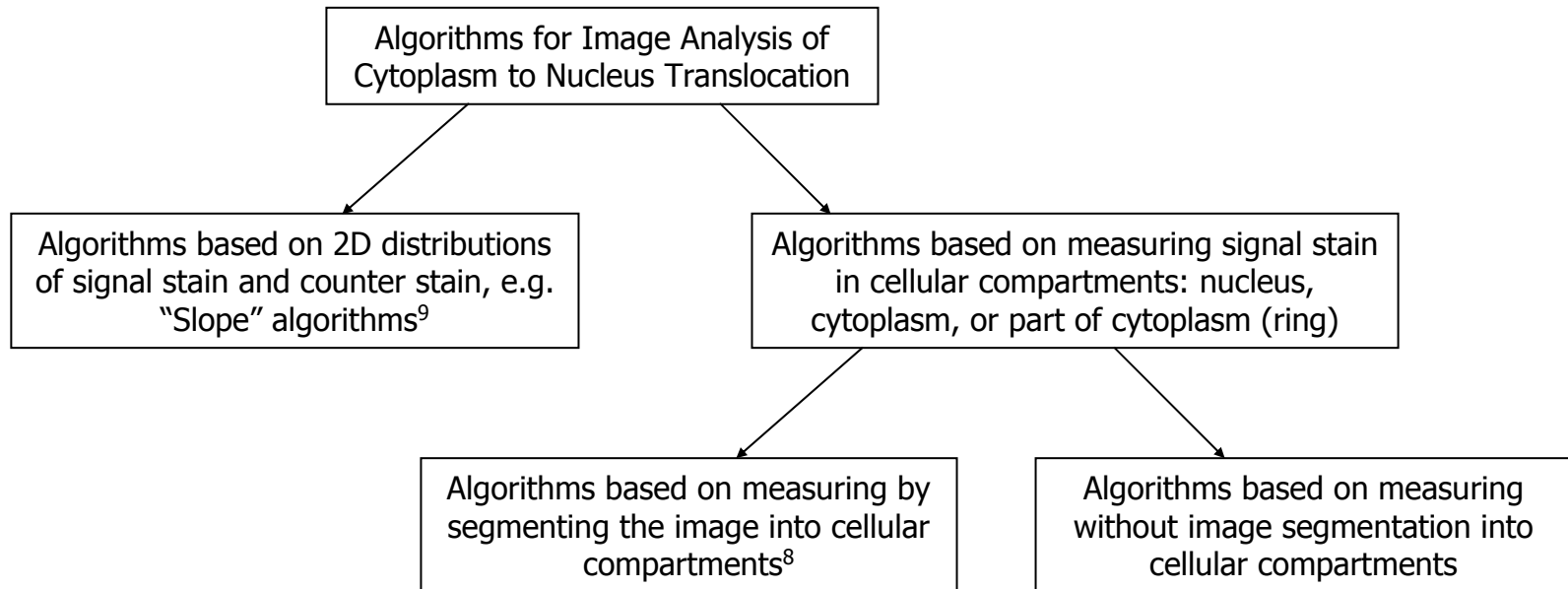


## **Comparison of Several Classes of Algorithms for Cytoplasm to Nucleus Translocation**

*Ilya Ravkin and Vladimir Temov, Vitra Bioscience, Inc.*

Cellular assays assessed by image analysis provide a new type of information but also present several new sources of variability compared to traditional screening. Understanding of these sources and of the behavior of image analysis algorithms under different conditions is important for obtaining reliable data from imaging assays. We have developed several new and implemented several commonly used measures of cytoplasm to nucleus translocation. The quality of these measures is analyzed with respect to user parameters, number of analyzed cells, uniformity of illumination, accuracy of focusing, and depth of field.

# Classes of Algorithms in this Study.



# Preprocessing Common to all Algorithms.

Before running specific algorithms for cytoplasm to nucleus translocation we perform several preprocessing steps:

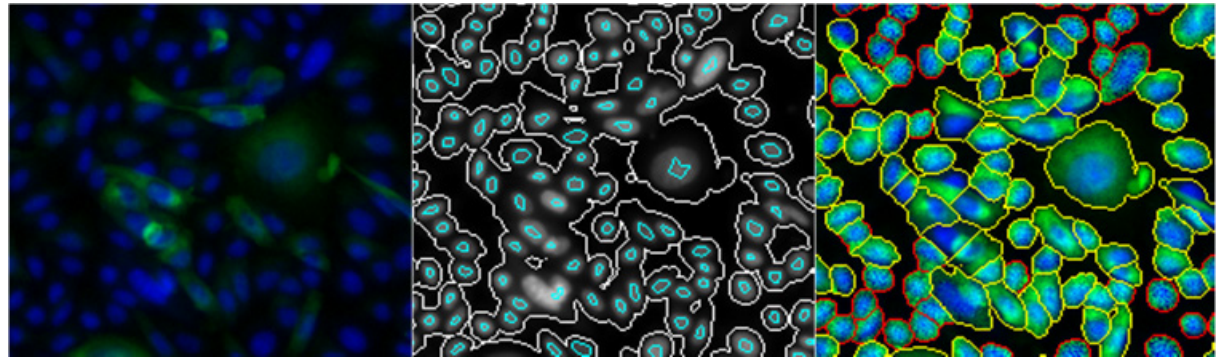
1. **Rejection of artifacts.** This removes areas of fluorescent debris determined primarily by saturation of intensity.
2. **Shading correction** is available, but was turned off in this study because we used intensity normalization (see below).
3. **Background removal.** This operation removes background by morphological opening. Saturated areas are not changed.
4. **Intensity normalization.** This is a complex operation based on partitioning the image into components; it consists of the following parts:
  - Detection of markers of components
  - Detection of watersheds, which serve as separating lines between components
  - Adjustment of intensity in individual components

Steps 1-3 are rather common in processing of microscope images. Step 4 will be described in some detail below.

# Partitioning into Components. Markers.

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

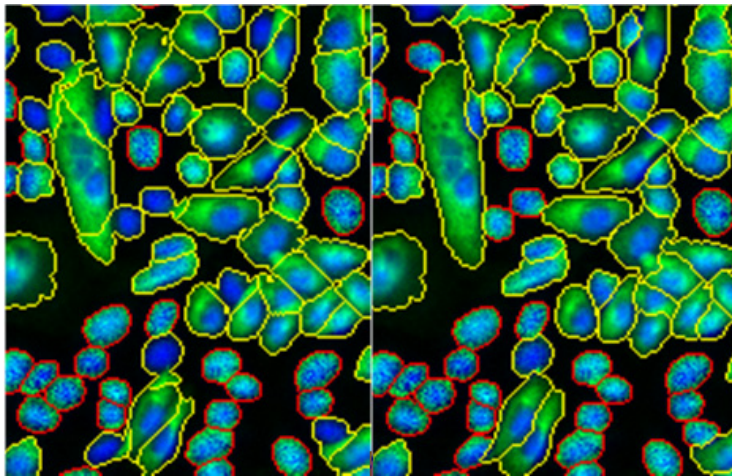
Markers are found by the following algorithm. A fixed value (marker contrast) is subtracted with saturation from the image of nuclear counterstain and the resulting image is reconstructed<sup>7</sup> within the image of nuclear counterstain. This image is then subtracted from the counterstain image and 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 in the original intensity image 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<sup>7</sup> may be used too.



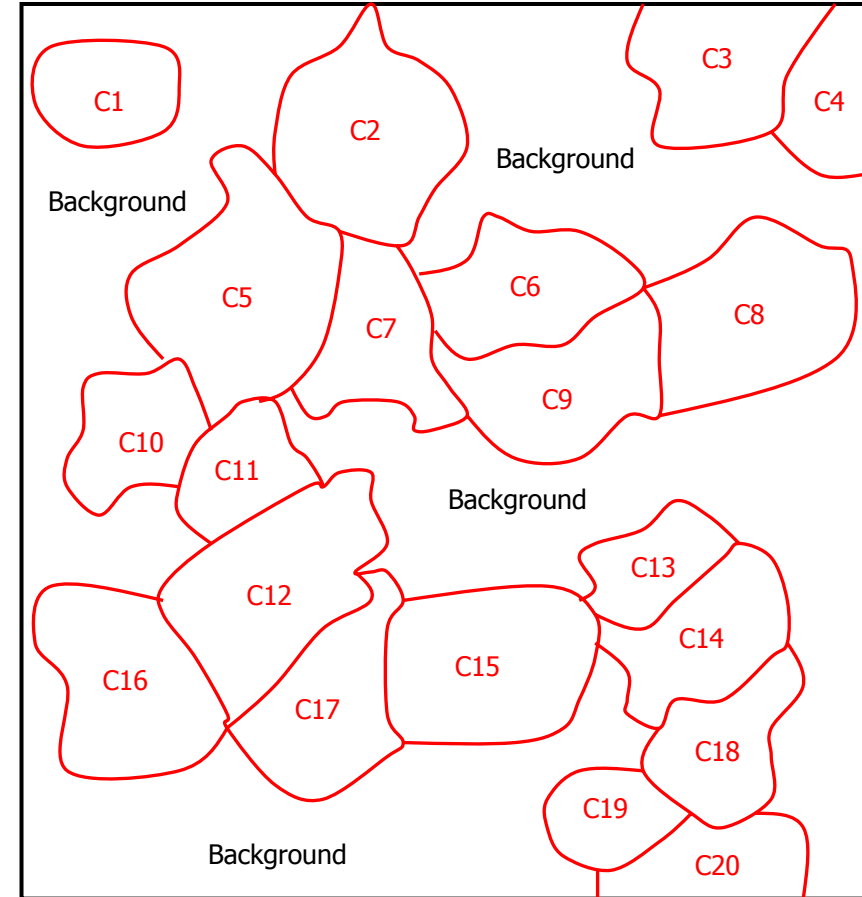
Left - original image of cells expressing GFP-fused p65 protein. Middle - smoothed nuclear counterstain image shown in grayscale with overlaid contours of markers (cyan) and contours of background areas (white). Right - partitioning of the image induced by the markers.

# Partitioning into Components. Watersheds of Combined Intensity Images.

Separation lines are defined as the watershed<sup>4,5,6</sup> 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 (for such nuclear stains as DRAQ5, which give a low level staining of the whole cell, the use of signal stain is not needed).

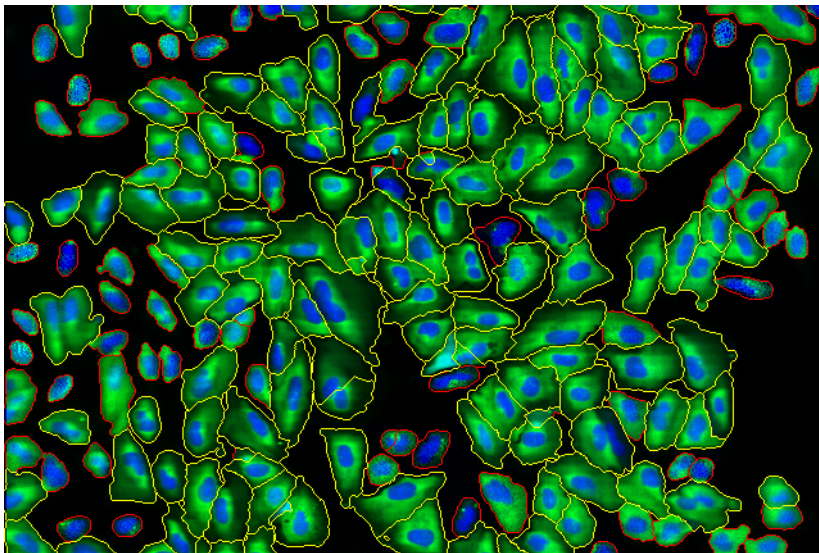
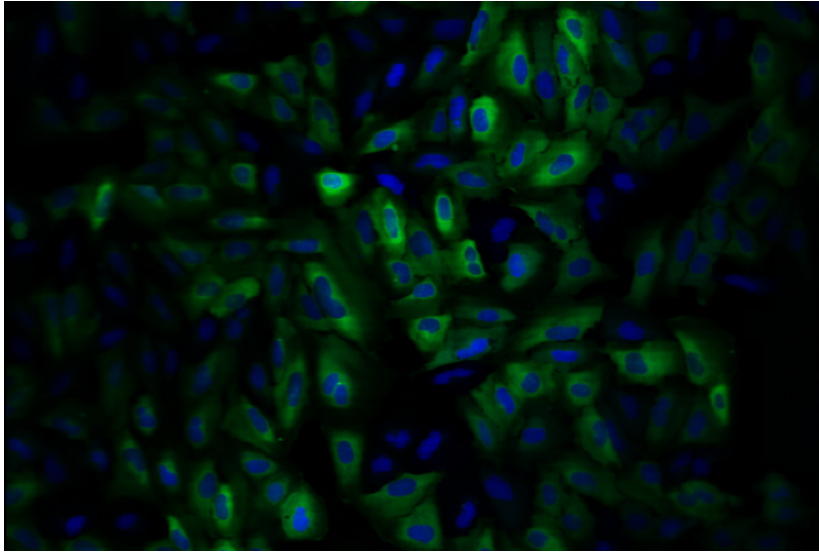


Left - image partitioning using only nuclear stain image (blue).  
Right - image partitioning using linear combination of nuclear and signal stains. Coefficient of linear combination for the nuclear image = 0.3, for the signal stain = 0.7



The result of partitioning is the labeling of the background and of the components. Components contain preferably individual cells, but may contain clusters of cells; they do not necessarily correspond to cell shape.

# Equalization of Intensities.



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

Equalization of intensities is not strictly required by the algorithms described below. Identical results can be achieved by adaptively choosing parameters (specifically, thresholds) for individual cells or by normalizing the 2D intensity distributions.

However, we find that equalizing the intensities of cells gives additional feedback to the user and may reveal features that were not seen before normalization. In addition, it simplifies the subsequent processing.

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

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

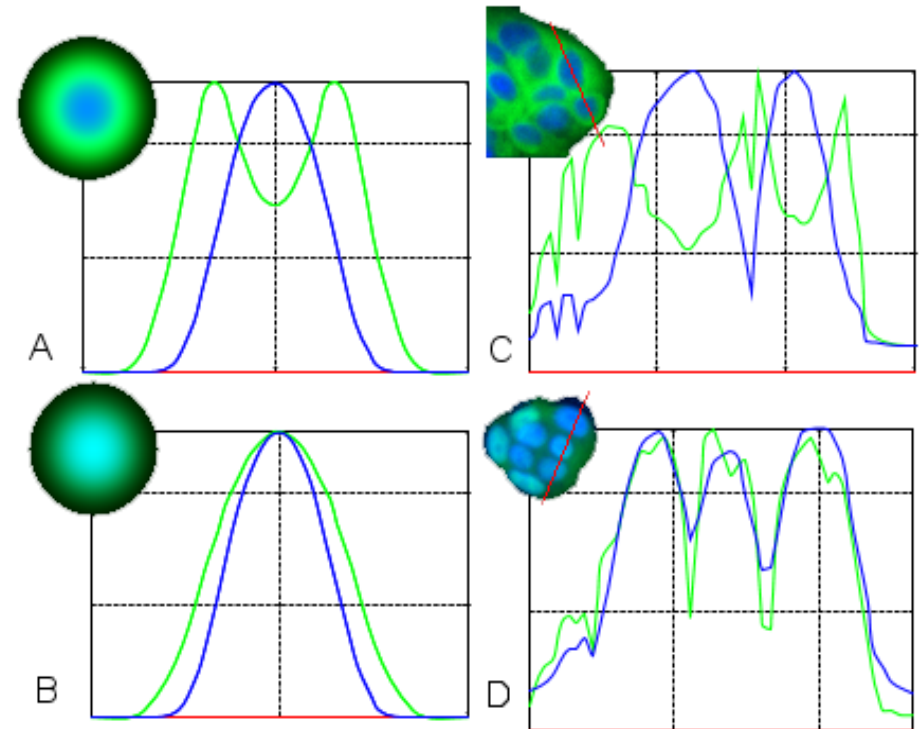
# Algorithms Based on 2D Distributions of Signal Stain and Counter Stain.



# 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.)



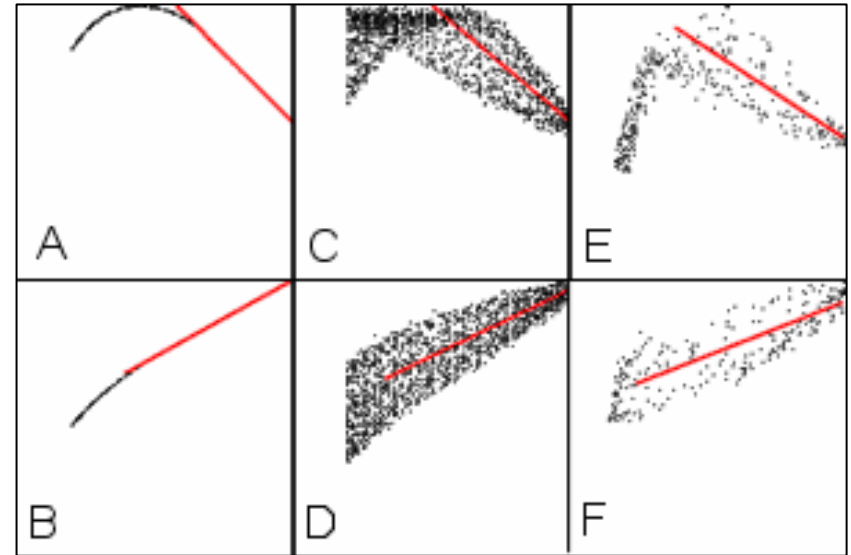
A,B – profiles through model distributions; C,D – profiles through real cells. A,C – negative, B,D – positive. Blue – counterstain, green – signal stain.



# Model and Experimental 2D Distributions of Stains. Translocation Measures Derived from 2D Distributions.

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 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 close to the center of the nucleus. The farther from the center, the more diffuse the distribution, and the less reliable the approximation becomes.

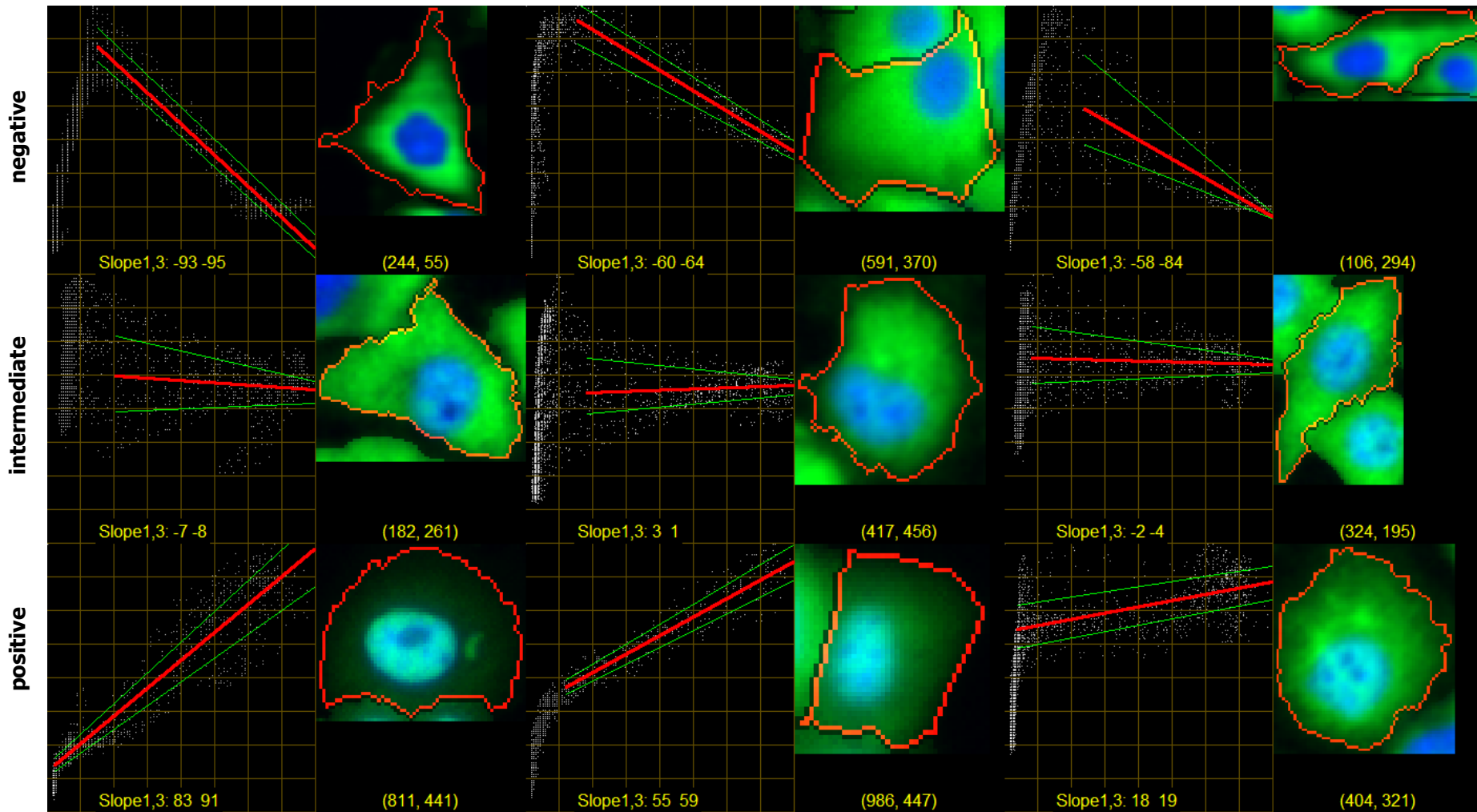


Cross-histograms of counterstain (X-axis) and signal stain (Y-axis) in ideal model (A,B), perturbed model (C,D) and a real cell (E,F). Scale on both axes is 0-255. Red lines show the calculated approximation.

The portion of the distribution that is used for approximation with the straight line segment is found by plotting the approximated slope going from right to left and selecting the range where this approximation is the most stable. This line segment is shown in red (next slide) and its tangent multiplied by 100 is called Slope1. Top panel – nuclear localization (positive), middle panel – intermediate localization, bottom panel – cytoplasm localization (negative) of protein.

A variation on this method is to calculate two more slopes (shown by green lines). The top line is the regression line calculated on all points above the original slope segment; 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 average of the three lines is used. This measure is called Slope3.

# Translocation Measures Derived from 2D Distributions.



Examples of Slope measures in cells with different protein localization.

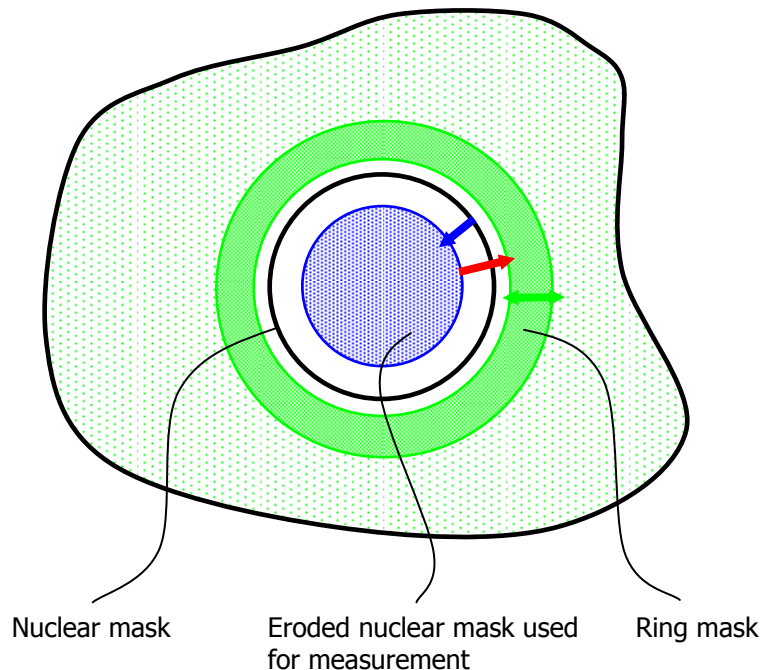
Top panel – nuclear localization (positive), middle panel – intermediate localization, bottom panel – cytoplasm localization (negative) of protein.

Slope1 - red line. Slope3 – one of green lines.

# Algorithms Based on Measuring Signal Stain in Cellular Compartments.

# Algorithms Based on Segmentation into Nucleus and Cytoplasm.

This class of algorithms<sup>8</sup> is the most widely used for analysis of cytoplasm to nucleus translocation. The approach is to use the nuclear stain image to identify the nuclear mask; then to find the mask of what is some approximation of the cytoplasm, usually a ring around the nucleus as shown in the figure. After the masks representing nucleus and cytoplasm are found, the average and/or total intensity of the signal stain are calculated in each mask. The measure of translocation is the ratio or difference between these intensities.



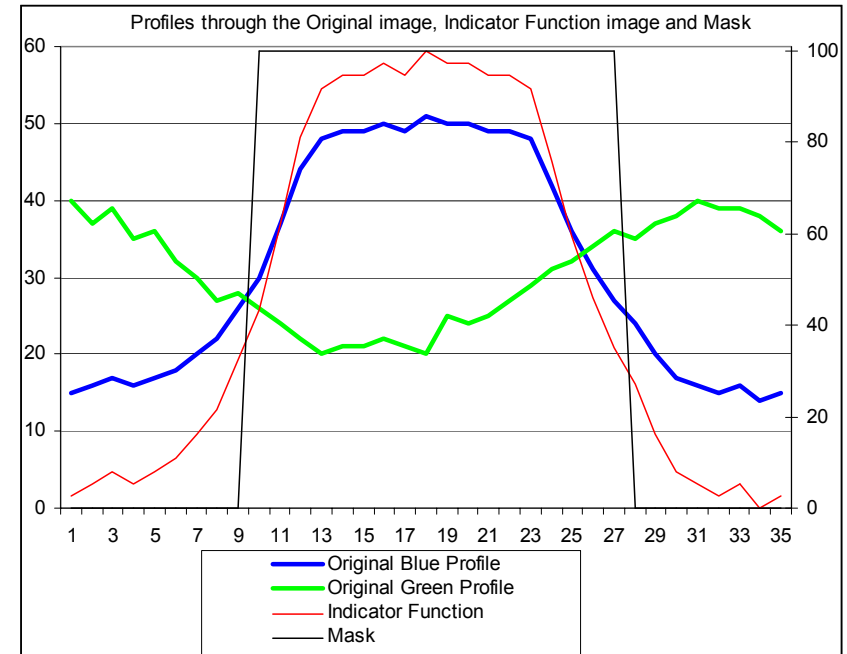
- A** – Erosion of the nuclear mask
- B** – Gap between the eroded nuclear mask and the ring mask
- C** – Width of the ring mask

# Algorithms Based on Indicator Functions. Measurement Without Segmentation.

Suppose the task is to measure the amount of fluorescence emitted by proteins in the nucleus of a cell. The figure on the right gives a representation of the images by their intensity profiles along a straight line through the cell.

The conventional approach is to first use the blue profile to generate a mask of the nucleus (the profile through the mask is shown in black); then - to integrate intensities under the green profile within this mask.

Our approach is to use the blue profile to generate an indicator function shown in red. This name was given because the function indicates the certainty we have in that a pixel belongs to the object. The indicator function may be produced by simply scaling the blue profile to the range  $[0-1]$ , or  $[0-100\%]$ , or  $[0-255]$  (for convenience of representation as image). Integrated intensity of green in the object is produced by integrating green in the whole image, using the indicator function as weight.



The application of this approach does not require having two measured images; green and blue in this example. It can be applied to a single image as well. In this case the image that generates the indicator function and the image which is being integrated is the same image.

This approach has several advantages: it avoids possible errors of segmentation, it is faster to perform and has better scalability with magnification.

The indicator function may be produced from one or more acquired images by a complex algorithm.

# Use of Indicator Functions for Analysis of Cytoplasm to Nucleus Translocation. 1.

The indicator function method can be applied to the analysis of images of the cytoplasm to nucleus translocation assay. In this assay, cells are visualized in two wavelengths corresponding to nuclear stain and to protein stain. The goal of analysis is to quantify translocation of proteins between the cytoplasm and the nucleus. Images representing extreme states of this assay are shown in the next slide in Fig. A,B,C.

The indicator function for the nucleus, in the simplest case, is the scaled nuclear stain image as described above and shown in Fig. D. It can also be a smoothed nuclear image or any other derivative of the nuclear image that may be beneficial. The amount of stained protein in the nucleus is the integral of image A with the weight given by image D over the area of the component containing the cell. For easier visualization, image A weighted by image D is shown as image G.

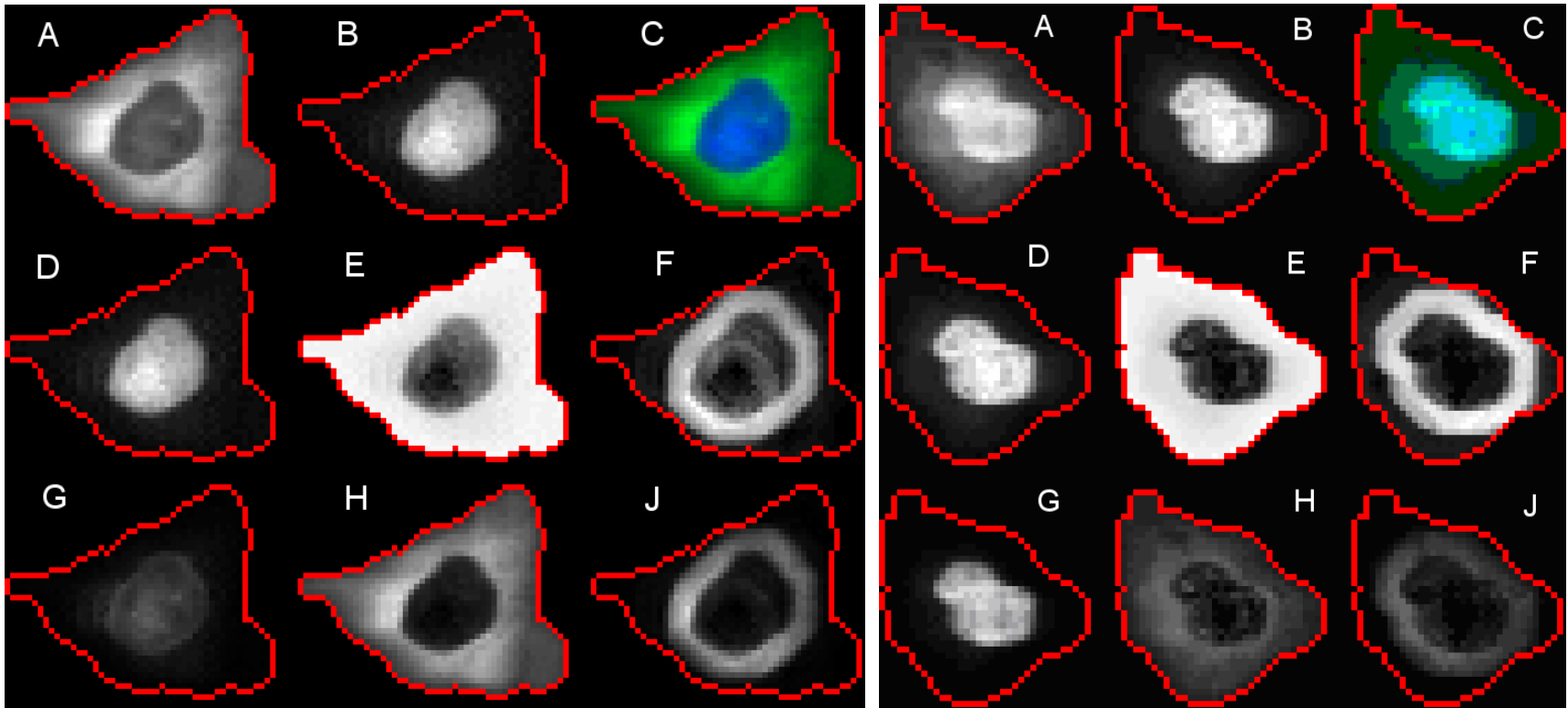
Thus, the amount of stained protein in the nucleus is the sum of all pixels of G. Another useful measure is the average intensity of protein staining in the nucleus. This measure is calculated as the sum of all pixels of G divided by the sum of all pixels of D.

In addition to the nucleus, the two other compartments that are useful to assess for this assay are the cytoplasm and the perinuclear region of cytoplasm, which we will call the ring. The indicator function for the cytoplasm is the complement of the indicator function for the nucleus within the compartment containing the cell (Fig. E). The indicator function for the ring is produced by dilating the indicator function for the nucleus by a given amount (width of the ring), subtracting the indicator function for the nucleus from it and normalizing to the given scale as described above. Other methods of producing the indicator function for the ring may be used too.

The amount of stained protein in the cytoplasm is the sum of all pixels of H. The average intensity of protein staining in the cytoplasm is the sum of all pixels of H divided by the sum of all pixels of E.

The amount of stained protein in the ring is the sum of all pixels of J. The average intensity of protein staining in the ring is the sum of all pixels of J divided by the sum of all pixels of F.

# Use of Indicator Functions for Analysis of Cytoplasm to Nucleus Translocation. 2.



Left panel corresponds to negative state of the assay when stained proteins are localized predominantly in cytoplasm.

Right panel corresponds to positive state of the assay when stained proteins are localized predominantly in nucleus.

A – image of stained proteins (green),

B – image of nuclear stain (blue),

C – pseudo-color composite of A and B,

D – same as B normalized to [0-255] and used as indicator function of nucleus,

E – complement of D within the component containing the cell (shown by red contour) – used as indicator function of cytoplasm,

F – dilation of D minus D normalized to [0-255] – used as indicator function of the ring around the nucleus.

G – A weighted (multiplied) by D – the image of protein content in nucleus,

H – A weighted by E – the image of protein content in cytoplasm,

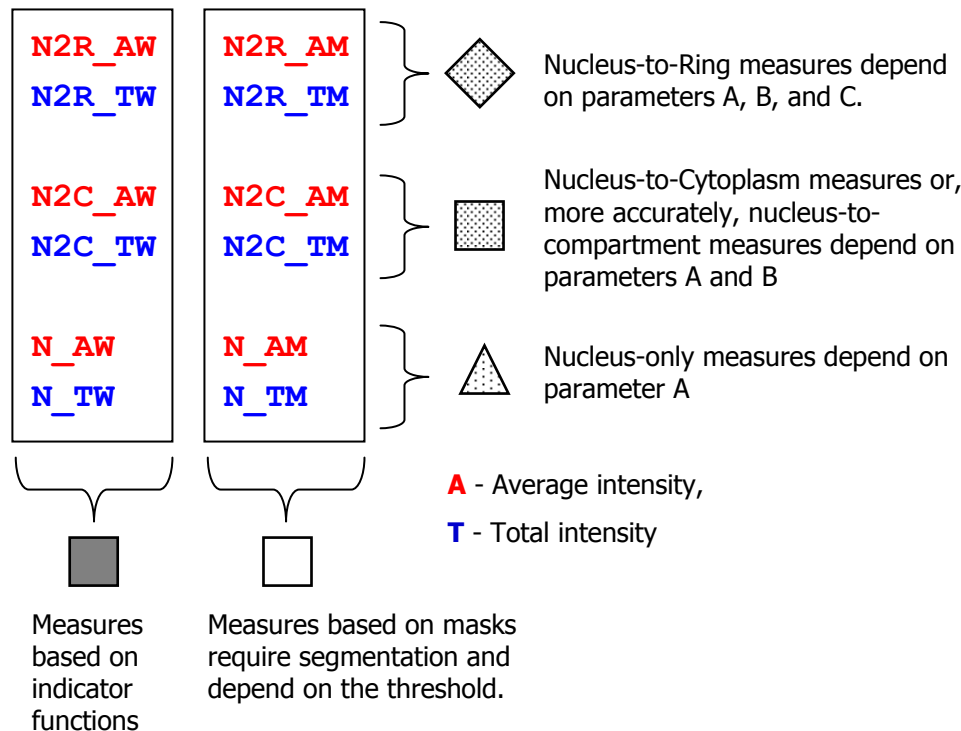
J – A weighted by F – the image of protein content in a ring around the nucleus.



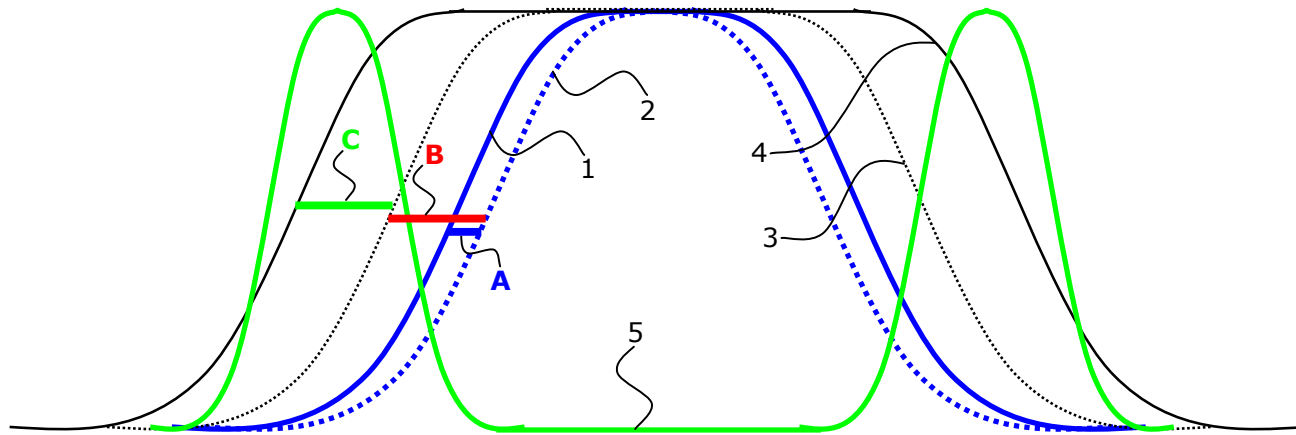
# Family of Measures for Algorithms Based on Measuring Signal Stain in Cellular Compartments.

A family of assay measures can be constructed as ratios of the average (A) or total (T) amounts of stained protein in the three compartments: nucleus (N), cytoplasm (C), and ring (R). For more convenient scaling we use  $100 \cdot \log_2()$  of all ratio measures. In addition, a non-ratio measure of the amount of signal stain in the nucleus may be also useful. The measures may be produced by segmentation into masks of nucleus and cytoplasm (M) or without segmentation by the use of indicator functions (W). Each measure may be calculated for each cell and then the well measure is some statistic of the cell population, e.g. mean or median; alternatively these measures may be calculated globally on the whole image (G). Global measures are not included in this study.

The family of measures is given by the formulas:  $N2\{R|C\}_{\{A|T\}\{W|M\}}\{ |G\}$  and  $N_{\{A|T\}\{W|M\}}\{ |G\}$ .



# Parameters of the Indicator Function Algorithm for Analysis of Cytoplasm to Nucleus Translocation.



Parameters of the indicator-function algorithm for analysis of cytoplasm to nucleus translocation:

**A** – Erosion of the nuclear indicator function

**B** – Gap between the eroded nuclear indicator function and the ring indicator function

**C** – Width of the ring indicator function

**1** (solid blue profile) – original nuclear indicator function

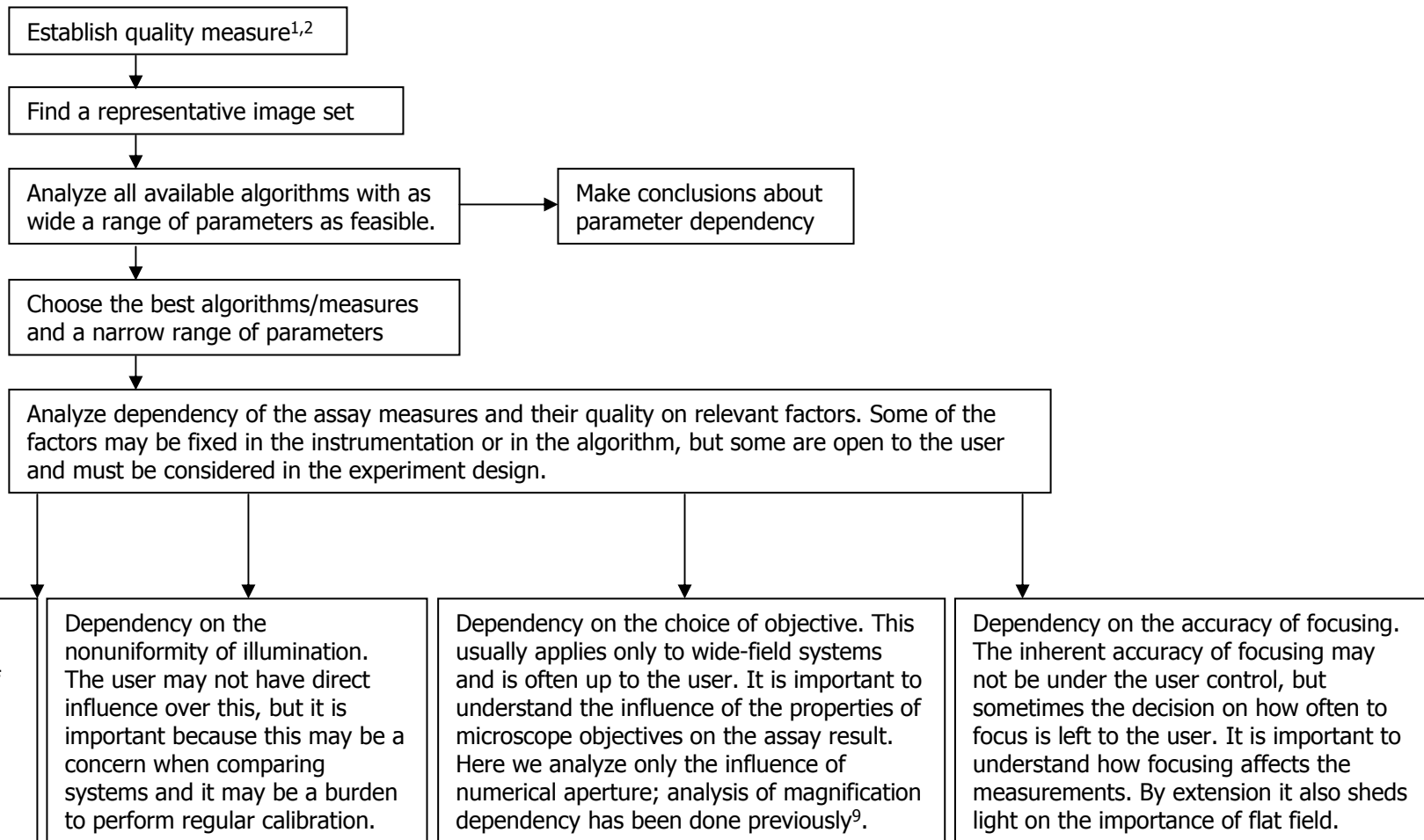
**2** (dotted blue profile) – eroded nuclear indicator function

**3** (dotted black profile) - nuclear indicator function dilated by B-A

**4** (solid black profile) - nuclear indicator function dilated by B-A+C; this is the indicator function of nucleus and ring together

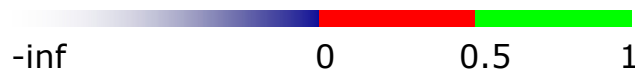
**5** (solid green profile) – indicator function of the ring produced by subtracting profile 3 from profile 4 and normalizing to full scale.

# Methodology of Algorithm Evaluation.

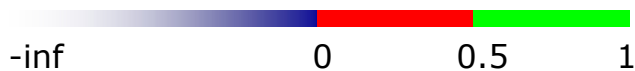


# V-factor – a Generalization of Z-factor.

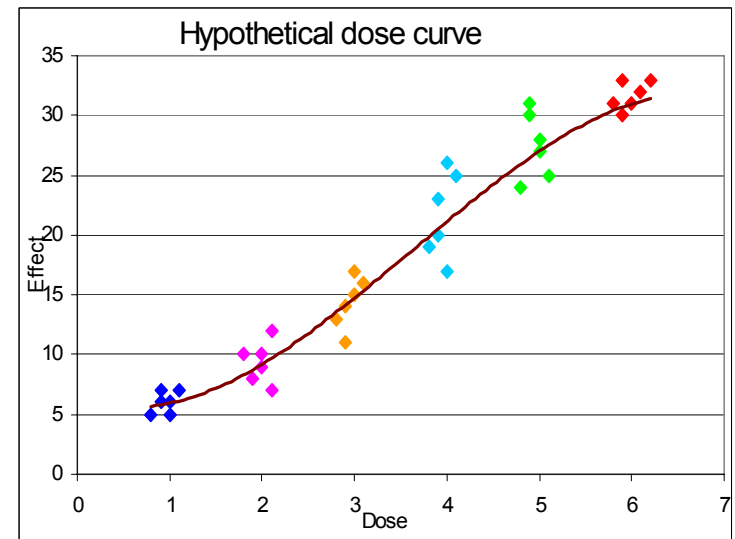
$$(1) \quad Z = 1 - 3 \left( \frac{SD_{pos} + SD_{neg}}{|M_{pos} - M_{neg}|} \right)$$



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



$$(3) \quad SD_{of\_fit} = \sqrt{\frac{1}{n} \sum_1^n (f_{model} - f_{experiment})^2}$$





For a complete discussion of v-factor see ref. 2,3.

# Description of Plates Used in the Algorithm Comparison.

## Vitra plate

	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-α 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 TNFα for 30 minutes. This treatment results in the activation of NFκB 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 acquired at Vitra Bioscience on the CellCard reader - a microscope-based system. Image size 1360\*1024 pixels, 8-bit per pixel.

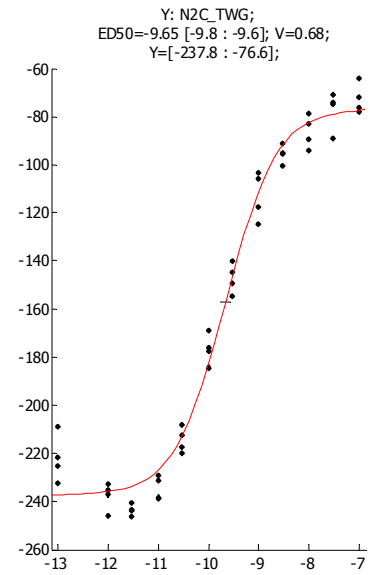
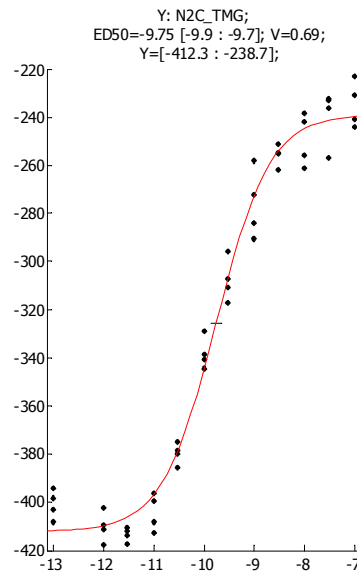
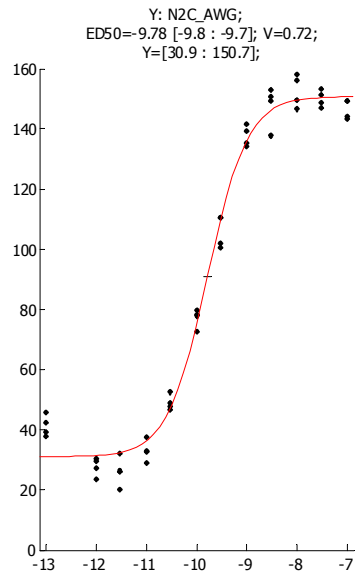
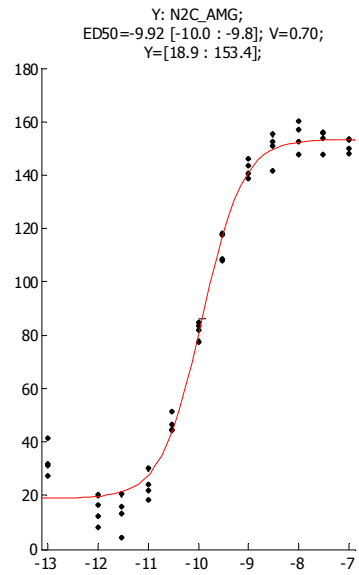
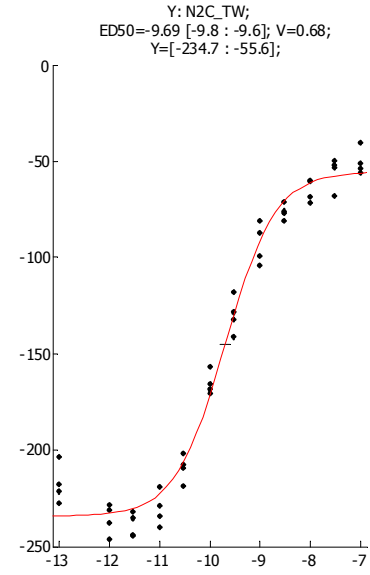
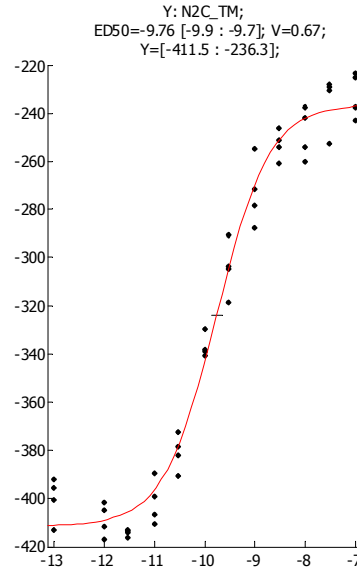
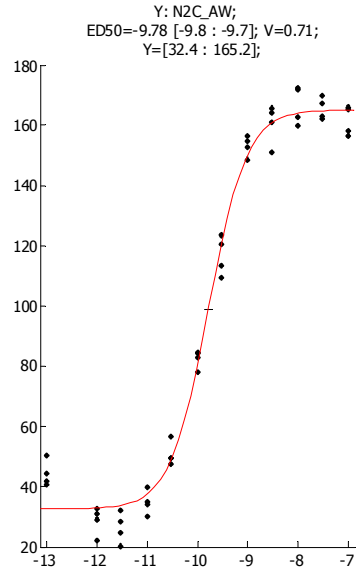
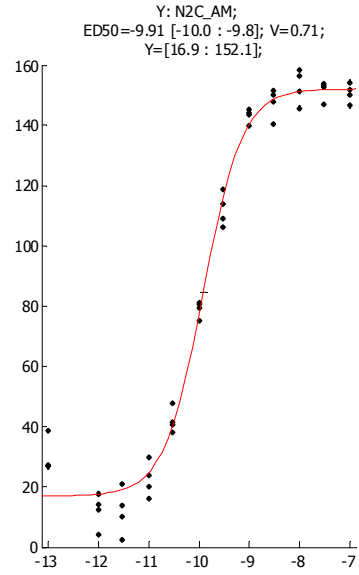
## BioImage plate

	1	2	3	4	5	6	7	8	9	10	11	12
A			0.97656	1.95313	3.90625	7.8125	15.625	31.25	62.5	125	250	
B			0.97656	1.95313	3.90625	7.8125	15.625	31.25	62.5	125	250	
C			0.97656	1.95313	3.90625	7.8125	15.625	31.25	62.5	125	250	
D			0.97656	1.95313	3.90625	7.8125	15.625	31.25	62.5	125	250	
E			0.3125	0.625	1.25	2.5	5	10	20	40	80	
F			0.3125	0.625	1.25	2.5	5	10	20	40	80	
G			0.3125	0.625	1.25	2.5	5	10	20	40	80	
H			0.3125	0.625	1.25	2.5	5	10	20	40	80	

 Wortmannin: 250 nM, 2 fold dilutions, 9 points  LY294002: 80 μM, 2 fold dilutions, 9 points

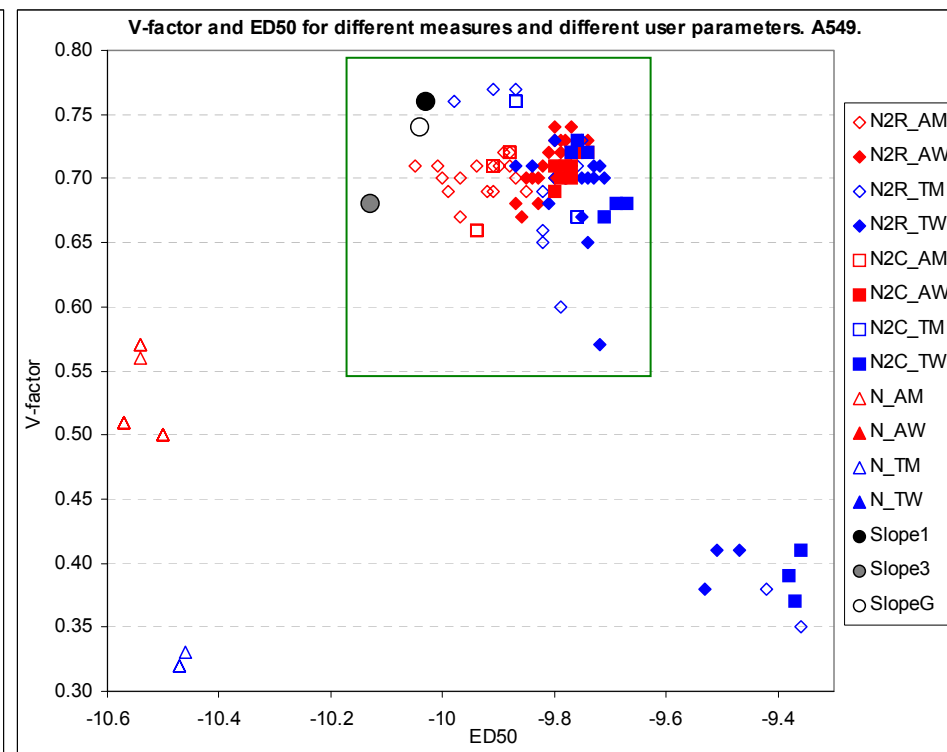
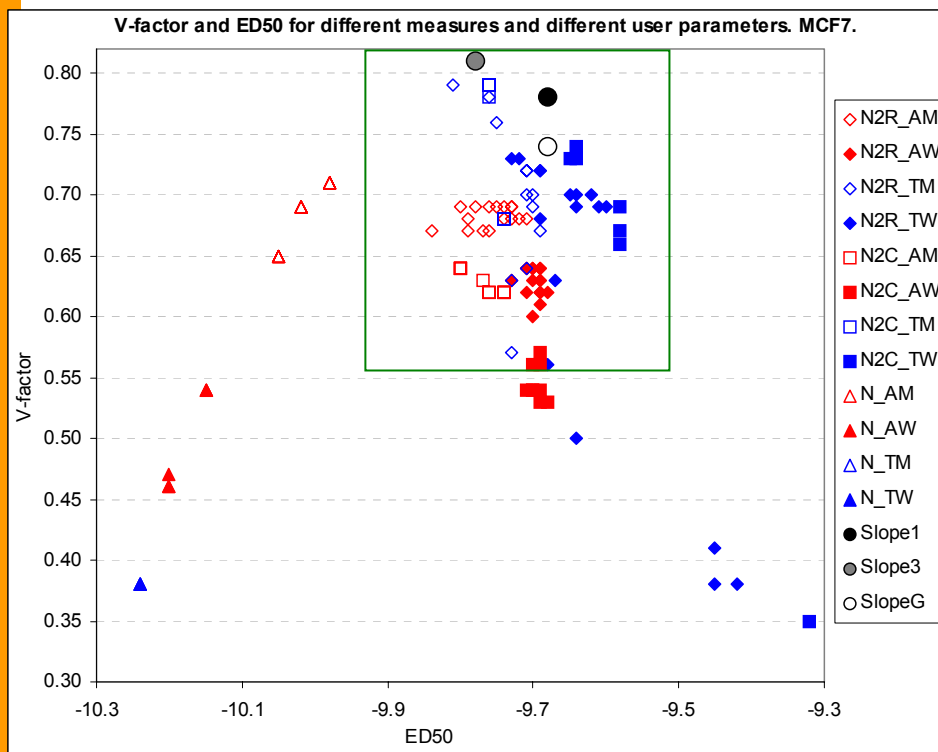
This assay quantifies the Forkhead fusion protein (FKHR-EGFP) accumulated in the nuclei of stably transfected human osteosarcoma cells, U2OS. In proliferating cells FKHR is localized in the cytoplasm. Even without stimulation, Forkhead is constantly moving into the nucleus, but is transported out by export proteins. Upon inhibition of nuclear export, FKHR accumulates in the nucleus. In this assay, export is inhibited by blocking PI3 kinase / PKB signaling by incubating cells for 1 hr with a drug as indicated in the plate map. Nuclear staining: DRAQ. The images were acquired at BioImage on the IN Cell Analyzer 3000. Image size 640\*640 pixels.

# Example of Dose Curve Analysis. X: Dose; CellType=A549; Erosion=1; Gap=2;



Vitro Plate. Concentration is in log scale.

# Evaluation of Algorithms. Parameter Dependency. Vitra Plate.



Concentration is in log scale

Symbol code for measures:

**N2R** (nucleus to ring)

**A** (average intensity) – Red,

**M** (mask-based) wire shape,

**N2C** (nucleus to cytoplasm)

**T** (total intensity) – Blue

**W** (weight-based) solid shape

**N** (nucleus only)

Parameters:

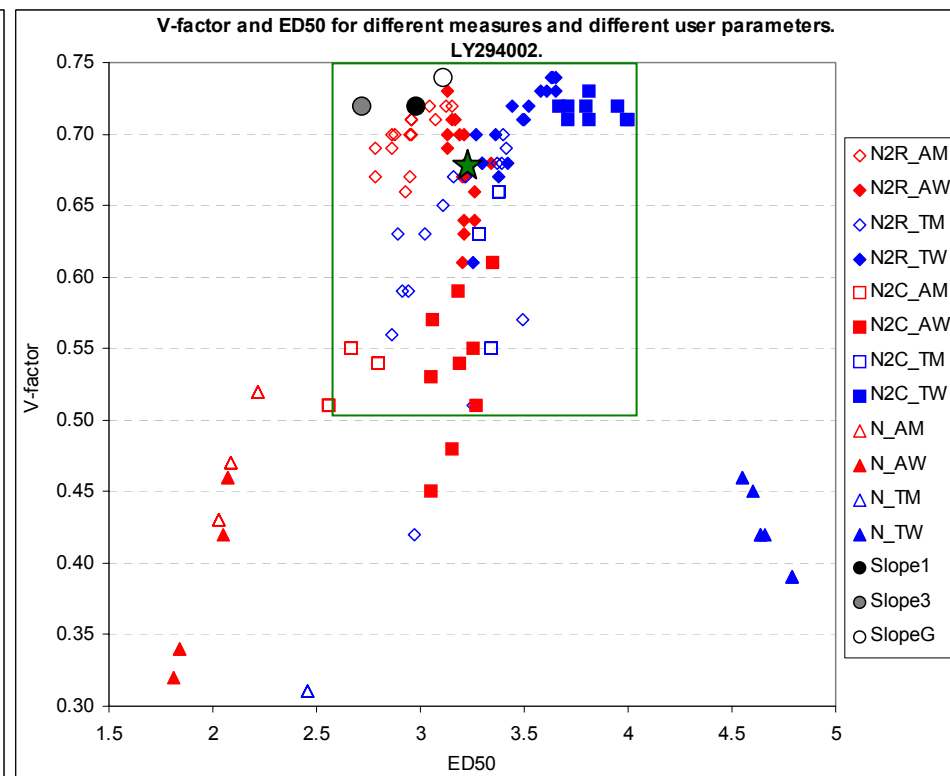
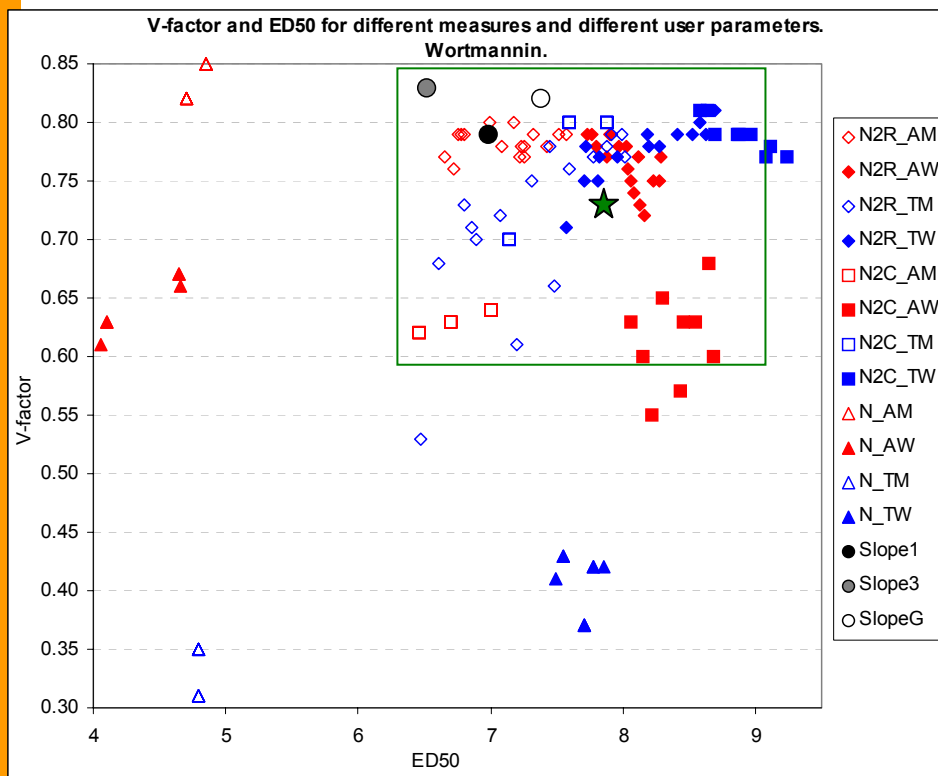
Nuclear erosion = [0 – 2]

Gap = [0 – 4]

Ring = [1 – 8]



# Evaluation of Algorithms. Parameter Dependency. BioImage Plate.



Symbol code for measures:

**N2R** (nucleus to ring)

**N2C** (nucleus to cytoplasm)

**N** (nucleus only)

**A** (average intensity) – Red,

**T** (total intensity) – Blue

**M** (mask-based) wire shape,

**W** (weight-based) solid shape

★ **Log(Nuc/Cyt)**, where Nuc/Cyt is the InCell3000 measure

Parameters:

Nuclear erosion = [0 – 2]

Gap = [0 – 4]

Ring = [1 – 8]

# Evaluation of Algorithms. First conclusions.

Measures based on assessing signal stain in the nucleus alone (N\_ measures) show higher variability and greater inconsistency than other measures, both in v-factor and in ED<sub>50</sub>.

Some measures show greater parameter dependency than others.

Choice of best measures is not consistent between drugs and cell types.

Each measure has maximum quality at different parameter values.

There is a systematic shift between ED<sub>50</sub> of different measures. If the goal is comparison among drugs or cell types, then a single measure should be used.

For further analysis we selected all measures except N\_ and narrowed the parameter range:

Nuclear erosion = [0 – 1]

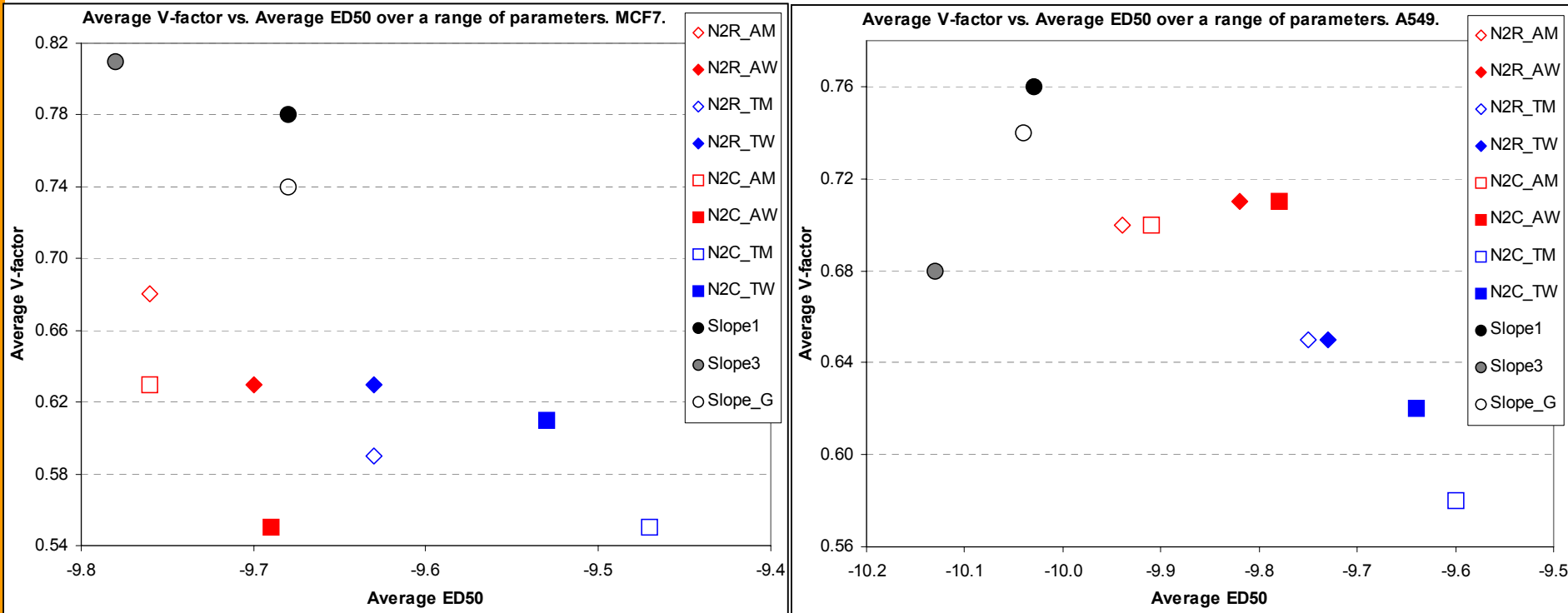
Gap = [0 – 3]

Ring = [1 – 8]

Ero=2 and Gap=4 were excluded because these parameters almost always result in the decline of v-factor.






For realistic assessment of quality of different measures we cannot assume that the parameters would always be optimal, therefore, we averaged V-factors (and ED<sub>50</sub>) for each measure over the parameter range determined above (except for slope measures, which do not have parameters).

# Evaluation of Algorithms. Quality Averaged over Parameter Range. Vitra Plate.



Concentration is in log scale

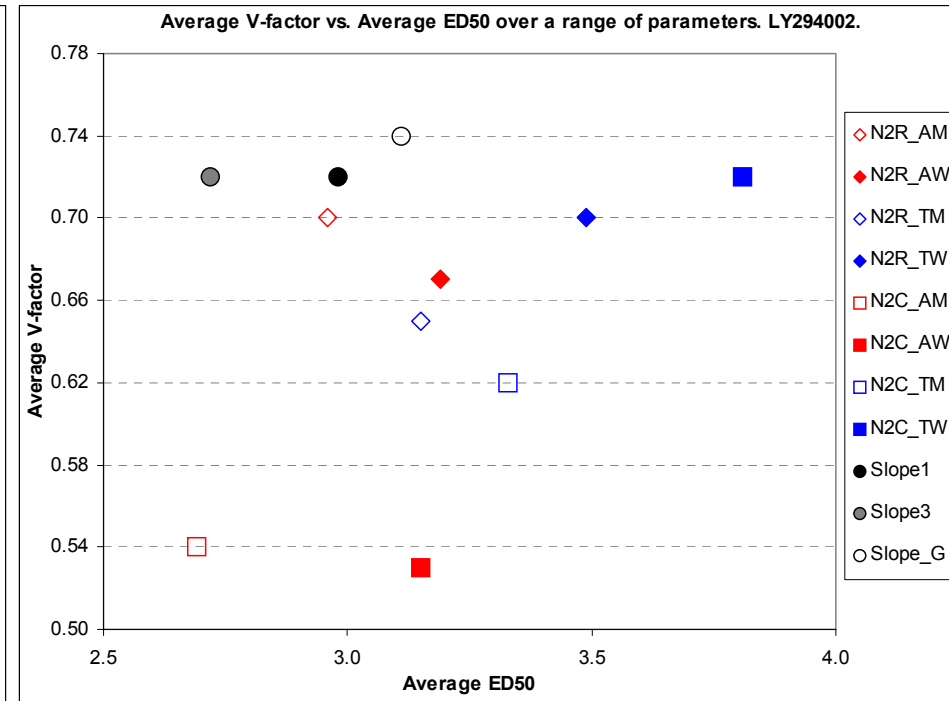
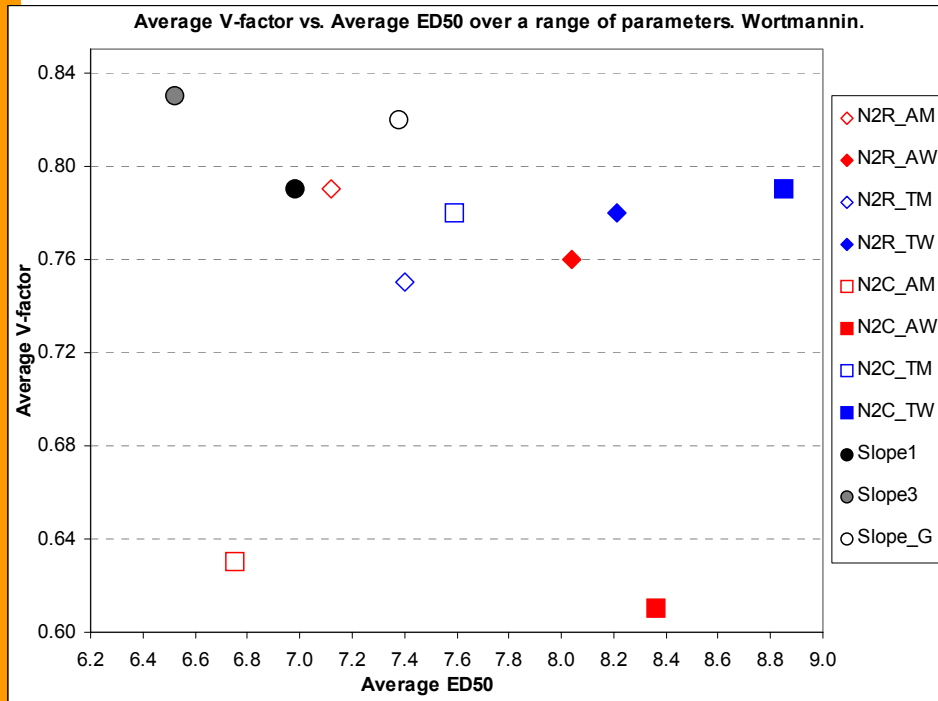
Symbol code for measures:

**N2R** (nucleus to ring) 
**N2C** (nucleus to cytoplasm) 
**N** (nucleus only)   
**A** (average intensity) – Red, **T** (total intensity) – Blue  
**M** (mask-based)  wire shape, **W** (weight-based)  solid shape

Averaged over Parameter range:

Nuclear erosion = [0 – 1]  
 Gap = [0 – 3]  
 Ring = [1 – 8]

# Evaluation of Algorithms. Quality Averaged over Parameter Range. BioImage Plate.



Symbol code for measures:

**N2R** (nucleus to ring)

**A** (average intensity) – Red,

**M** (mask-based) wire shape,

**N2C** (nucleus to cytoplasm)

**T** (total intensity) – Blue

**W** (weight-based) solid shape

**N** (nucleus only)

Averaged over Parameter range:

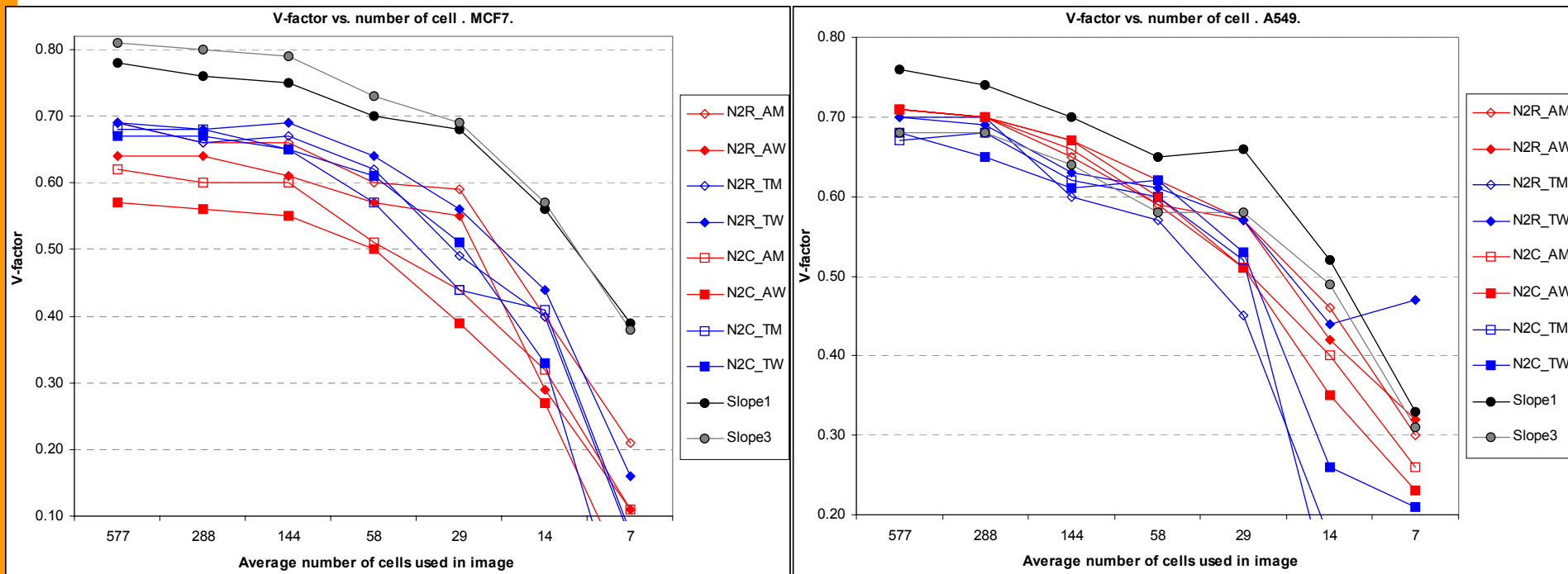
Nuclear erosion = [0 – 1]

Gap = [0 – 3]





Ring = [1 – 8]

# Analysis of Dependencies.

# Dependency of V-factor on Cell Number. Vitra Plate.



Symbol code for measures:

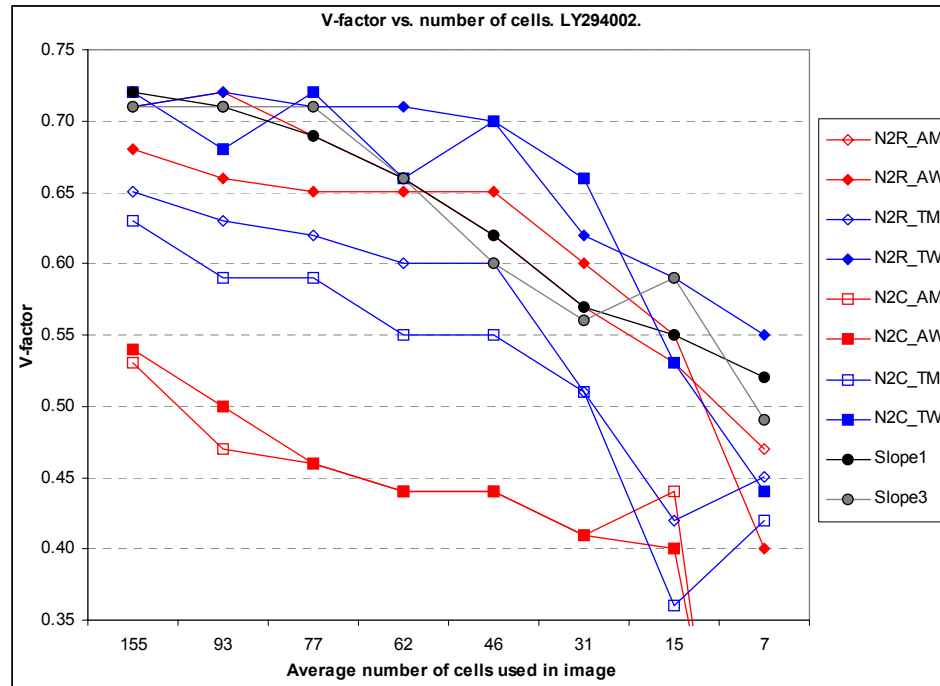
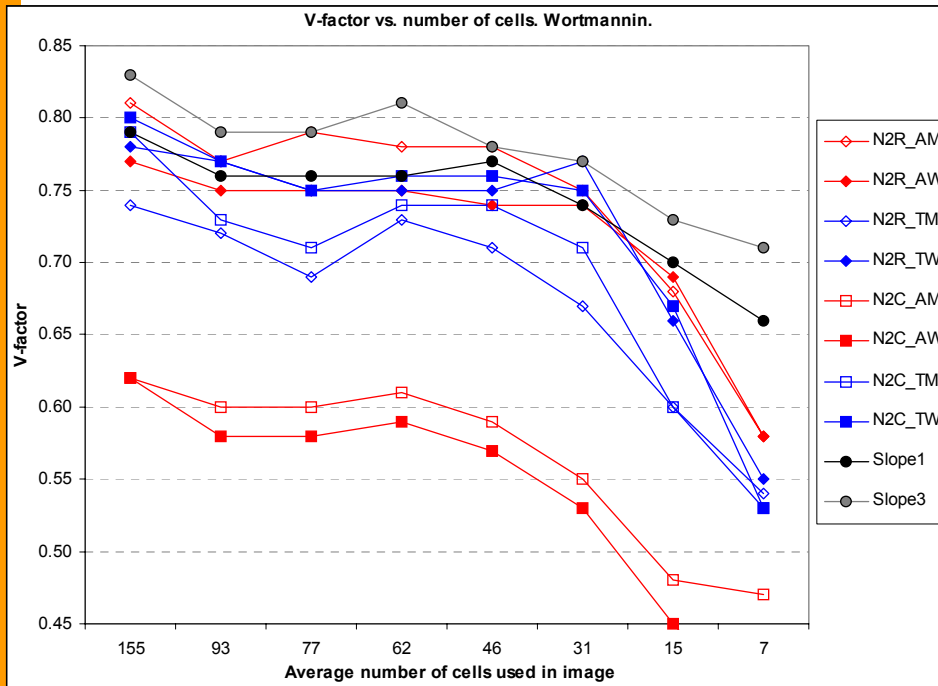
**N2R** (nucleus to ring)  **N2C** (nucleus to cytoplasm)   
**A** (average intensity) – Red, **T** (total intensity) – Blue  
**M** (mask-based)  wire shape, **W** (weight-based)  solid shape

Parameters:

Nuclear erosion = 1  
 Gap = 2  
 Ring = 4

To study the cell number dependency we selected random subsets of cells from the whole image to eliminate possible positional dependency.

# Dependency of V-factor on Cell number. BioImage Plate.



Symbol code for measures:

**N2R** (nucleus to ring)

**A** (average intensity) – Red,

**M** (mask-based) wire shape,

**N2C** (nucleus to cytoplasm)

**T** (total intensity) – Blue

**W** (weight-based) solid shape

Parameters:

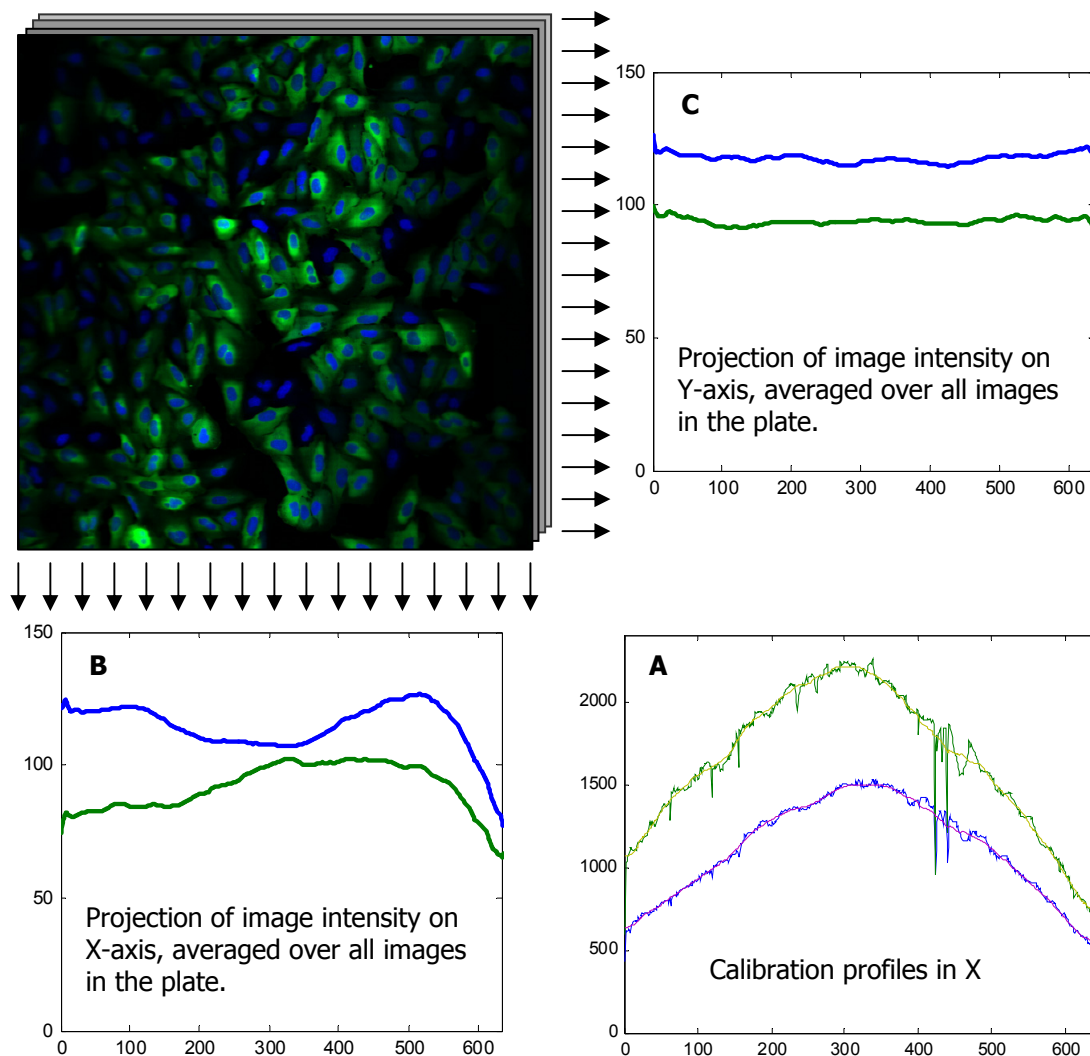
Nuclear erosion = 1

Gap = 2

Ring = 4



# Is Uniformity of Illumination a Factor in CNT Algorithms? 1.



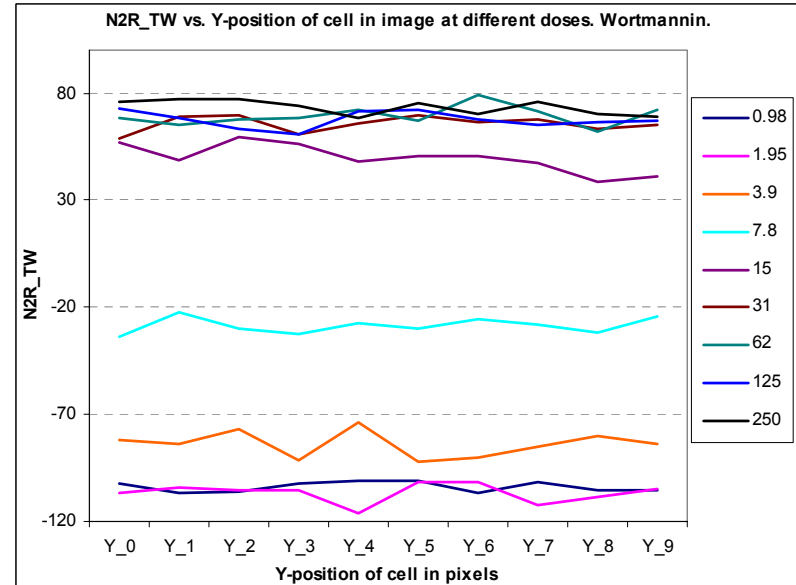
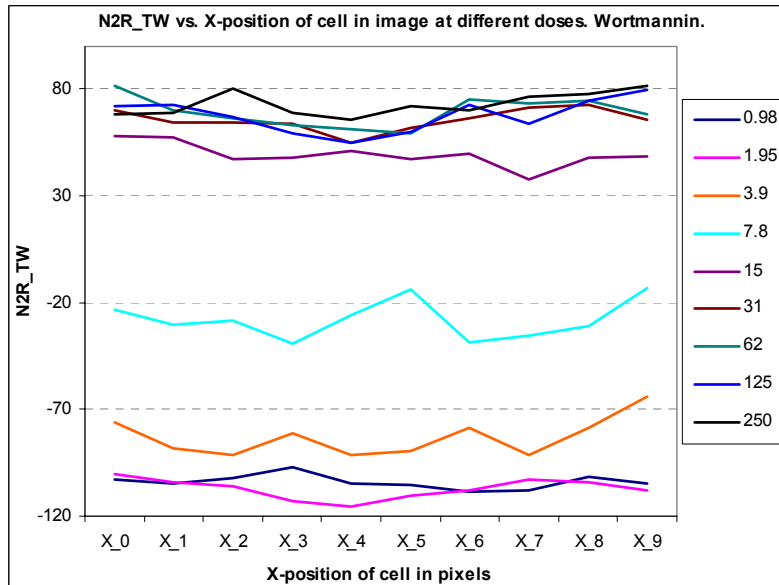
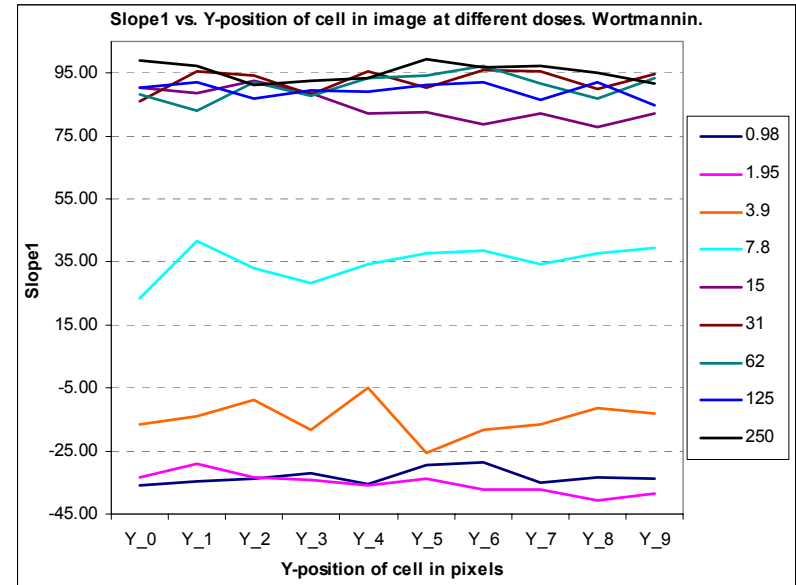
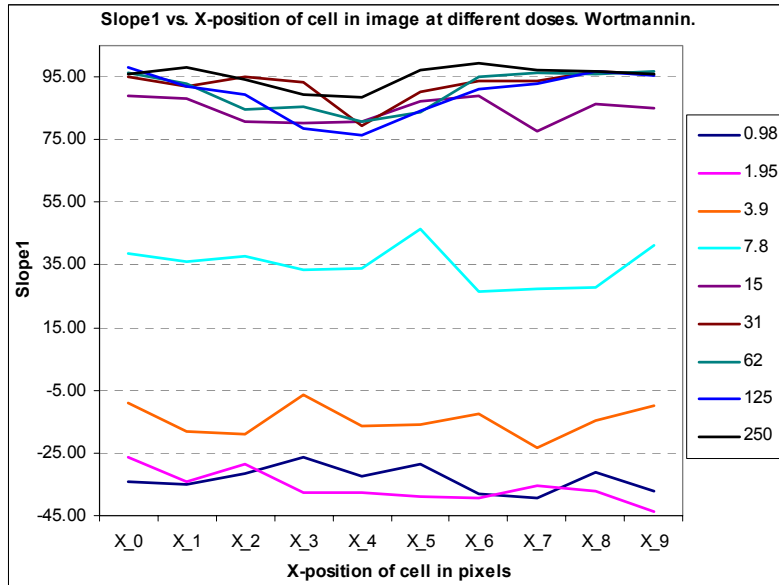
Images acquired on InCell3000 provide a convenient model to study the dependency of cellular measures on non-uniformity of illumination. Y-direction, which corresponds to the stage movement has uniform brightness, but X-direction, which corresponds to the scan of the laser beam has variable brightness. Calibration curves provided with the images are shown in Figure A.

Actual brightness in the two channels averaged over the entire plate to eliminate random errors is shown in figures B (X-projection) and C (Y-projection). The plots show clearly uniform intensity in Y and non-uniform intensity in X.

To assess if the brightness non-uniformity affects the cellular measures produced by the presented family of CNT algorithms, we plotted the measures against X-coordinates of cells and against Y-coordinates of cells. Two of the measures are shown below. The plots show that there is no dependency on either X or Y position of cells. Other measures, which are not shown, exhibit similar behavior.

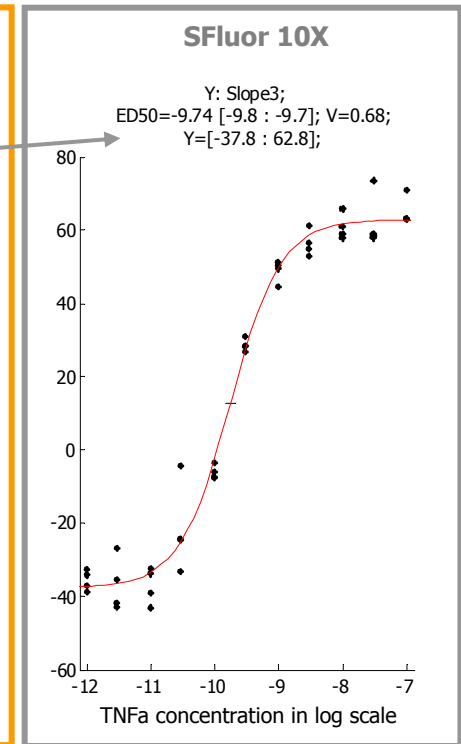
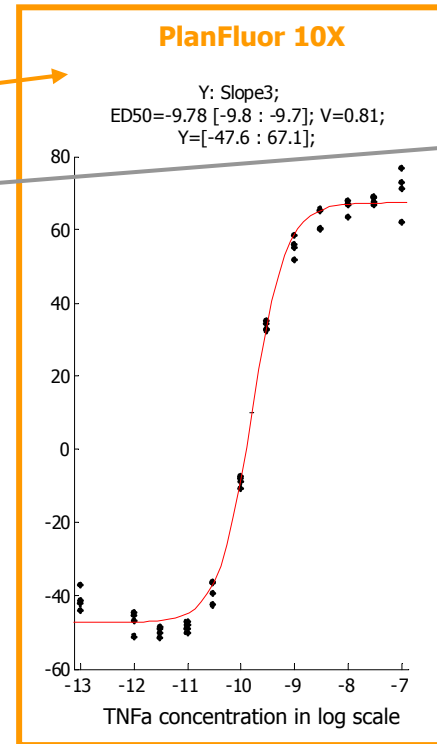
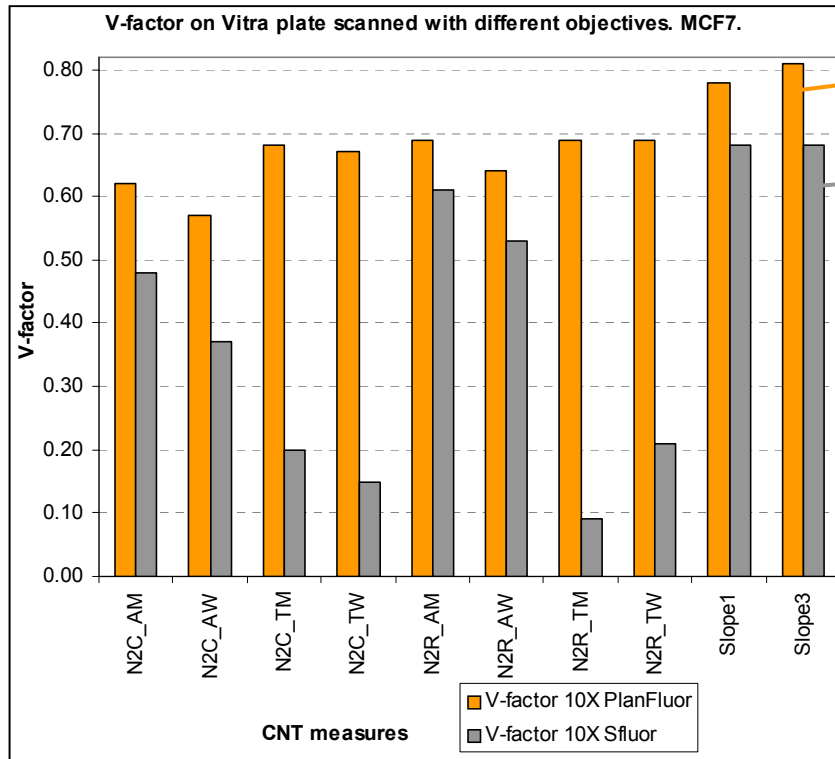
This allows us to conclude that variation in brightness alone does not affect cellular measures if the algorithm is specifically designed to compensate for it.

# Is Uniformity of Illumination a Factor in CNT Algorithms? 2.



# Dependency of CNT Measures on the Depth of Field in a Non-confocal System. 1.

The Vitra plate was scanned two times: with a PlanFluor 10X 0.3NA objective and with a SFluor 10X 0.5NA objective (Nikon). The images were processed with the same parameters: Nuclear erosion=1, Gap=2, Ring=4. V-factor shows that PlanFluor always gives better results. Note that measures based on total amount of fluorescence (as opposed to average) suffer the most.

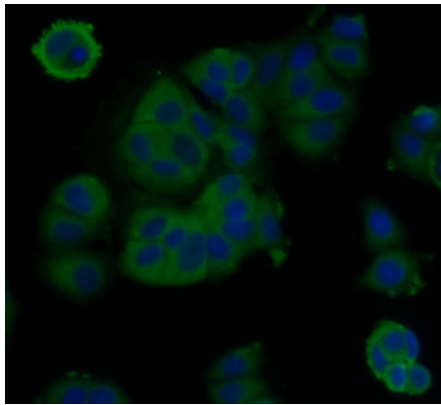


Our hypothesis is that this result is due to the relative thickness of cells and the depth of field of the two objectives. The depth of field of 0.3NA PlanFluor 10X is about 8.5  $\mu\text{m}$ , and the depth of field of 0.5NA SFluor 10X is about 3.6  $\mu\text{m}$ . The thickness of attached cells is probably between these two values. If the cell is thicker than the depth of field of the objective it will appear fuzzy. This will make negative cells look less negative and reduce the dynamic range of the measurements. The following illustration supports this hypothesis.

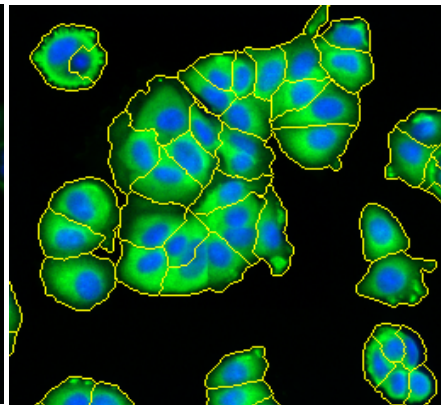
Analysis of the minimal and maximal values of the fitted logistic curves indeed shows reduced dynamic range with the SFluor objective. This does not explain, in full, the decrease in V-factor. The variation among replica wells also was higher with the SFluor objective. This may be attributed to focusing artifacts due to shallow depth of field.

# Dependency of CNT Measures on the Depth of Field in a Non-confocal System. 2.

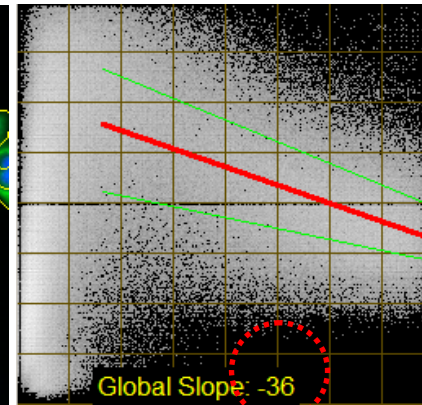
**PlanFluor10X, 0.3NA, Depth of field = 8.5  $\mu\text{m}$**



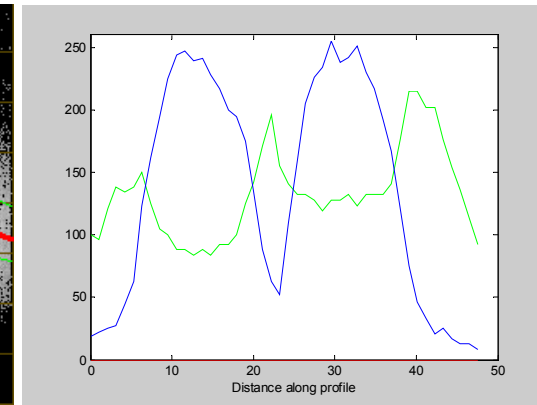
Original



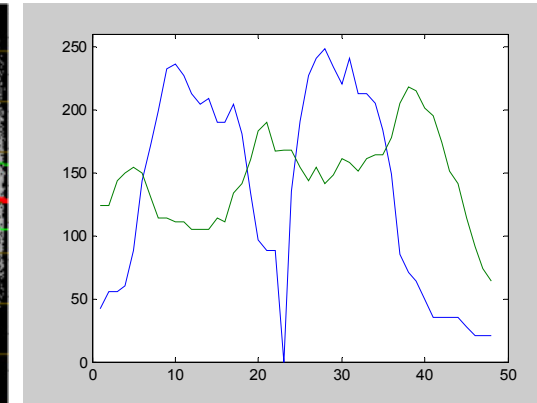
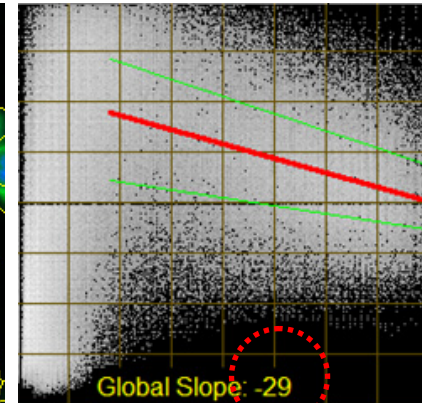
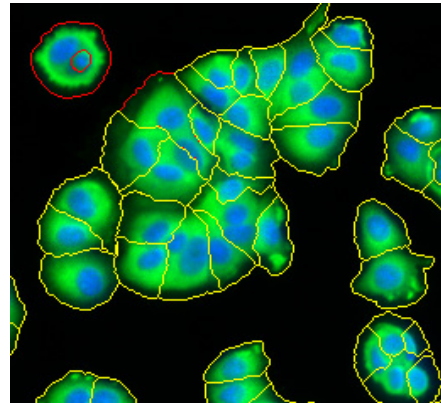
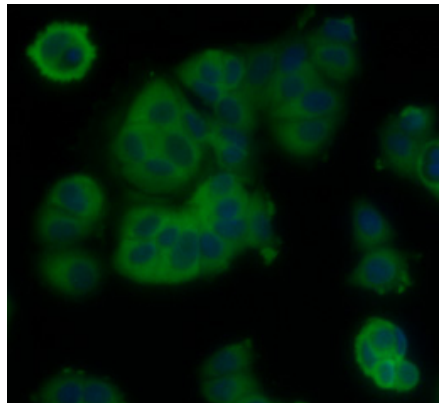
Normalized



2D distribution



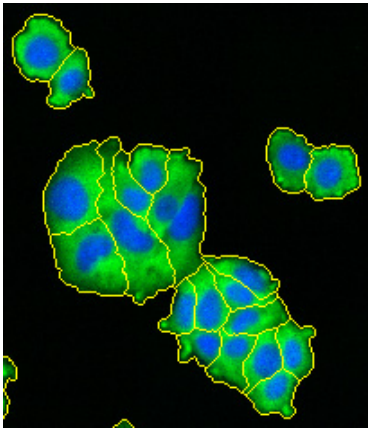
Profiles



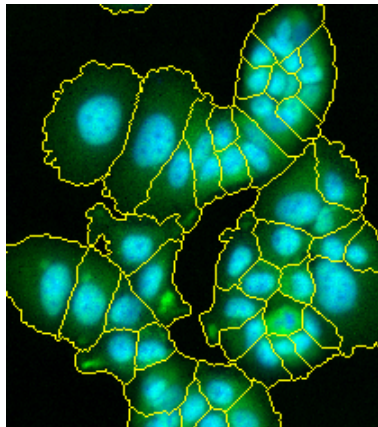
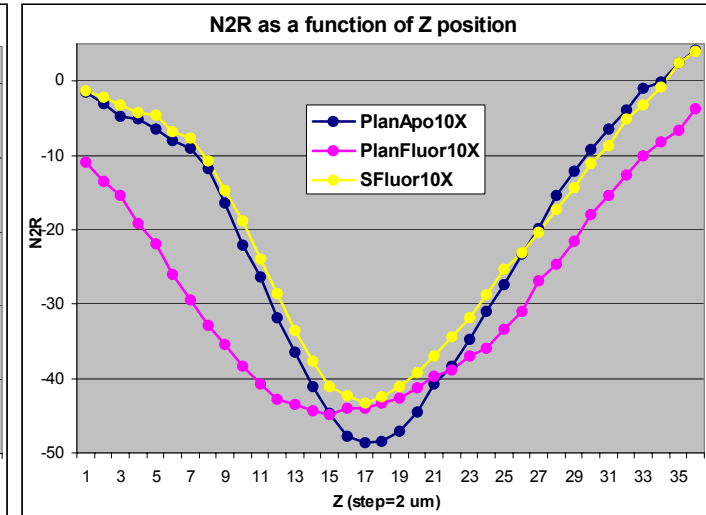
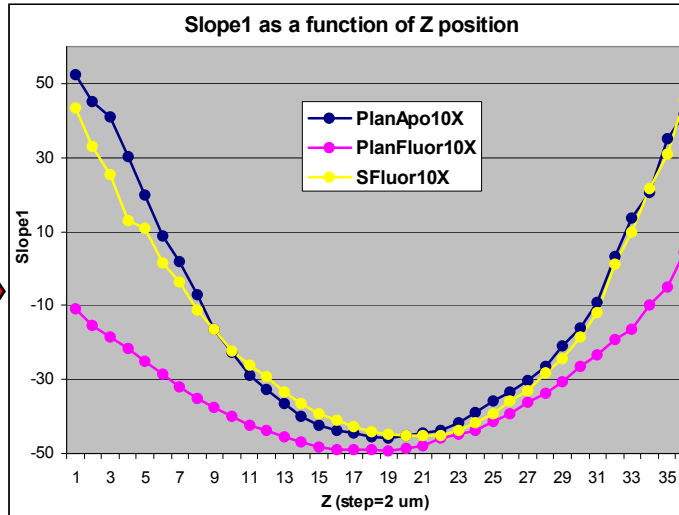
**SFluor10X, 0.5NA, Depth of field = 3.6  $\mu\text{m}$**

# Dependency of CNT Measures on Focus Position.

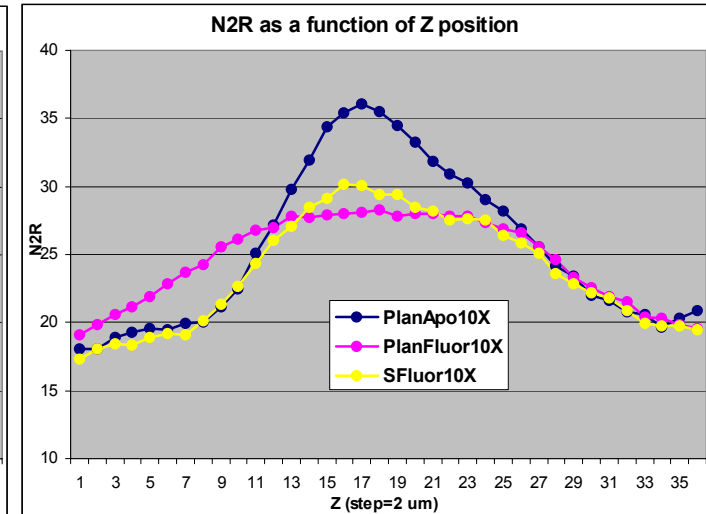
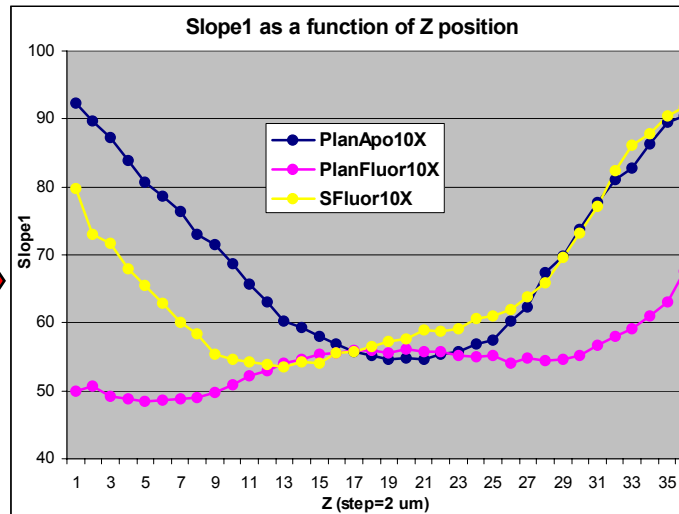
Images at 10X



Negative



Positive



- CNT measures are almost identical within 10-15  $\mu\text{m}$  from the best focus position in 10X (less for PlanApo with N2R\_AM).
- Objectives with lower NA are more tolerant to focusing inaccuracy.
- Slope1 and N2R\_AM show similar dependency on Z-position on negative cells and opposite dependency on positive cells.
- For Slope1 PlanFluor gives the greatest range (marginally); for N2R\_AM PlanApo gives the greatest range.
- The importance of this dependency is not only in setting focusing requirements, but also in establishing the need for flat field.

# Conclusions and Recommendations.

## *Algorithms and measures:*

1. Measures based on assessing signal stain in nucleus alone show much higher variability than other measures, both in v-factor and in ED50.
2. There is a systematic shift between ED50 of different measures. If the goal is comparing drugs or cell types, then a single measure should be used for the whole experiment. This points to the need for a uniformly best measure across experimental conditions.
3. Different measures have maximum quality at different parameter values. For a realistic comparison, the quality of different measures must be averaged over a reasonable parameter range. Therefore, measures which do not depend on parameters, have a big advantage.
4. Conceptually similar measures implemented by different vendors can show significant difference in quality.

## *Choosing an instrument:*

1. Illumination non-uniformity alone does not play an important role if the algorithms can compensate for it.
2. Flatness of field and accuracy of focusing are critical and cannot be easily compensated for algorithmically.

## *Experiment design:*

1. Objectives in a wide-field system should be chosen based on their depth of field and cell thickness.
2. Focusing accuracy of the system must be evaluated and focusing strategy established to provide for small error (e.g.,  $<10\mu\text{m}$  in 10X).
3. Required number of cells can be established by analyzing a fraction of images and selecting a number above which quality does not significantly improve with cell number.



# 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
8. US Pat. 6,671,624 "Machine readable storage media for detecting distribution of macromolecules between nucleus and cytoplasm in cells".
9. Ilya Ravkin, Vladimir Temov, "Image analysis without segmentation - A new method to measure cytoplasm to nucleus translocation", poster, High Content Analysis Conference, San Francisco 2005 .