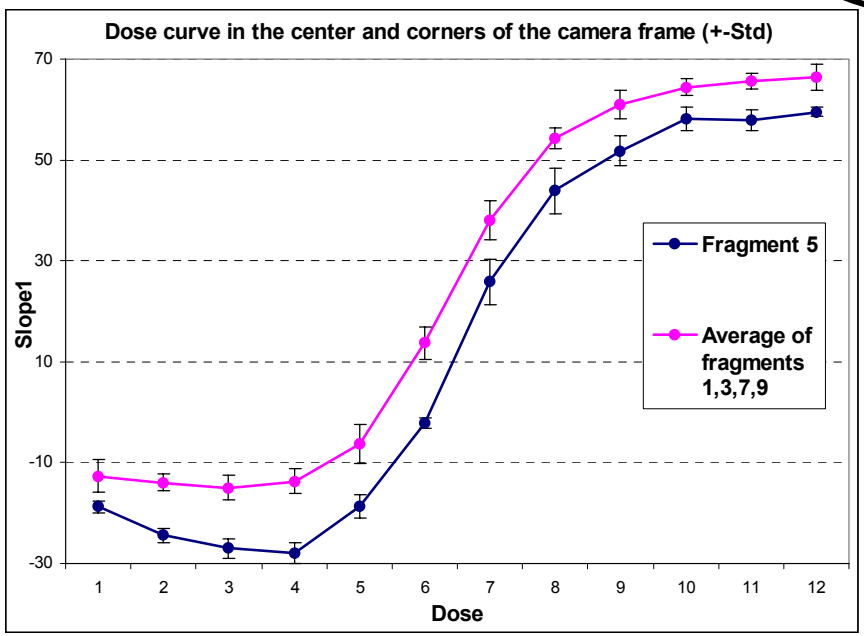
Example: Slope1 for MCF7 cells at 4X. 3*3 fragments.

Source	DF	Sum of squares	Mean square	Fisher's F	Pr > F
Image fragment	8	6202	775	58	<0.0001
Replica (row)	3	342	114	9	<0.0001
Dose (column)	11	531539	48322	3617	<0.0001

ANOVA shows that image fragment is a significant source of variation



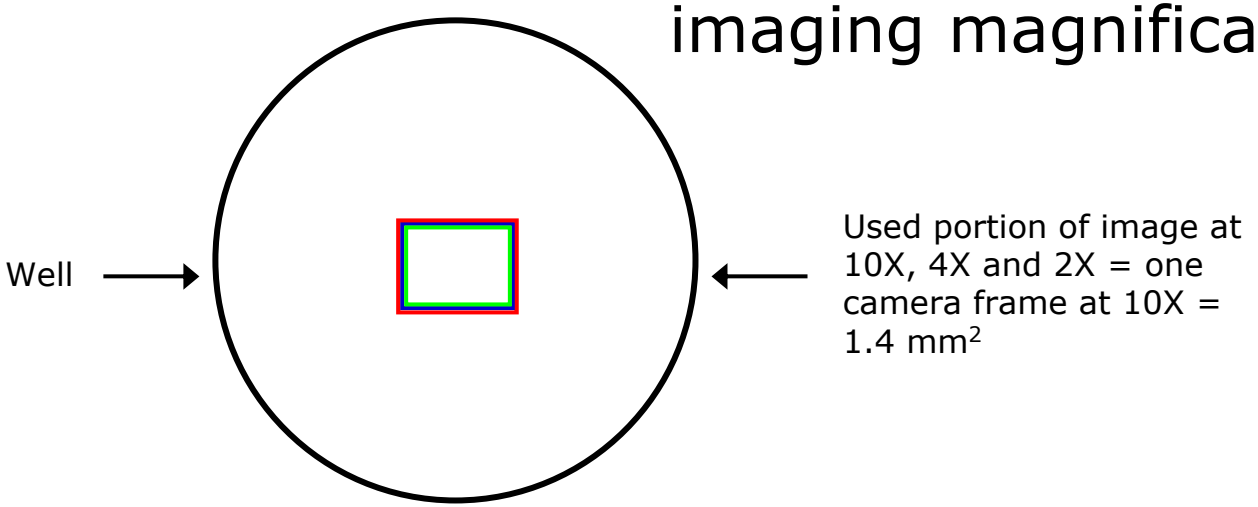




Due to non-flat well bottom the center and periphery of the image are in different focal positions. This leads to the shift of the dose curve and increased variation. Using subset of the image gives better quality than the whole image.

VITRA BIOSCIENCE

Dependency of assay quality on the number of cells and imaging magnification. Take 2.

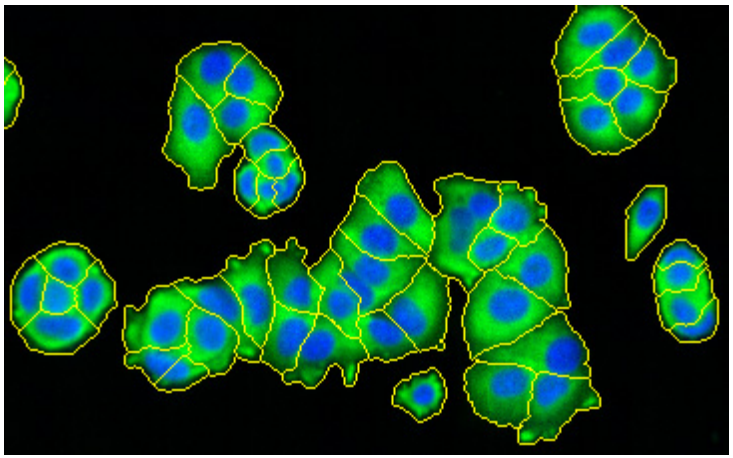


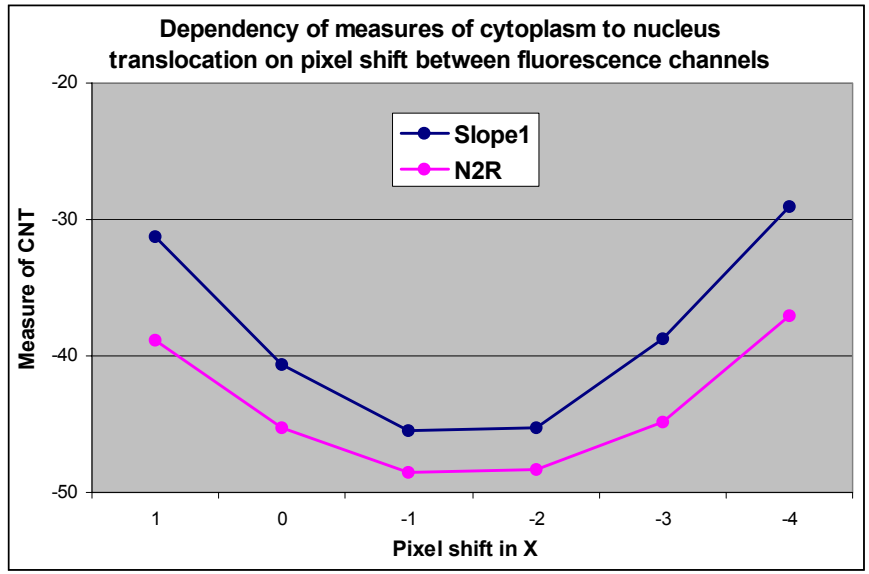
	MCF7		Magnification		
Required CellCard carriers	Image size (sq.mm)	Number of cells	2X	4X	10X
14	1.40	~700	0.78	0.87	0.89
3-4	0.35	~180	0.58	0.65	0.66
1-2	0.16	~80	0.43	0.55	0.59
	A549		Magnification		
Required CellCard carriers	Image size (sq.mm)	Number of cells	2X	4X	10X
14	1.40	~500	0.74	0.83	0.83
3-4	0.35	~130	0.58	0.65	0.66
1-2	0.16	~60	0.36	0.58	0.59

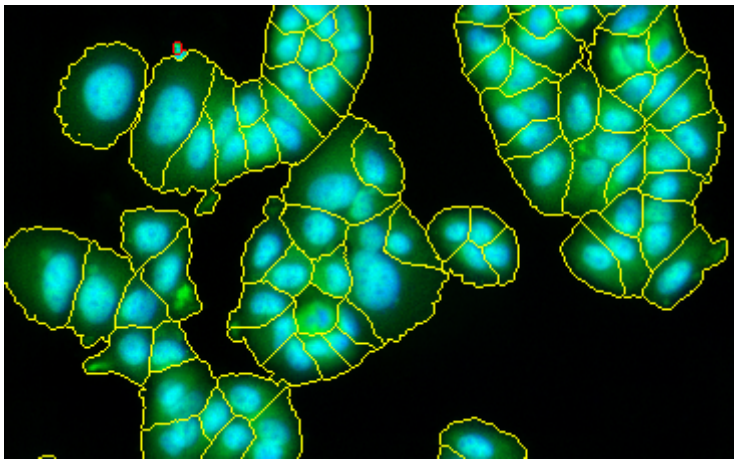
VITRA BIOSCIENCE

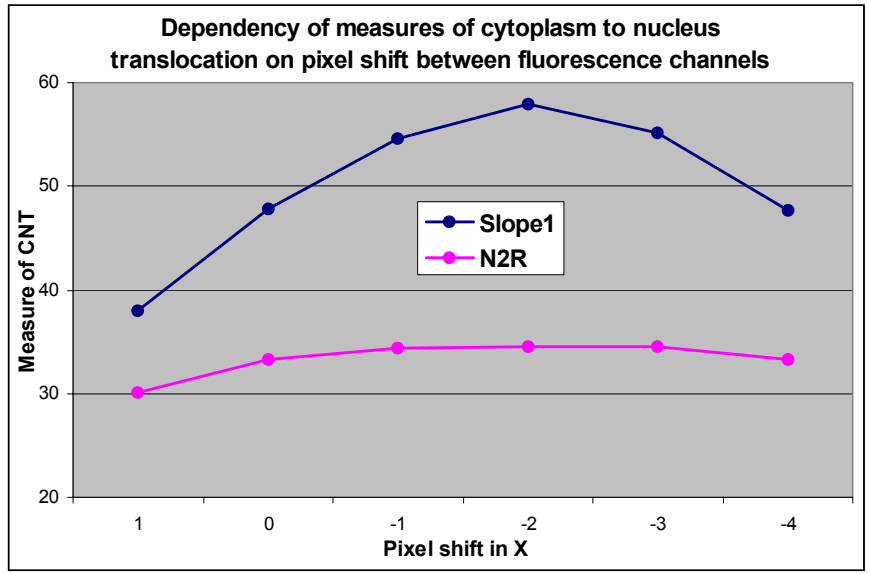
Dependency of CNT measures on pixel shift

Images at 10X







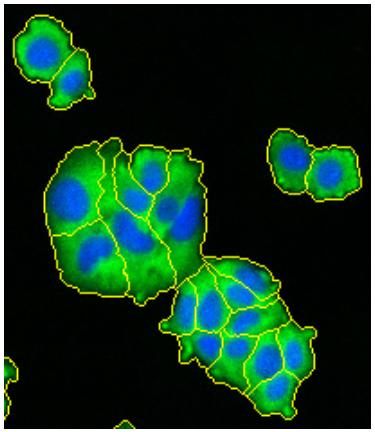


- CNT measures show strong dependency on pixel shift
- On negative cells Slope1 and N2R behave similarly relative to pixel shift
- On positive cells Slope1 is more dependent on pixel shift than N2R

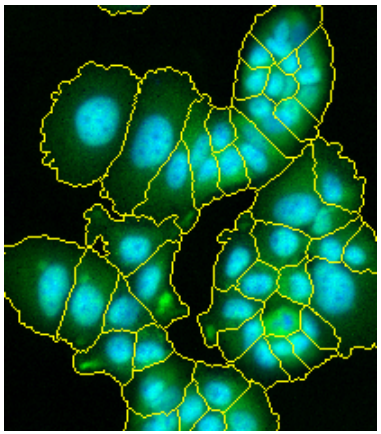
VITRA BIOSCIENCE

Dependency of CNT measures on focusing accuracy

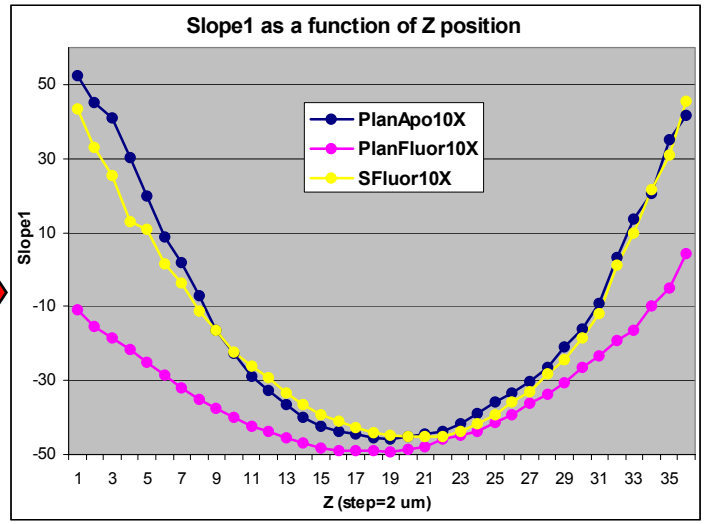
Images at 10X

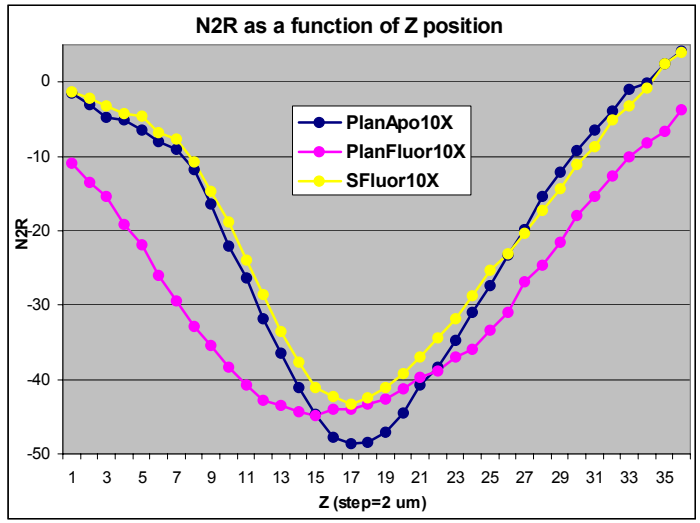


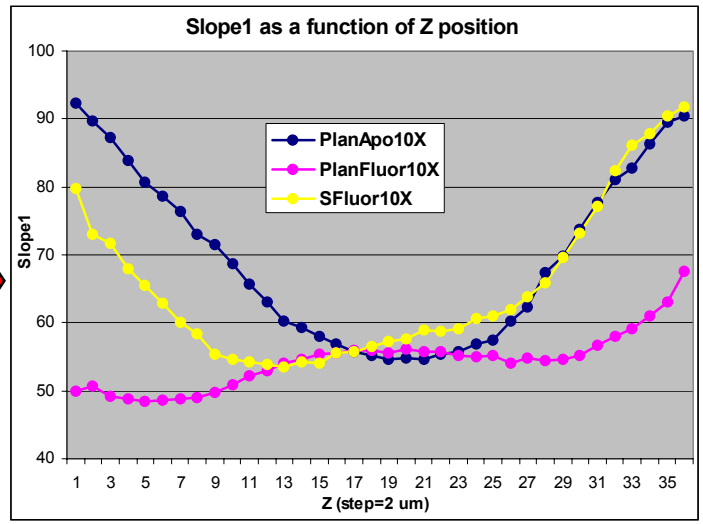
Negative

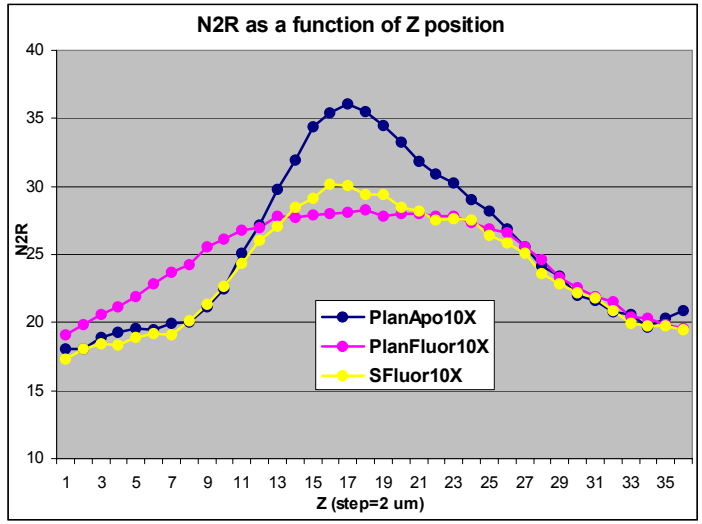


Positive







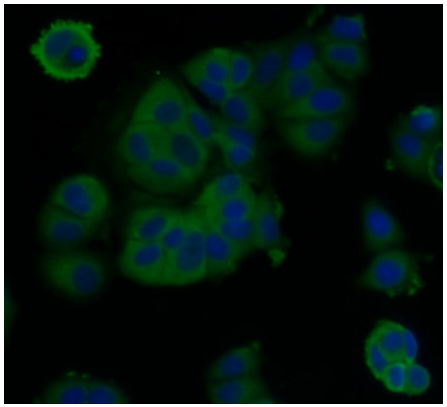


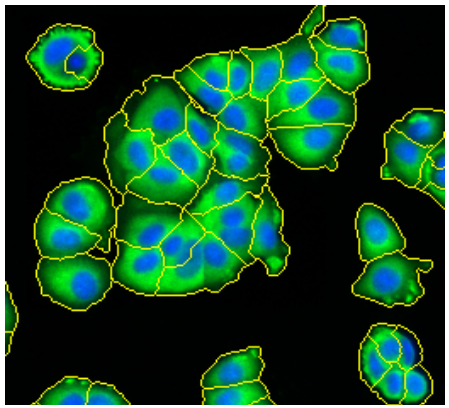
- Slope1 and N2R show similar behavior on negative cells and opposite behavior on positive cells.
- CNT measures are almost identical within 10-15 µm from the best focus position in 10X and within 40-50 µm in 2X (data for 2X not shown).
- Objectives with lower NA are more tolerant to focusing inaccuracy.
- For Slope1 PlanFluor is (marginally) the best objective; for N2R PlanApo is the best objective.

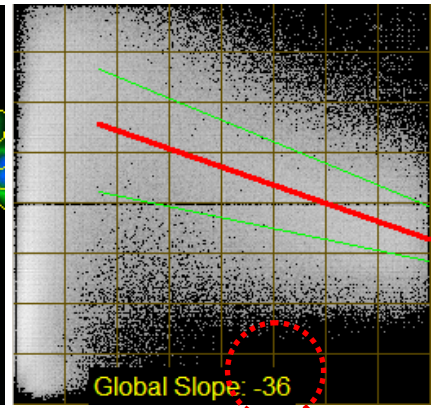
VITRA BIOSCIENCE

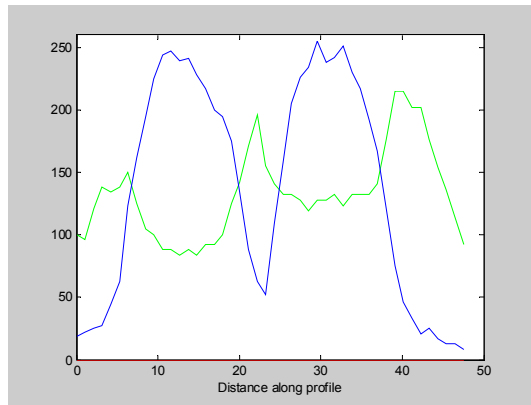
Dependency of CNT measures on the depth of field in a non-confocal system

PlanFluor10X, 0.3NA, Depth of field = 8.5 µm

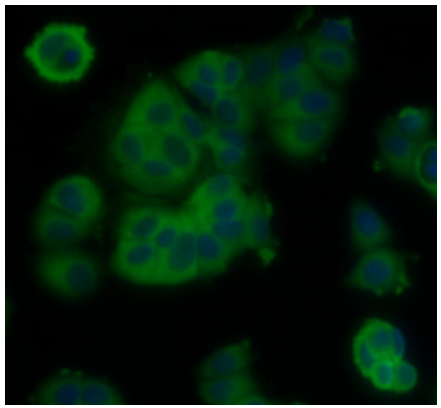


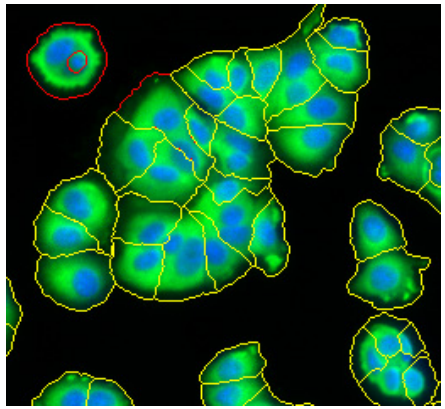


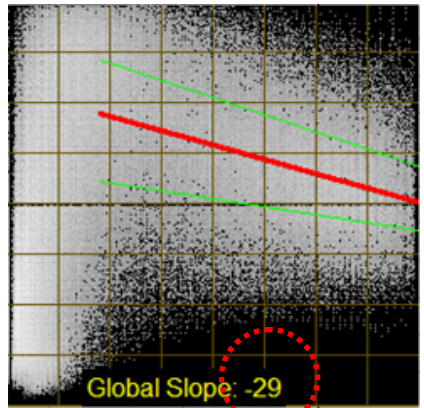


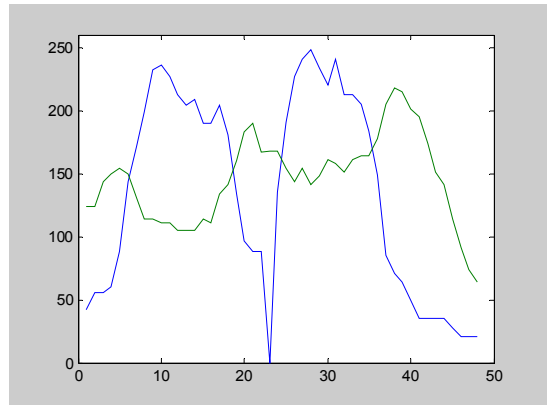


Original | Normalized | 2D distribution | Profiles







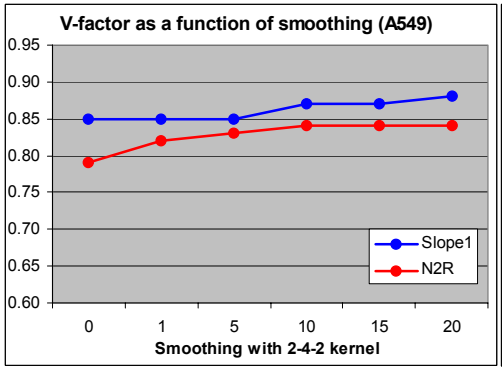


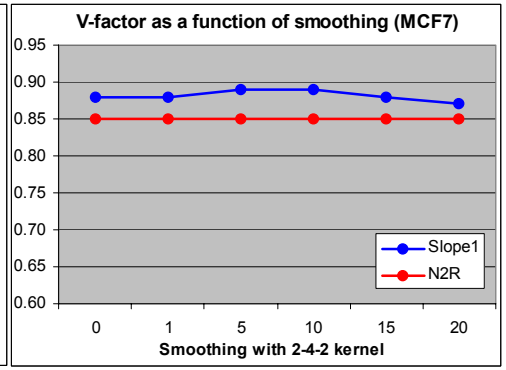
SFluor10X, 0.5NA, Depth of field = 3.6 µm

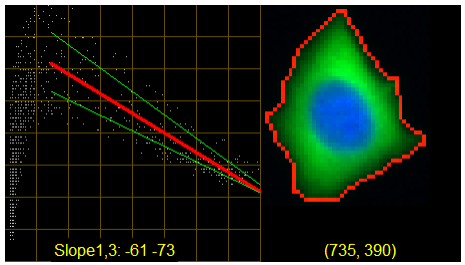
If cells are thicker than depth of field of objective, they may appear fuzzy and the CNT measurements for negative cells may be underestimated.

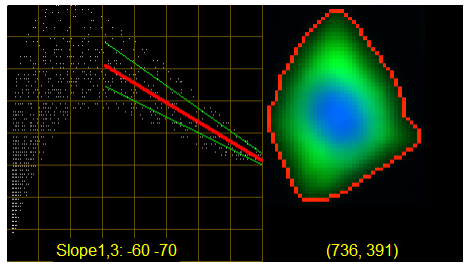
VITRA BIOSCIENCE

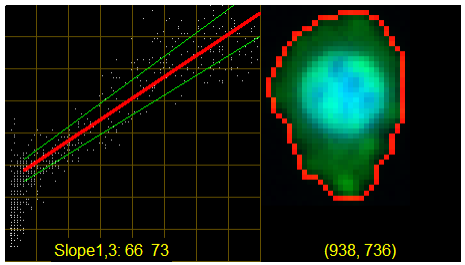
Is a confocal system better for CNT? (1)

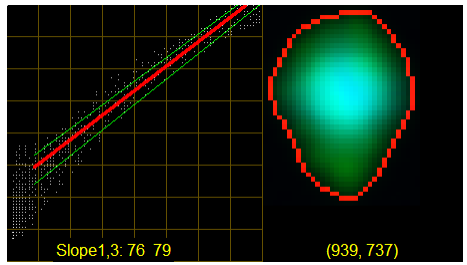


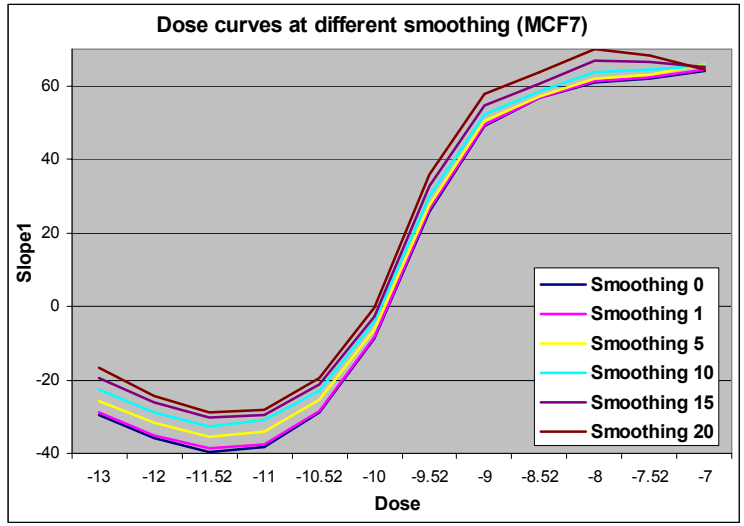


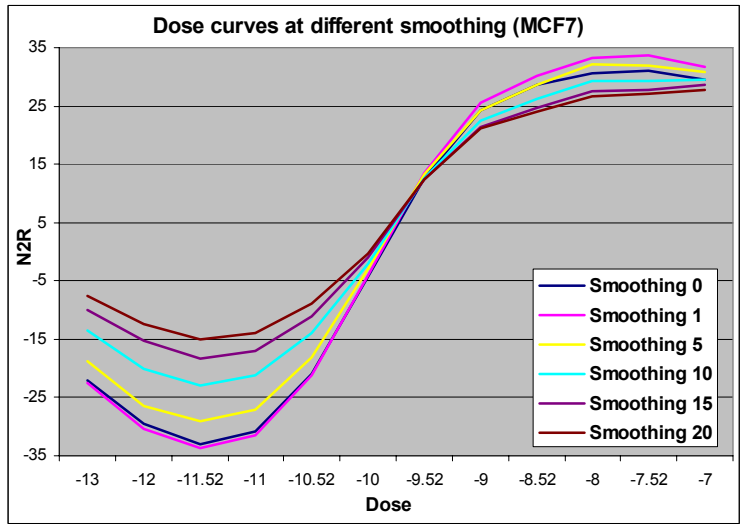










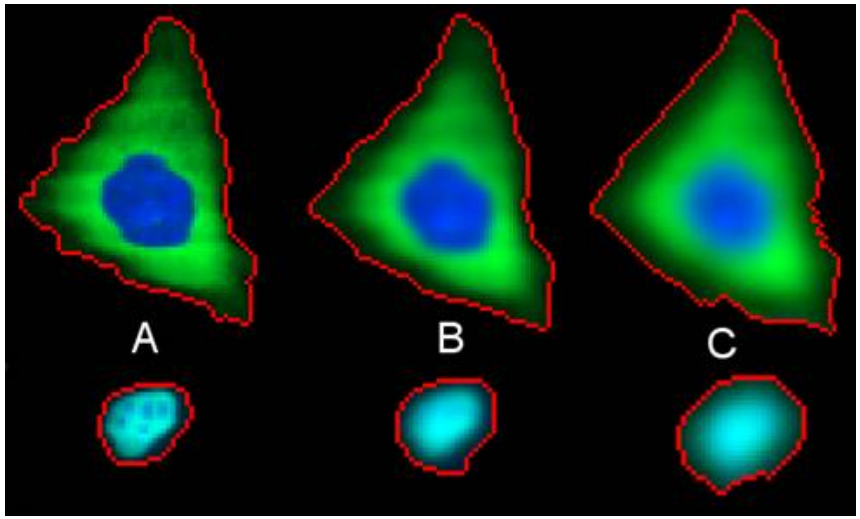


Simulation of enlarged depth of field by smoothing (10X).

For slope measures the dose curve shifts up. For mask measures the dynamic range is reduced. For both measures, however, V-factor is not reduced with smoothing. This result suggests that reducing the depth of field does not improve the assay quality.

VITRA BIOSCIENCE

Is a confocal system better for CNT? (2)



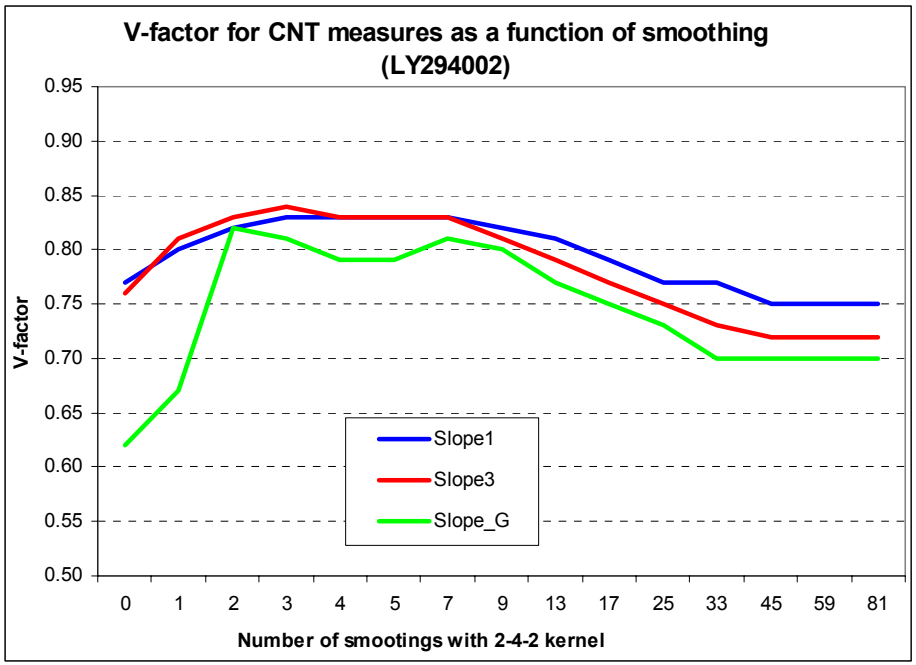
A - Smoothing = 0,0
B - Smoothing = 4,4
C - Smoothing = 25,25

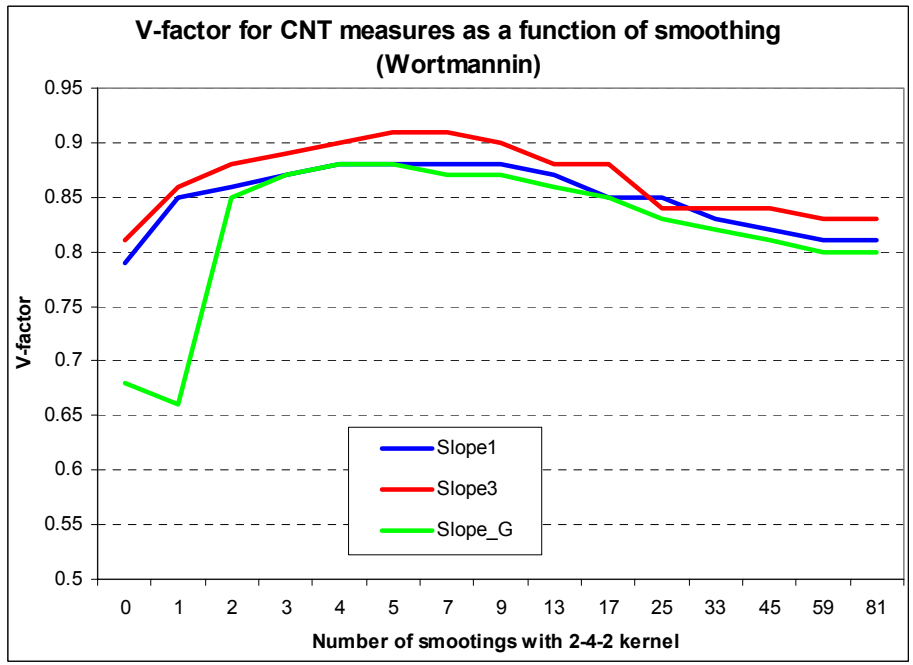
First - number smoothing of green, second number - smoothing of blue.

Images acquired on a confocal scanner; depth of field ~1μm.

Top row – "negative" cells,

Bottom row – "positive" cells.





VITRA BIOSCIENCE

Conclusions

- A new method for the analysis of cytoplasm to nucleus translocation images is presented; it does not have user parameters and provides superior quality of data compared with mask-based method.
- Segmentation of intracellular compartments may not be necessary for quantitation of intracellular processes.
- V-factor proved to be a useful benchmark to compare image analysis algorithms/measures and to determine image resolution and image size/cell number requirements.
- Better assay design and analysis can be achieved by using presented methodologies for the analysis of dependency of assay measures on cell type, magnification, number of analyzed cells, accuracy of focusing, depth of field, pixel shift and plate flatness.
- Less is more: choosing a subset of cells in the same focal plane may improve assay quality; alternatively, increasing the cell number only helps if all cells are in focus.
- The cell number and magnification requirements of cytoplasm to nucleus translocation assay put it within the parameter range of the CellCard system.
- Future developments of this algorithm will focus on enabling CellPlex assays on the CellCard System.

VITRA BIOSCIENCE

References

1. J.-H. Zhang, T.D.Y. Chung, K.R. Oldenburg "A simple statistical parameter for use in evaluation and validation of high throughput screening assays", J. Biomol. Screening 4: pp. 67-73, 1999

2. I. Ravkin "Quality Measures for Imaging-based Cellular Assays" SBS 2004 conference poster #P12024

3. I. Ravkin, V. Temov, A.D. Nelson, M.A. Zarowitz, M. Hoopes, Y. Verhovsky, G. Ascue, S. Goldbard, O. Beske, B. Bhagwat, H. Marciniak "Multiplexed high-throughput image cytometry using encoded carriers", Proc. SPIE Vol. 5322, pp. 52-63, 2004 (Imaging, Manipulation, and Analysis of Biomolecules, Cells, and Tissues II; Dan V. Nicolau, Joerg Enderlein, Robert C. Leif, Daniel L. Farkas; Eds.)

4. S. Beucher and F. Meyer, "The Morphological Approach to Segmentation: The Watershed Transformation" in: Mathematical Morphology in Image Processing, E.R. Dougherty – Ed., pp. 433 – 481, Marcel Dekker, New York, 1993

5. L. Vincent, P. Soille, "Watersheds in Digital Spaces: An Efficient Algorithm Based on Immersion Simulations", IEEE Transactions of Pattern Analysis and Machine Intelligence, 13, No. 6, pp. 583-598, 1991

6. Image Processing Toolbox, The MathWorks, Inc. http://www.mathworks.com/products/image/

7. J. Serra, Image Analysis and Mathematical Morphology, Vol. 1. Academic Press, London, 1989

VITRA BIOSCIENCE