Cytometry PART A
Journal of the International Society for Advancement of Cytometry

Quantifying Colocalization by Correlation: The Pearson Correlation Coefficient is Superior to the Mander's Overlap Coefficient

Jeremy Adler, Ingela Parmryd*

The Wenner-Gren Institute, Stockholm University, 106 91 Stockholm, Sweden

Received 19 October 2009; Revision Received 4 March 2010; Accepted 5 March 2010

Additional Supporting Information may be found in the online version of this article.

Grant sponsor: Carl Trygger's Foundation.

*Correspondence to: Ingela Parmryd, The Wenner-Gren Institute, Stockholm University, 106 91 Stockholm, Sweden.

Email: ingela.parmryd@wgi.su.se

Published online 30 March 2010 in Wiley InterScience (www.interscience.wiley.com)

DOI: 10.1002/cyto.a.20896

© 2010 International Society for Advancement of Cytometry

• Abstract

The Pearson correlation coefficient (PCC) and the Mander's overlap coefficient (MOC) are used to quantify the degree of colocalization between fluorophores. The MOC was introduced to overcome perceived problems with the PCC. The two coefficients are mathematically similar, differing in the use of either the absolute intensities (MOC) or of the deviation from the mean (PCC). A range of correlated datasets, which extend to the limits of the PCC, only evoked a limited response from the MOC. The PCC is unaffected by changes to the offset while the MOC increases when the offset is positive. Both coefficients are independent of gain. The MOC is a confusing hybrid measurement, that combines correlation with a heavily weighted form of co-occurrence, favors high intensity combinations, downplays combinations in which either or both intensities are low and ignores blank pixels. The PCC only measures correlation. A surprising finding was that the addition of a second uncorrelated population can substantially increase the measured correlation, demonstrating the importance of excluding background pixels. Overall, since the MOC is unresponsive to substantial changes in the data and is hard to interpret, it is neither an alternative to nor a useful substitute for the PCC. The MOC is not suitable for making measurements of colocalization either by correlation or co-occurrence. © 2010 International Society for Advancement of Cytometry

• Key terms

correlation; colocalization; co-occurrence; Mander's overlap coefficient; Pearson correlation coefficient

Introduction

The quantification of colocalization between two fluorescence channels broadly divides into two categories: (1) methods that simply consider the presence of both fluorophores in individual pixels, which we call co-occurrence and (2) those that examine the relationship between the intensities, correlation. The two categories are different and full co-occurrence is compatible with zero correlation, while a high correlation can be found among the co-occurring pixels even when co-occurrence is rare.

The co-occurrence of fluorophores may simply reflect physicochemical similarities between two fluorescent molecules or antigens: hydrophobic molecules will partition into membranes, hydrophilic molecules to the cytoplasm while amphiphilic molecules are mostly found at interfaces. Co-occurrence can be quantified by expressing the number of co-occurring pixels as a fraction of the total number or by using the M_1 and M_2 coefficients which, separately for each fluorophore, record the fraction of the total fluorescence that co-occurs (1).

A correlation between the intensities could reflect a direct molecular interaction or an indirect interaction, with a third molecule or with subdomains of a cellular compartment. The variability of the fluorescence and therefore the potential for correlation arises from inhomogeneities within a domain. A correlation between two fluorophores is likely to be of greater biological significance than co-occurrence, though any change

ISAC International Society for Advancement of Cytometry

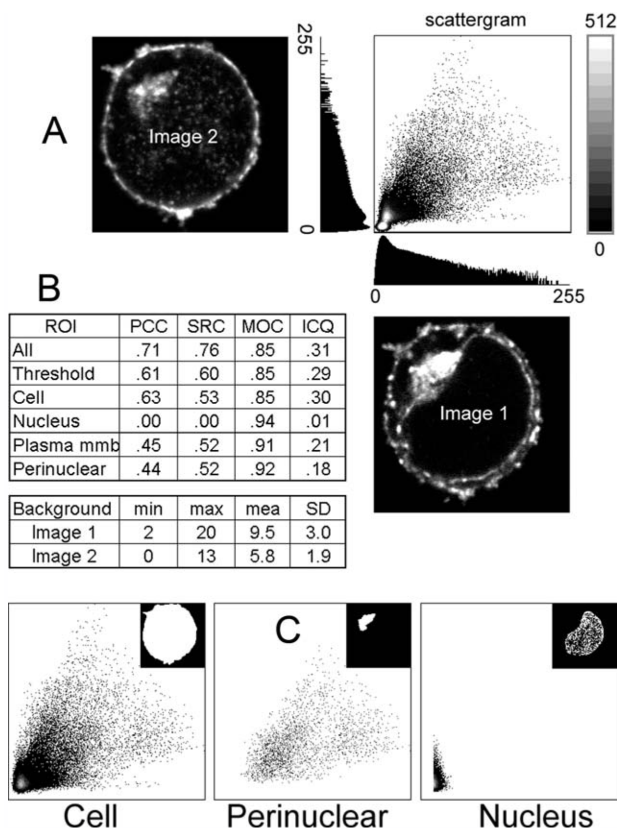
in colocalization that can be related to an experimental intervention is of interest. There is a need to measure colocalization and the accuracy with which measurements can be made sets the limits for an observable physiological response.

Two measures of correlation appear in most software, the Pearson correlation coefficient (PCC) and the Mander's overlap coefficient (MOC) (1). The PCC is a well-established measure of correlation, originating with Galton in the late 19th century (2), but named after a colleague, and has range of +1 (perfect correlation) to −1 (perfect but negative correlation) with 0 denoting the absence of a relationship. Its application to the measurement of colocalization between fluorophores is relatively recent (3). The MOC lacks the pedigree of the PCC and was created to meet perceived deficiencies in the PCC, principally that the PCC "is not sensitive to differences in signal intensity between the components of an image caused by different labeling with fluorochromes, photobleaching or different settings of amplifiers" and "the negative values of the correlation coefficient (PCC) are difficult to interpret when the degree of overlap is the quantity to be measured" (1), much repeated claims (4,5).

The two measures are mathematically similar, differing only in the use of either the absolute intensities (MOC) or the departure from the mean (PCC) in both the numerator and the denominator. The numerator is the sum of the products of the two intensities (which we will for convenience refer to as red and green) in homologous pixels and the denominator computes the maximal product, corresponding to perfect colocalization. The method works because the numerator is maximized when the relative intensities of the two fluorophores coincide: high with high and low with low, while combination of high with low reduces the sum of their products. The denominator acts to limit the range of the coefficients: 0 to +1 for the MOC and −1 to +1 for the PCC.

Two other measures of correlation have been used to quantify colocalization, the intensity correlation quotient (ICQ) (6,7) and the Spearman rank correlation (SRC) (8,9), both derived from the PCC. The SRC is a well-established statistical test and is simply the PCC applied to ranked data: intensities are replaced by the order in which they occur. The ICQ goes a step further than the SRC and only considers the sign: whether each of the two intensities are above or below their respective mean intensity. The numerator is the number of pairs of intensities that have a common sign, either minus and minus or plus and plus. The denominator is just the number of intensity pairs. This would give the ICQ method a range from 0 to 1 but, to align negative correlations with a negative coefficient, 0.5 is subtracted from the calculated values, creating a range from −0.5 to +0.5 (6).

It is hard to visually assess the degree of colocalization from a pair of images even when they are overlayed. A more informative alternative is to display the intensities of the pairs of homologous pixels in a scattergram (Fig. 1A). Each axis covers the intensity range of one of the fluorophores and the scattergram shows the frequency of occurrence of each pair of intensities, which reveals any correlation between the fluorophores.



ROI	PCC	SRC	MOC	ICQ
All	.71	.76	.85	.31
Threshold	.61	.60	.85	.29
Cell	.63	.53	.85	.30
Nucleus	.00	.00	.94	.01
Plasma mmb	.45	.52	.91	.21
Perinuclear	.44	.52	.92	.18

Background	min	max	mea	SD
Image 1	2	20	9.5	3.0
Image 2	0	13	5.8	1.9

Figure 1. Colocalization, scattergrams, and regions of interest (ROI). **A**: Two images and their log frequency distribution histograms. For each pixel in the pair of fluorescent images, the two intensities are used as the coordinates of an entry in the scatterplot. This shows the relationship between the two fluorophores. Pixels from the whole area, including areas outside the cell are included. A grayscale look up table shows the frequency of occurrence of each pairs of intensities. Note that a white background has been used for the scattergrams and that the fluorescent images have been contrast stretched for display purposes, but that the histograms and scattergrams show the original distribution. **B**: Correlation measurements and the background intensity. **C**: Scattergrams for different ROIs (inserted top left), showing which pixels were included in the analysis: the nuclear ROI is speckled because an intensity threshold (mean plus twice the standard deviation) was also employed.

Despite the appearance of several reviews on colocalization (4,10–12) and related literature, it is surprising that a critical comparison of the methods used to measure colocalization using correlation has not been undertaken. This we seek to remedy.

Methods

Measurement of Colocalization

PCC (r) is given as follows:

$$r = \frac{\sum (R_i - R_{av}).(G_i - G_{av})}{\sqrt{\sum (R_i - R_{av})^2 . \sum (G_i - G_{av})^2}} \quad (1)$$

The Spearman rank correlation is the same as the PCC except that the original intensities are replaced by their rank.

MOC (R) is given as follows:

$$R = \frac{\sum (R_i).(G_i)}{\sqrt{\sum (R_i)^2 . \sum (G_i)^2}} \quad (2)$$

where R_i is the intensity of the first (red) fluorophore in individual pixels and R_{av} the arithmetic mean, whereas G_i and G_{av} are the corresponding intensities for the second (green) fluorophore in the same pixels.

ICQ is given as follows:

$$\text{ICQ} = \frac{\sum (R_i > R_{av}) = (G_i > G_{av})}{N} - 0.5 \quad (3)$$

where $R_i > R_{av}$ and $G_i > G_{av}$ are the sign (plus or minus) of the difference from the respective mean values, whereas "=" indicates that the signs are the same. N is the number of pixels.

Simulated Images

Simulations were principally performed as described previously (8).

(a) Varying the correlation. A pair of uncorrelated images, referred to as red and green, with similar population distributions, were combined to generate a new red image whose intensities (R_{rg}) had varying degrees of correlation with the green image. This was achieved by altering each red pixel, replacing a fraction (the copy fraction, C_f) of the original intensity (R_i) with the same fraction copied from the homologous pixel in the green image (G_i). G_{av} and R_{av} are the mean intensities in the green and red images.

$$R_{rg} = R_{av} + (R_i - R_{av})(1 - C_f) + (G_i - G_{av})C_f \quad (4)$$

As the copy fraction progressively changes from 0 to +1, a wide variety of correlated image pairs are generated (Fig. 2A). Negative copy fractions produced inverse relationships and were generated by subtracting, rather than adding, the copied intensity (Fig. 2A).

$$R_{rg} = R_{av} + (R_i - R_{av})(1 - C_f) - (G_i - G_{av})C_f \quad (5)$$

(b) Changes in offset were simulated by the global addition of a constant to one or both images. Alterations in gain were achieved by multiplication with a constant.

(c) Image pairs with two different relationships were created by altering a subset of the original population. When the gain of the subset in one image was altered, the new images then contained two correlated populations. Alternatively, when the subset of both images was replaced by uncorrelated values the pair of images then had both a correlated and an uncorrelated component.

Weightings

The contribution made to the PCC and MOC by different combinations of intensities was calculated separately for the numerator and denominator over an 8-bit range. The calculations for the PCC and ICQ assume that the mean for both the red and green population was 128. For the denominator of the MOC, it was assumed that for each intensity pair there was a second pair with the red and green intensities reversed, making the sum of intensities the same for red and green. In a calculation based on a single pair of values, the denominator and numerator would be equal.

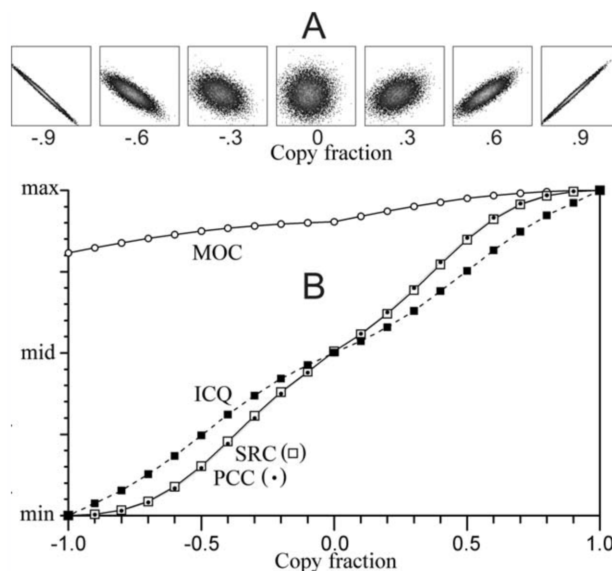


Figure 2. Four correlation coefficients. **A**: The scattergrams show a range of correlations. They were generated from two uncorrelated images, by replacing a fraction of the intensity of one of the images with the same fraction copied (the copy fraction) from the other image. When the copy fraction was zero the images remain unaltered. **B**: Four correlation coefficients applied to the data shown in (**A**). Since the four coefficients make use of three different numerical ranges, the scale covers the full range of each coefficient.

Image Analysis and Processing

All analysis and image processing was performed using software that runs on a Semper6w kernel (Synoptics, Cambridge, UK), which is included in the Supporting Information. Figures were prepared using Adobe Photoshop and graphs were created using PSI-Plot (Poly Software International, NY).

Results

Correlation Analysis on a Cell

Correlation measurements were made on a single cell (Fig. 1A) using four coefficients: the PCC, SRC, MOC, and ICQ. Measurements were made on the whole image, including pixels outside the cell, the cell alone and subregions within the cell (Fig. 1B). Analysis of the whole image produced high values for each coefficient, which were reduced when the analysis was restricted to the cell or when background was removed by thresholding, using the intensity range found in an area outside the cell (Fig. 1B). Analysis of subregions (Fig. 1C) show different patterns within the cell, with the fluorescence within the nucleus having no correlation, except when measured using the MOC.

It is noteworthy that the MOC is almost unaffected by the choice of ROIs and that the removal of background pixels from the analysis reduced the measured correlation using the PCC, SRS, and ICQ. It is also surprising that the correlation for the cell was higher than that of any subregion.

Response Range

The coefficients were tested using a range of paired images that incrementally changed from a completely positive correlation to a fully negative correlation (Fig. 2A). The original image pair was uncorrelated, a copy fraction of zero, with a PCC that was 0 and a MOC that was 0.9. As the copy fraction changed from −1 to +1, the image pair shifted from being negatively correlated to positively correlated, producing a change in the PCC, SRC, and ICQ over their full range, a change that was symmetrical around a copy fraction of 0 (Fig. 2B). The corresponding change in the MOC was from 0.8 to 1.0, a fraction of its nominal range. The symmetry shown by the PCC was not seen with the MOC, which declined as the copy fraction fell from +1 to 0 and then continued to fall as the copy fraction became negative.

Offset

The PCC is unaffected by an additive offset applied to one or both images, unless the offset compresses the intensity range, while an offset can substantially alter the MOC (Fig. 3). When the offset was increased in both images, the MOC increased and moved progressively toward the upper limit of the measurement range, to such an extent that changes to the copy fraction from −1 to +1 had minimal effect on the MOC (data not shown). An offset applied to one of the pair of images can produce an increase, a decrease or even a small increase followed by a decrease in the MOC (Fig. 3). The SRC and the ICQ are, like the PCC, unaffected by gain (data not shown).

Gain

Alterations to the gain change the appearance of scattergrams (Fig. 4B), but neither the MOC nor the PCC are affected by multiplicative changes applied to the intensities in one image of the pair, except at very low gains: <0.05 (Fig. 4A). The presence of noise, with a uniform distribution that is unaffected by the gain, increased the sensitivity to low gain, with the PCC being altered appreciably more than the MOC.

The decrease in the MOC and PCC at low gains arises from the use of 8-bit integer numbers and is a consequence of compressing the signal range. This problem is minimized when 10- or 12-bit numbers are used. In practice, we adjust the image acquisition settings to make good use of the detector response range and very low signals are easily avoided.

Multiple Relationships: Adding Uncorrelated Pixels

The effects of replacing some of the original pixels with pixels that are uncorrelated and have low intensities, equivalent to including background pixels, on the PCC is noticeable, while the limited response of the MOC is not appreciably altered (Fig. 5A). The PCC always became more positive, a change that was especially marked with negative correlations (a negative copy fraction): a perfect negative PCC of −1.0 became −0.1 when just 25% of the pixels were uncorrelated. The positive shift in the PCC increases as the fraction of pixels with background values increases from 0 to 25%. Similarly, when the copy fraction is zero and the original pair of images are uncorrelated, the insertion of background pixels, that are

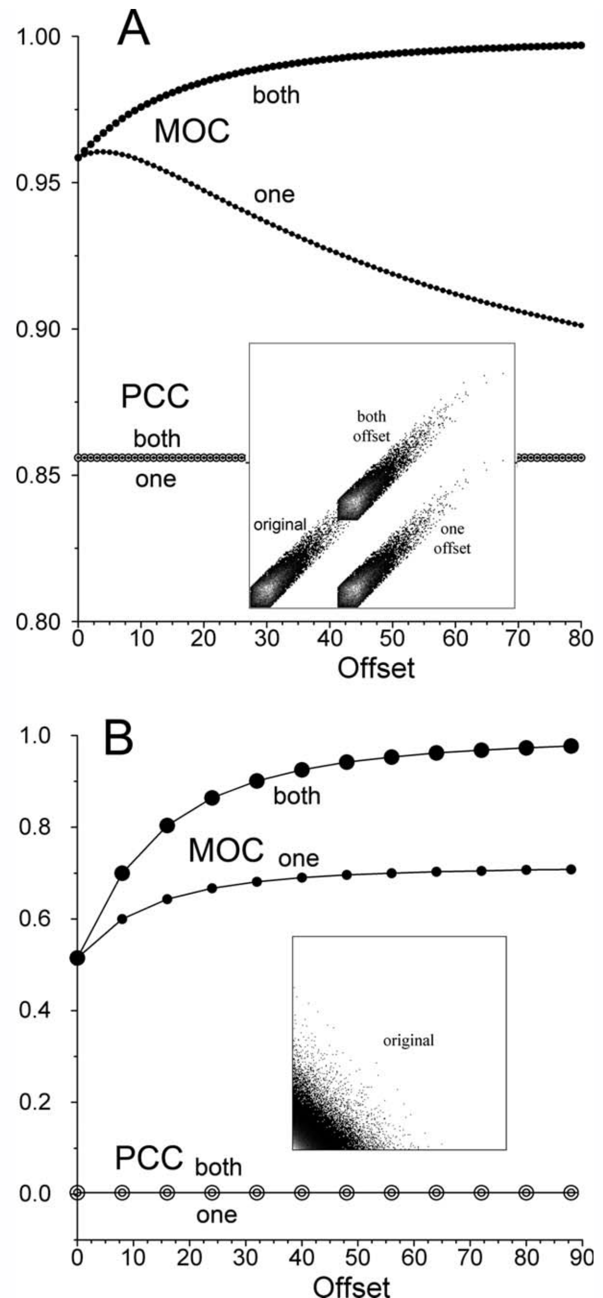


Figure 3. Positive offsets were applied to either one or both images in a correlated pair (**A**). The overlayed scattergram shows the original distribution (bottom left) and the distribution after an offset was applied to a single image or to both images. A poorly correlated pair of images (**B**) with a scattergram. After the application of the range of offsets all intensities remained within the 8-bit range.

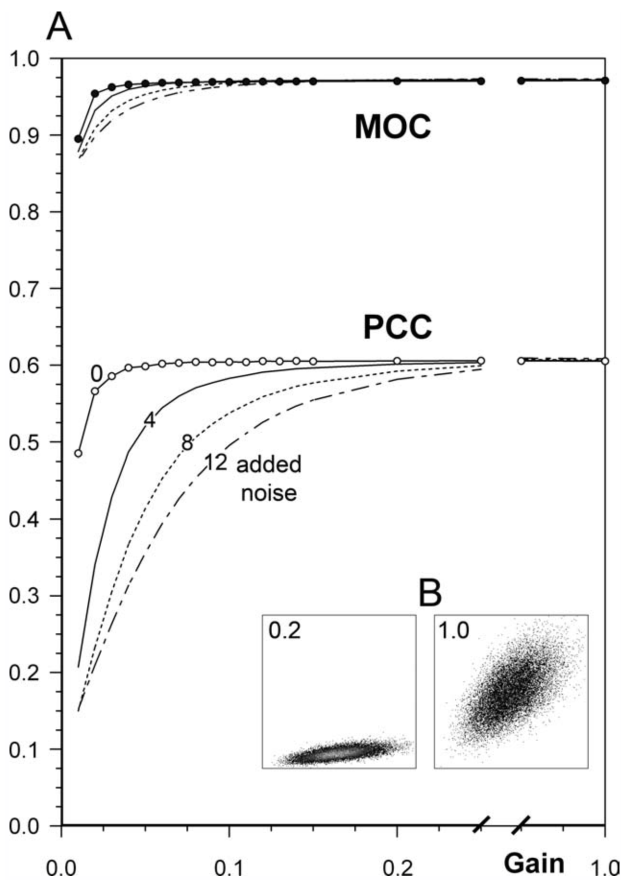


Figure 4. Gain. The intensities of one image of a colocalized pair was progressively altered and the response of the PCC and the MOC followed (**A**). Note that the *x*-axis has a break The effect of changes in the gain is illustrated by two scattergrams (**B**). A gain of 1.0 kept the data at the limits of the 8-bit range. In addition to varying the gain the effect of including noise that is independent of the gain is shown. The noise was added after any alteration in the gain and had a uniform distribution with different widths, which are marked on the PCC plots. For reasons of clarity, the noise levels are not marked on the MOC plots, but correspond to those of the PCC. The original images had a mean intensity of 128 and a Gaussian distribution with a width of 28.

themselves uncorrelated, produces an overall positive correlation: a combination of two uncorrelated populations produces a positive PCC (Fig. 5A).

When the replacement population of pixels has a mean intensity equal to that of the original population, rather than a lower intensity, the pattern of change is reversed (Fig. 5B), the PCC remains unchanged while the MOC increases, most noticeably when the correlation (negative copy fractions) is negative.

Multiple Relationships: Combining Two Correlated Subsets

Combining two correlated sets of pixels into a single population reduces the measured PCC and MOC, with the greater reduction in the PCC (Fig. 5C). The second-correlated population was produced by reducing the intensity of a subset of the pixels in one of the original images. Increasing the fraction of the original population that was altered progressively reduced the colocalization between the pair of images. Similarly, progressively reducing the gain applied to the subset, also progressively reduced the colocalization measured by both the PCC and MOC. The MOC was less responsive to the presence of a second-colocalized subset of pixels, falling from nearly 1 to around 0.75 whereas the PCC fell from nearly 0.9 to around 0.25. However, the fall in the MOC is appreciable, considering that the MOC is generally unresponsive to changes in the data and that measurements below 0.75 are hard to achieve.

Influence of a Single Datapoint

The coefficients were examined by adjusting a single datapoint out of 256, by rotating it around the mean values at different radii (Fig. 6). The patterns are quite different, the PCC varies with the angle and the radius. The SRC was similarly varied, in that the peaks were clipped once the two intensities hit the highest or lowest ranks. The ICQ jumped between two values and was independent of the radius. The MOC was unresponsive.

Weighting of the PCC, MOC, and ICQ

The coefficients all use the same data, the intensities of two fluorophores in homologous pixels. To explore the origins of the differences between the coefficients, the numerator and denominator were considered separately and the values for every combination of intensities were superimposed on a standard scattergram (Fig. 7). This displays the weight attached to each combination for the numerator, the denominator and finally shows their combined impact as a ratio (Fig. 7). The numerator of the PCC is based on the difference from the population mean and is nonlinear. The maximum weight of +16,256 applies to a pair of intensities at either the upper or lower limits of the intensity range, while pairs with values at the population mean carry no weight (zero). Since the PCC employs the difference from the mean, every positive weighting has a matching negative weighting. The MOC employs the absolute intensity, making every weighting positive, and is also nonlinear. A pair of intensities in the midrange have only a quarter of the weight (16,256) of a pair at the top of the range (65,025), while a combination of low intensities has little influence. The numerator of the ICQ is binary, either the differences from the mean have the same sign or they do not.

The denominator is similar to the numerator with a nonlinear pattern for the MOC and PCC. The pattern seen in the denominator and in the numerator are similar for the MOC, but are quite different for the PCC. The denominator for the ICQ is unweighted, every pixel pair has the same weight.

The ratio (numerator/denominator) shows combinations that the coefficients treat as having a similar degree of colocalization. The ratios show that both the PCC and the MOC have quite a broad band of maximal colocalization, where small changes to the intensities have little effect on the ratio. The surrounding gradient is relatively gentle for the MOC but changes more sharply for the PCC. Note that the "ratio"

panels on the right display areas of equal colocalization but that the numerator provides a better indication of the weight given to different combinations of intensities in the overall calculation: for the MOC, a high- and low-intensity pair may be equally colocalized but the high intensity carries a heavy weight whereas the low-intensity pair has a negligible influence. The scattergram, that was overlayed on the ratio panels, shows that almost all the points fall into the highly colocalized category for MOC whereas with the PCC they spread more widely, illustrating why the MOC returns high values and is unresponsive to changes in the data. The ratio for the ICQ is identical to the numerator.

The weights for the SRC were not calculated because the pattern, although basically similar to the PCC, will vary between datasets, an effect of ranking.

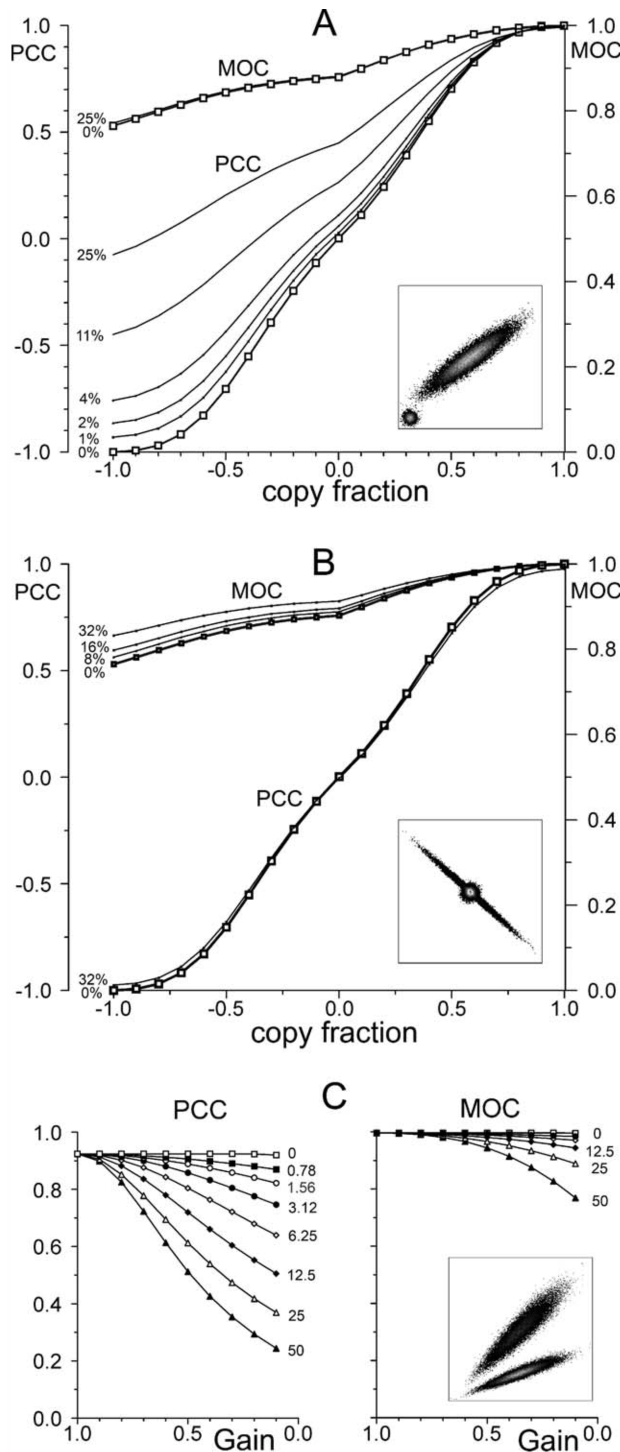


Figure 5. Colocalization and multiple relationships. The PCC and the MOC are affected by the presence of a subpopulation with a different degree of colocalization. In the upper panel (**A**) the subpopulation is uncorrelated and of low intensity, intended to emulate the inclusion of background pixels. The size of the subpopulation is expressed as a percentage of the whole population and the colocalization of the main population was varied by changing the copy fraction (see Fig. 1A). The MOC was unaffected even when background pixels formed 25% of the whole population and intermediate values are not shown. In the middle panel (**B**) the subpopulation has a mean equal to that of the main population. The bottom two panels (**C**) include a second correlated subpopulation, which is illustrated by the scattergram. The changes in gain (*x*-axis) apply only to the subpopulation and the size of subpopulation is shown as percentage of the total.

Discussion

It is self-evidently worthwhile to quantify colocalization but the plethora of available coefficients (PCC, MOC, ICQ, SRC, M_1, M_2, k_1, and k_2) and their differing meanings, can be confusing. We have made a detailed examination of two of the coefficients used for correlation measurements, the PCC and its derivative the MOC, to establish how they work and whether they are useful. The two coefficients are almost identical and differ only in the use of the absolute intensity, by the MOC, or the deviation from the mean, by the PCC, a small but significant change.

The MOC was created as an improvement on the PCC, to be "...especially applicable when the intensities of the fluorescence of detected antigens differ" (12) and because the PCC "is not sensitive to differences in signal intensity between the components of an image caused by different labeling with fluorochromes, photobleaching or different settings of amplifiers" (1), claims that have been repeated uncritically (5,11). We find that both the PCC and MOC are, within wide limits, independent of the magnitude of the signal. Therefore, the major claim made for the MOC falls.

Two further coefficients, k_1 and k_2, were derived from the MOC, by using only the intensity of one fluorophore in the denominator (1). The product of the intensities of both fluorophores is then related to the intensity of a single fluorophore, hence the need for two coefficients. Absolute intensity is embedded in the k_1 and k_2 coefficients, but image acquisition is almost always adjusted to fit the detector's response range and the actual intensities have little meaning. k_1 and k_2 really require the actual number of molecules present in each pixel. Even comparisons between cells imaged under standard conditions are problematic because uptake and expression of fluorescent molecules varies widely. The k_1 and k_2 seem to have no

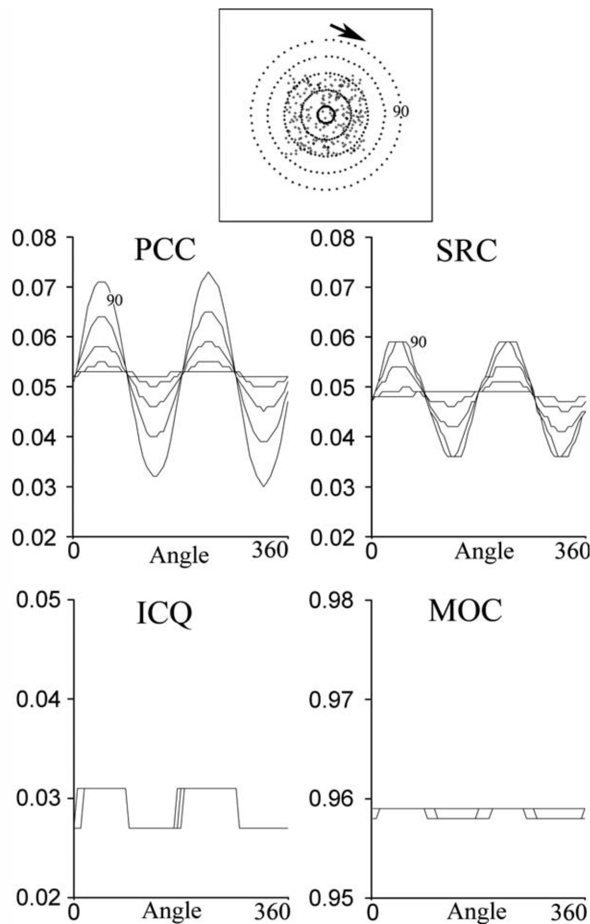


Figure 6. Effect of a single datapoint. One datapoint was moved radially about the mean of an uncorrelated population. Radii of 90, 70, 50, 30, and 10 were used, which are marked on the scattergram. The start position (angle of zero) is shown by the base of the arrow, which also shows the rotation direction. The peak response was seen with the largest radius (90). The scales used for each coefficient cover an equal fraction of the full response.

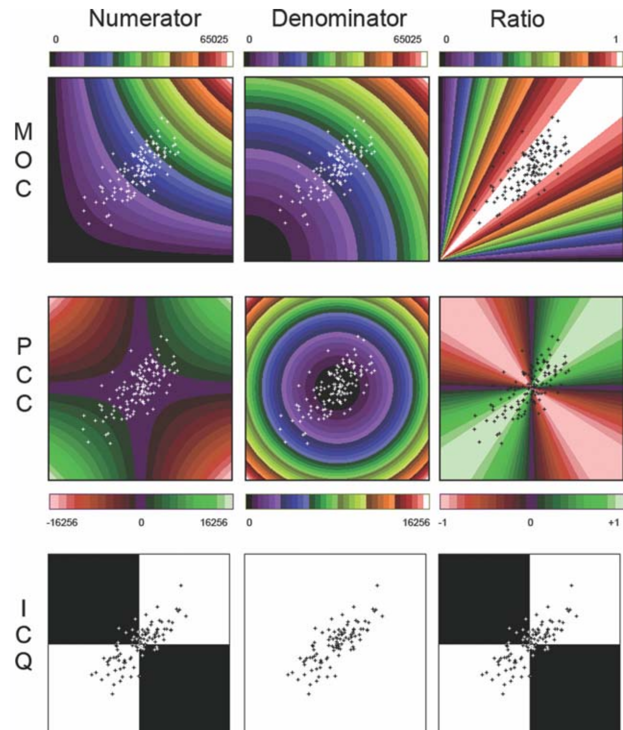


Figure 7. The weightings inherent in the PCC, the MOC and the ICQ for every combination of intensities found in an 8-bit range and based on the numerator and denominator of Eqs. (1), (2), and (3). Two false color scales are used, a red-blue-green scale that covers negative values and a rainbow scale for positive range. The ICQ range is binary, with white for 1 and black for 0. The values for an 8-bit intensity range are shown adjacent to each false color scale. A scattergram is superimposed.

advantages over the M_1 and the M_2 coefficients that were concurrently launched (1). M_1 and M_2 calculate for each fluorophore the fraction of the total intensity that co-occurs. The absolute values of k_1 and k_2 would only become meaningful if intensities were replaced by an estimate of the number of molecules present. However, even photon counting methods are not used routinely in biological imaging and estimating the number of molecules is difficult.

Offset has a differential effect on the PCC and the MOC. The PCC is completely independent of shifts in the signal but the MOC can either increase or, more surprisingly, be decreased, by positive offsets. The MOC works because the product of the intensities (the numerator) is less than or equal to the denominator. A positive offset increases the numerator more than it increases the denominator and the MOC rises. A fall in the MOC after a positive offset to one of the images is therefore unexpected. It arises when the correlation is negative and high intensities in one image correspond to low intensities in the second image, and vice versa. Then, the increase in the numerator is less than the increase in the denominator and the MOC falls. An increase in the MOC followed by a progres-

Table 1. Comparison of correlation coefficients

	PCC	MOC	ICQ	SRC
Theoretical range	−1 to +1	0 to +1	−0.5 to +0.5	−1 to +1
Gain	Independent	Independent	Independent	Independent
Offset background subtraction	Unimportant	Important	Unimportant	Unimportant
Weighting	Departure from the mean	Magnitude	None	Departure from mean rank
Inclusion of background pixels	Sensitive	Insensitive	Sensitive	Sensitive
Inclusion of midrange pixels	Insensitive	Sensitive	Sensitive	Slightly sensitive
Sensitivity to correlation	Good	Poor	Good	Good

sive fall is also possible when a single image is offset. This occurs with low intensities, low enough that the product with homologous pixels is nearly zero, and when a small increase, say from 1 to 2, has a bigger effect on the numerator than on the denominator. In this limited sense, the MOC is sensitive to the absolute signal. It is therefore important that the offset be set correctly, i.e., zero fluorescence produces zero detection. Correctly setting the offset is important since the position of the intercept contains useful information: a line that does not pass through the origin indicates that part of the fluorescence is independent of the second fluorophore and of the correlation between the two fluorophores. Since the size of any offset is not reported by the PCC, it could be considered a limitation.

The inclusion of uncorrelated pixels with low intensities, which emulate background pixels, has a profound effect on the PCC but leaves the MOC unchanged. The PCC becomes more positive and the effect on low or negatively correlated PCCs is substantial even when the percentage of background pixels is very small. The practical consequences are that the accurate measurement of the PCC requires the exclusion of background pixels, which should be standard practice. The failure to exclude pixels devoid of fluorescence transforms an apparently uncorrelated relationship into a highly positive PCC (10). The corollary is that pixels with intensities close to the mean affect the MOC but not PCC.

Combining two positively correlated populations appreciably reduces both the PCC and the MOC, although the absolute change in the MOC is smaller. This is a limitation of both coefficients. The coefficients summarize what may actually be a complex relationship that might include differently correlated subpopulations and nonlinear relationships. The PCC underestimates nonlinear relationships and the rank Spearman coefficient is a viable alternative (8,9). The original images and the scattergram should always be examined, even though visual assessment is imperfect (13).

When a scattergram suggests a complex relationship, it is tempting to select and then separately analyze any subpopulation (14). However, this is a fundamental error in data analysis, since the selection of the subpopulation is based on the very relationship for which an objective measurement is required. A legitimate alternative is to select biologically meaningful areas for analysis, e.g., individual cells rather than a tissue or to separate the cytoplasm from the nucleus. This might initially involve selecting a distribution from the scattergram and establishing its spatial origin in the specimen, but if a physiologically relevant area is highlighted then all the pixels in that area must be considered in the correlation analysis, i.e., if the "interesting" pixels come from the cell nucleus it is not legitimate to analyze only the selected pixels.

The explanation for the different properties and sensitivities of the PCC and the MOC lies in the different weighting given to the intensities of the two fluorophores. Since the PCC is based on differences from the mean, intensity pairs near the mean are of little consequence whereas those at the extremes of the intensity range are highly influential (8), hence the consequences of including background pixels. In the MOC, combinations of high intensities carry significant weight while combinations, where one or both of the pair is/are of a low intensity, have little influence on the numerator and a small influence on the denominator. This seems like an attractive feature, for a correlation coefficient, but a strong correlation requires that a match exists across the whole intensity range, including low intensities and the MOC is blind in this region. One high-intensity pair can produce a MOC that is almost unaffected by any number of blank or low-intensity combinations, which undermines its value as an overlap coefficient and makes the MOC a poor measure of co-occurrence. The biggest difference between the MOC and the PCC is apparent in the pattern of weightings for the numerator and denominator, they are similar for the MOC but differ with the PCC. The ratio of the numerator to the denominator shows one main axis for the MOC but two axes for the PCC, one strongly negative. This makes the PCC an effective measure of correlation. The different pattern of weighting explains the quite different meanings of a coefficient of zero: the PCC reports zero when there is no relationship between the intensities whereas the MOC reports zero only when the two fluorophores totally avoid each other.

The SRC is attractive because, unlike the PCC, it does not require a normally distributed population, a prerequisite that many biological specimens may not meet. The SRC also detects nonlinear correlations and is less sensitive to outlying datapoints than the PCC (8,9). It might be good practice to compare SRC and PCC and examine the raw data should they differ.

A new correlation method that counts only whether intensities are above or below the mean has been developed (6). The ICQ method simply expresses the number of matching pixels, when both are either above or below their mean, as a fraction of the total, and then subtracts 0.5. The subtraction ensures that negative correlations have a negative quotient, within a -0.5 to $+0.5$ scale. This scale differs from the more common -1 to $+1$ generally used for correlation. A remedy is simply to double the ICQ (15). The ICQ is a simple and therefore intelligible coefficient. The disadvantage is that pixels marginally above the mean carry exactly the same weight as pixels with more extreme intensities. Therefore, the ICQ is sensitive to changes in pairs of pixels that fall near the mean intensity of either fluorophore. By comparison, the PCC is almost unaffected by changes in this subset of pixels. Surprisingly, the ICQ was not rigorously compared with established coefficients when introduced (6). The ICQ performed well over the range of correlations produced by changing the copy fraction, being similar to the PCC and SRC. A tendency to flip between values was seen when a single pixel was moved and examination of the weightings suggests that there are datasets, which could undergo substantial changes without affecting the ICQ. The ICQ is nevertheless an interesting innovation.

A mistake often arises when two fluorophores that do not co-occur, with perhaps one in the cytoplasm and the other in the nucleus, are nonetheless tested for correlation. The PCC then reports a negative correlation, whereas the MOC reports a plausibly low value, the one occasion it delivers. This PCC is clearly spurious but these negative correlations are not always recognized as artifacts (10). It is important to differentiate

between a true negative correlation, where high intensities are matched with low intensities, and this "not in the same place" error. The lack of co-occurrence could be detected by the M_1 and M_2 coefficients. We strongly advocate thresholding to exclude pixels which do not contain both fluorophores and the separate analysis of biologically distinct regions. Automatic thresholding, using the idea that the background pixels are uncorrelated (16) or based on the background mean and standard deviation, are alternatives to operator controlled thresholding.

The MOC is considered to be easier to interpret than the PCC since it only reports positive values (4,5). Since negative correlations can arise, for example, an enzyme that converts a fluorescent molecule into a nonfluorescent form, quenching, FRET or localized avoidance, it seems appropriate to record them. The original case for the MOC "the negative values of the correlation coefficient (PCC) are difficult to interpret when the degree of overlap is the quantity to be measured" (1) is much more restricted and includes the important caveat "when the degree of overlap is of interest." Like many caveats, this one has been overlooked in the discussion of the PCC and the MOC (4,5).

The question arises as to what specifically "overlap" refers to in the context of the MOC, it remains undefined in the original article (1), unless the equation for the MOC is taken to be the definition and "an overlap coefficient equal to 0.5 implies that 50% of both components of the image overlap" is accepted, a claim for which there appears to be no justification. The assumption is that overlap is some measure of the degree of similarity in the distribution of two fluorophores, but the MOC is a curious hybrid measure combining elements of correlation with a highly weighted form of co-occurrence. It is in no way comparable with either the percentage of pixels in which co-occurrence is found nor to the M_1 and M_2, coefficients which report the fraction of each fluorophore's intensity that co-occurs.

It has been suggested that a threshold exists for values of the PCC (10,12) and the MOC (12) that mark biologically meaningful colocalization and, conversely, below which colocalization is deemed unimportant. It has been stated that no conclusions can be drawn from a PCC between −0.5 and 0.5 (10) and the MOC's threshold is apparently 0.6, for which no supporting evidence or rationale has been presented (12). Our results show that a MOC of <0.6 cannot be obtained even from datasets that show minimal or even negative correlation and that low values of the PCC have biological meaning (17). Even after randomly shuffling the pixel intensities, the MOC can still return values above 0.6 while randomization, predictably, reduces the PCC to zero but more surprisingly leaves the ICQ positive (15). However, since the PCC and MOC are graded measures the very idea of a threshold is strange (18), especially for values close to the nominal threshold, where a minor shift in the measurement would reverse the interpretation. The relevant biological consideration is whether the measured colocalization is changed experimentally. Even small changes "half an eye is just 1% better than 49% of an eye" (Richard Dawkins) and "information is any difference that makes a difference" (Gregory Bateson) (19) can be important. A more pertinent consideration is the accuracy and precision with which measurements can be made. It is acknowledged that the quality of the images influences the accuracy of colocalization measurements (1,16,20) and that noise reduces the measured colocalization. A correction for noise has been demonstrated for the PCC and SRC (8,21). The MOC is as insensitive to noise as to most other features of the data.

Overall, the PCC and the MOC produce values that differ widely for both the simulated datasets we have employed and with biological images (15,22) and there is little correlation between these two measures of correlation. The PCC does measure correlation, the degree to which the intensity variations of one fluorophore follows variation in the second fluorophore, but since only pixels containing both signals are analyzed, the PCC should be qualified by the M_1 and M_2 coefficients, which report the fraction of the total intensity that co-occur (1,23). The MOC provides a highly weighted measure of co-occurrence, is also affected by correlation and is sensitive to offset. For measurements of co-occurrence, the MOC should be replaced by M_1 and M_2. Given that colocalization is well supplied with coefficients, it would be productive to abandon the MOC and the related k_1 and k_2 pair of coefficients. The PCC, SRC, and perhaps the ICQ provide useful measures of correlation (Table 1).

Additional Equations

$$k_1 = \frac{\sum (R_i).(G_i)}{\sum \left(R_i^2\right).}$$

$$k_2 = \frac{\sum (R_i).(G_i)}{\sum \left(G_i^2\right).}$$

$$M_1 = \frac{\sum R_{i,\mathrm{coloc}}}{\sum R_1}$$

$R_{i,\mathrm{coloc}}$ is the intensity of the red fluorophore in pixels where the green fluorophore is present.

$$M_2 = \frac{\sum G_{i,\mathrm{coloc}}}{\sum G_1}$$

$G_{i,\mathrm{coloc}}$ is the intensity of the green fluorophore in pixels where the red fluorophore is present.

Literature Cited

1. Manders EMM, Verbeek FJ, Aten JA. Measurement of co-localization of objects in dual-colour confocal images. J Microscopy 1993;169:375–382.
2. Gillham EM. A Life of Sir Francis Galton, Oxford: Oxford University Press; 2001.
3. Manders EM, Stap J, Brakenhoff GJ, van Driel R, Aten JA. Dynamics of three-dimensional replication patterns during the S-phase, analysed by double labelling of DNA and confocal microscopy. J Cell Sci 1992;103 (Pt 3):857–862.
4. Zinchuk V, Zinchuk O, Okada T. Quantitative colocalization analysis of multicolor confocal immunofluorescence microscopy images: Pushing pixels to explore biological phenomena. Acta Histochem Cytochem 2007;40:101–111.
5. MediaCybernetics. Colocalization of fluorescent probes. Available at: http://www.mediacy.com/pdfs/colocfluorprobes.pdf; 2002. pp1–5.
6. Li Q, Lau A, Morris TJ, Guo L, Fordyce CB, Stanley EF. A syntaxin 1, Galpha(o), and N-type calcium channel complex at a presynaptic nerve terminal: Analysis by quantitative immunocolocalization. J Neurosci 2004;24:4070–4081.

7. Khanna R, Sun L, Li Q, Guo L, Stanley EF. Long splice variant N type calcium channels are clustered at presynaptic transmitter release sites without modular adaptor proteins. Neuroscience 2006;138:1115–1125.
8. Adler J, Pagakis SN, Parmryd I. Replicate-based noise corrected correlation for accurate measurements of colocalization. J Microsc 2008;230 (Pt 1):121–133.
9. French AP, Mills S, Swarup R, Bennett MJ, Pridmore TP. Colocalization of fluorescent markers in confocal microscope images of plant cells. Nat Protoc 2008;3:619–628.
10. Bolte S, Cordelieres FP. A guided tour into subcellular colocalization analysis in light microscopy. J Microsc 2006;224:213–232.
11. Oheim M, Li D. Quantitative colocalisation imaging: Concepts measurements and pitfalls. In: Shorte SL, Frischknecht F, editors. Imaging Cellular and Molecular Biological Functions. Springer; 2007. pp 117–155.
12. Zinchuk V, Grossenbacher-Zinchuk O. Recent advances in quantitative colocalization analysis: Focus on neuroscience. Prog Histochem Cytochem 2009;44:125–172.
13. Cleveland WS, Diaconis P, Mcgill R. Variables on scatterplots look more highly correlated when the scales are increased. Science 1982;216:1138–1141.
14. Penarrubia PG, Ruiz XF, Galvez J. Quantitative analysis of the factors that affect the determination of colocalization coefficients in dual-color confocal images. IEEE Trans Image Process 2005;14:1151–1158.
15. Nakamura K, Watakabe A, Hioki H, Fujiyama F, Tanaka Y, Yamamori T, Kaneko T. Transiently increased colocalization of vesicular glutamate transporters 1 and 2 at single axon terminals during postnatal development of mouse neocortex: A quantitative analysis with correlation coefficient. Eur J Neurosci 2007;26:3054–3067.
16. Costes SV, Daelemans D, Cho EH, Dobbin Z, Pavlakis G, Lockett S. Automatic and quantitative measurement of protein-protein colocalization in live cells. Biophys J 2004;86:3993–4003.
17. Parmryd I, Adler J, Patel R, Magee AI. Imaging metabolism of phosphatidylinositol 4,5-bisphosphate in T-cell GM1-enriched domains containing Ras proteins. Exp Cell Res 2003;285:27–38.
18. Adler J, Parmryd I. In support of the Pearson correlation coefficient. J Microsc 2007; 227 (Pt 1):83.
19. Kay A. Computer software. Sci Am 1984;251:41–47.
20. Demandlox D, Davoust J. Multicolour analysis and local image correlation in confocal microscopy. J Microscopy 1997;185:21–36.
21. Adler J, Bergholm F, Pagakis SN, Parmryd I. Noise and colocalization in fluorescence microscopy: Solving a problem. Microsc Anal 2008;155:7–10.
22. Galloway S, Takechi R, Pallebage-Gamarallage MM, Dhaliwal SS, Mamo JC. Amyloid-beta colocalizes with apolipoprotein B in absorptive cells of the small intestine. Lipids Health Dis 2009;8:46.
23. Smallcombe A, McMillan D. Co-localisation: How is it determined, and how is it analysed with the Bio-Rad LaserPic image analysis software. Bio-Rad. pp1–8.