AUCTORES
Globalize your Research

Journal of Nutrition and Food Processing
Arne Torbjørn Høstmark*

Open Access

Research Article

Intended Ranges and Correlations between Relative Amounts: a Review

Arne Torbjørn Høstmark

Institute of Health and Society, Faculty of Medicine, University of Oslo, Norway, Box 1130 Blindern, 0318 Oslo, Norway, Telephone: +47 22844629/Fax: +47 22850590.

Corresponding Author: Arne Torbjørn Høstmark, Institute of Health and Society, Faculty of Medicine, University of Oslo, Norway, Box 1130 Blindern, 0318 Oslo, Norway.

Received date: November 25, 2021; Accepted date: December 07, 2021; Published date: January 01, 2022

Citation: Arne T. Høstmark, (2022). Intended Ranges and Correlations between Relative Amounts: a Review, *J. Nutrition and Food Processing*; 5(1); DOI:10.31579/2637-8914/077

Abstract

The present article is a review of our previously suggested concepts of "*Distribution Dependent Correlations" (DDC),* and "*Intended Ranges*". DDC concern associations between relative amounts of positive scale variables, in unit systems where sum of the percentages is 100%. Such correlations arise mathematically on the condition that the variables in question have particular ("intended") ranges. For example, with three variables, two of which (A, B) having very low variability relative to a third one (C), we should expect a positive association between percent A and percent B, the slope being estimated by the B/A ratio. In addition, we should anticipate a negative relationship between %C and %A (%B). On the other hand, if A and B have high numbers and variability relative to C, then %A should relate inversely to %B. Furthermore, alterations in the ranges may have appreciable effects to change the associations. We present examples from physiology, where ranges seem to give strong DDC (positive and negative). The examples relate to body fatty acids, and white blood cell counts. Possibly, *Intended Ranges* could represent a case of evolutionary selection, to ensure proper balance between particular metabolites.

Definitions and Abbreviations:

Variability: the width or spread of a distribution, measured e.g. by the range and standard deviation.

Distribution: graph showing the frequency distribution of a variable within a particular range. In this article, we also use distribution when referring to a particular range, a – b, on the scale.

Uniform distribution: every value within the range is equally likely. In this article, we may write, "Distribution was from a to b", or "Distributions of A, B, and C were a - b, c - d, and e - f, respectively".

OA = Oleic Acid (18:1 c9); LA = Linoleic Acid (18:2 n6); ALA = Alpha Linolenic Acid (18:3 n3); AA = Arachidonic Acid (20:4 n6); EPA = Eicosapentaenoic Acid (20:5 n3); DPA = Docosapentaenoic Acid (22:5 n3); DHA = Docosahexaenoic Acid (22:6 n3); DGLA= dihomo-gammalinolenic acid (20:3 n6)

"Low–number variables" have very low numbers relative to "high-number variables".

Keywords: *correlation rules; relative amounts; ranges; biological regulation; fatty acids; white blood cells*

Introduction

This article is a review and extension of our previously suggested concepts of "*Distribution Dependent Correlations" (DDC),* and *"Intended Ranges*". Below, we first present some theoretical considerations to explain mathematically the phenomenon of DDC. Next, we show results of computer experiments with random numbers, to test the hypotheses. Finally, we show some examples from physiology, where "intended ranges" seem to govern DDC.

Particular background

The idea of "intended ranges" and "distribution dependent correlations" originated from a diet trial in chickens, carried out for a specific purpose, without any bearing on the present subject [1]. During post-trial analyses of the data, we observed remarkably strong positive and negative correlations between relative amounts of the measured fatty acids (in breast muscle). The present work relates to some of these correlations, and our efforts to explain them, as well as our attempts to find some general rules concerning associations between relative amounts.

A major part of the current article relates to body fatty acids. In brief, dietary intake is a major factor to regulate the concentration of fatty acids in blood and tissues, and these lipids are important in health and disease [2-4]. For example, poly-unsaturated fatty acids with 20 or 22 carbon atoms are precursors of eicosanoids and docosanoids, which are important regulatory molecules formed in most organs and cell types, through the actions of cyclooxygenases, lipoxygenases, and epoxygenases [5]. EPA (20:5 n3) - derived eicosanoids may decrease inflammatory diseases [6,7], decrease the risk of coronary heart diseases [8,9], and cancer [10], but the beneficial effects of long-chain n3 fatty acids on all-cause and cardiovascular mortality have been questioned [11].

On the other hand, eicosanoids derived from AA (20:4 n6), such as thromboxane A2 (TXA2) and leukotriene B4 (LTB4), have strong proinflammatory and prothrombotic properties, and are involved in allergic reactions and bronchoconstriction [2, 3, 5]. Furthermore, AA-derived endocannabinoids may have a role in adiposity and inflammation [12]. Additionally, low serum EPA/AA ratio may be a risk factor for cancer death [10]. Thus, the EPA and AA antagonism could explain many of the alleged positive health effects of EPA.

Also docosanoids, originating from C22 fatty acids (DPA, DHA), have strong metabolic effects. Among these latter compounds are protectins, resolvins, and maresins, which may strongly counteract immune- and inflammatory reactions [5]. Also eicosatrienoic acid, i.e. 20:3 n6 (dihomo-gammalinolenic acid, DGLA) may give eicosanoids [5].

To obtain a proper balance between the metabolic influences of the many eicosanoids and docosanoids, we should anticipate a coordinated regulation of their precursor fatty acid percentages, e.g. of % EPA, %AA, and %DGLA. Indeed, we might expect in general that these particular percentages of the total sum of fatty acids were positively associated, so that an increase (decrease) in e.g. %AA would be accompanied by a concomitant increase (decrease) in other fatty acid precursor percentages as well, in order to obtain the required balance. We previously reported that that %AA, %EPA, %DHA, as well as other eicosanoid (docosanoid) percentages were positively associated in breast muscle lipids of chickens [12 - 15], as discussed below. We also showed that the correlation outcomes related to the particular *concentration distributions* of the fatty acids. This finding seemed to be in line with the remarkably similar outcomes with true values and with surrogate random numbers, found on the condition that we sampled the numbers within the true ranges of the fatty acids [14 - 18]. Furthermore, computer experiments showed that *altering* the ranges strongly influenced the correlation outcomes, in support of our suggested name: *Distribution Dependent Correlations*, DDC [14, 18 - 27].

In addition to the situation with body fatty acids, we observed that also the ranges of white blood cell (WBC) counts influenced the inverse relationship between e.g. the relative amounts of blood neutrophil granulocytes and lymphocytes [27]. We subsequently suggested the name *"intended ranges"* [28] to indicate ranges that might possibly serve to make strong correlations (positive and negative) between relative amounts of biological variables in "*unit systems"*, as exemplified by WBC, and by particular fatty acids

Since DDC rules are general, they should apply to any unit system in nature. However, investigations specifically focusing upon this issue seem hard to find, in a literature search. The apparent lack of interest might possibly relate to a methodological concern encountered when correlating percentages of the same sum, since the associations arise mathematically. On the other hand, it may not be obvious whether we should reject strong positive (negative) associations between percentages of the same sum as correlation bias. Rather, we previously suggested that intended ranges could be a case of evolutionary selection to obtain strong DDC [20 - 22]. The first part of this review article concerns mathematical explanations of correlations arising between percentages of the same sum. The presentation does not necessarily reflect our opinions in chronological order. Rather, we try to give a systematized presentation, based upon our present knowledge. However, in the second part, a synopsis of papers related to this topic, appear chronologically. We present some examples from physiology where evolution seems to have selected particular ranges to make relative amounts of some variables to be positively or negatively associated, mathematically, i.e. *Distribution Dependent Correlations.*

Materials and Methods

The present review concerns associations between relative amounts of positive scale variables in "*unit systems*", where sum of relative amounts is 100%. We define A, B, C… to be positive scale variables and S their sum, i.e. S = A + B + C +… All of the variables should have the same unit, for example, g/kg, g/L, moles/L, or counts/L, i.e. the *absolute* amounts of muscle fatty acids may appear as *g/kg wet weight*. Furthermore, each of the variables should have particular ranges.

Previously [21], we investigated the association between relative amount of e.g. arachidonic acid (AA, 20:4 n6) and percentage of eicosapentaenoic acid (EPA, 20:5 n3), in chicken lipids. From histograms, the physiological concentration distributions (g/kg wet weight) for the fatty acids were determined. Next the sum (S, g/kg wet weight) of all fatty acids was computed, as well as and the remaining sum (R) when omitting the couple of fatty acids under investigation, thereby apparently obtaining 3 positive scale variables. With these variables, and with surrogate random number variables, generated with the true concentration distributions, computer analyses as described in detail below, were carried out. Our previous analyses [20, 21] demonstrated that correlations between e.g. %A and %B depended upon the particular range of each of the variables involved, and we obtained qualitatively similar correlations using the true (measured) values, or random numbers, *if ranges were like the measured ones.*

A major part of the present work consists of computer experiments using random numbers to explore further, how ranges of A, B, and C might influence correlations between relative amounts of the sum, S = A + B + C. This equation implies that sum of the A (B, C) *percentages* of S is 100; i.e. %A + %B + %C = 100, showing dependency between the percentages. We studied histograms, scatterplots, and correlations (Spearman's rho). Computer experiments were performed, to study how *alterations* in ranges might change associations between %A, %B, and %C. Several repeats were carried out, with new sets of random numbers (for simplicity, n = 200 each time); the general outcome was always the same, but corresponding correlation coefficients and scatterplots varied slightly. We present the results mainly as scatterplots with correlation coefficients. In most of the computer experiments, the random numbers had *uniform* distribution, but we used random numbers with *normal* distribution as well, however obtaining qualitatively similar results. We used SPSS 27.0 for the analyses, and for making figures. The significance level was set at $p < 0.05$. We present further details under Results and Discussions.

Results and Discussions

Unit systems with two variables only

We name the two variables A and B, respectively, and choose their ranges arbitrarily to be e.g. 11-37 for A, and 51 - 83 for B. Thus, A + B = S. Within these ranges, and for each of for example 200 cases, we generate random numbers of the variables. Thereby, we obtain 200 S-values: S_1, S_2, S_3,. .. S_{200}, each of which representing one of the 200 cases. We do not assume any relationship between the absolute amounts (Fig. 1, left panel). In contrast to this, the *relative* amounts of A and B must vary inversely, since sum of the relative amounts will always be 100%. A (B) percentages of S are %A = (A/S) x 100, and %B = (B/S) x 100, respectively. Thus, %A + %B = 100, or %B = -%A + 100, showing a perfect inverse linear relationship (Fig. 1, right panel). All %A values must correspond inversely to %B, since %B = - %A + 100, irrespective of which of the allowed values of A (B) appearing for a certain case. For example, the lowest (highest) %A must correspond to the highest (lowest) %B. Unlike this strict requirement put upon A and B *percentages*, for each case, the *absolute* A-value could have any values from 11 to 37, and B any value from 51 to 83, explaining the lack of correlation between A and B.

Accordingly, in a unit system with two variables only, each of which with a particular range, we should find a perfect inverse relationship between their relative amounts, irrespective of their ranges. We may raise the question of whether the inverse relationship between relative amounts, in a "two-variable unit system", could be of any physiological interest.

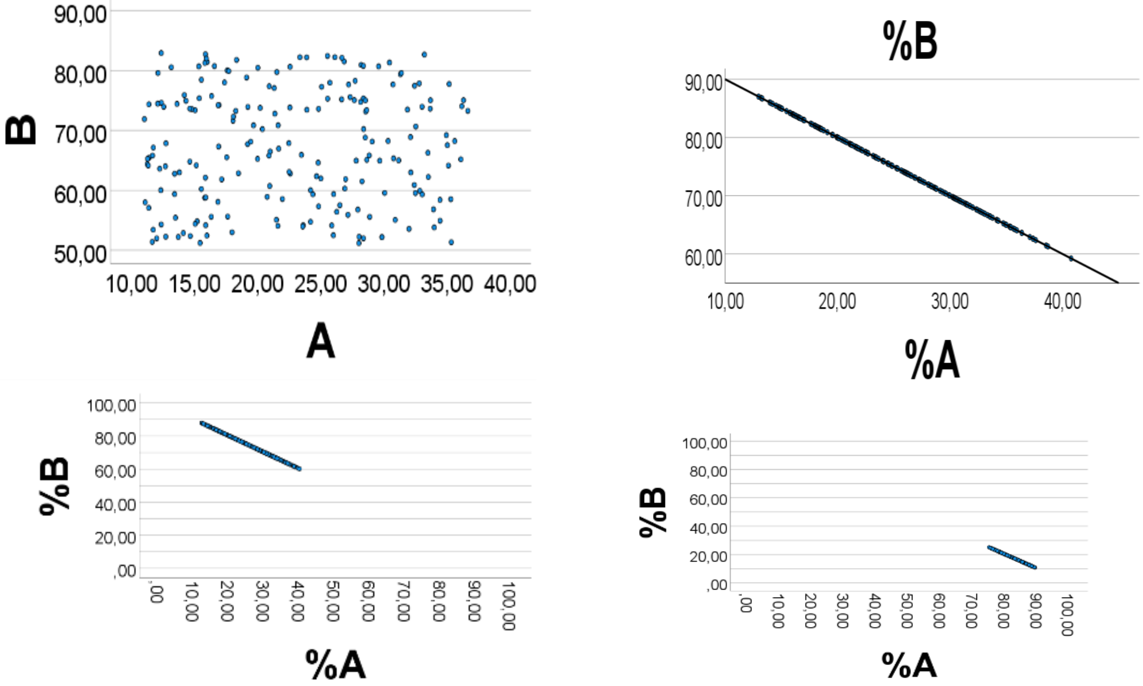

Figure 1: *Association between absolute amounts (upper panel, left) and relative amounts (upper panel, right, and lower panels) of A and B, in a unit system, i.e. %A +% B = 100%, see text. We generated 200 uniformly distributed random numbers of both variables. In all panels, except lower, right panel, A had range 11-37, and B 51 - 83. In lower panel, right: A 70 - 90, B 10 - 25. %A vs. %B: rho = -1.000 in all, p<0.01.*

The equation %B= -%A + 100 shows that the regression line for %A vs. %B should pass through 100% on both axes, irrespective of the sizes of the A and B values. However, the A and B *ranges* determine where we find the points on the %A vs. %B scatterplot. Thus, if A has low (high) numbers relative to B, then the points will appear in the lower (upper) part of the %A scale, implying that the %B to %A *ratio* is distributed among high (low) values, respectively (Fig. 1). The %B to %A ratio is equal to the B/A ratio, as seen from the calculation of the percentages. In medicine, we might encounter close to "two variable unit systems" where distribution of the B/A ratio among high or low values could be of clinical interest [29, 30].

Below, we shall consider "unit systems" involving more than two positive scale variables, each of which having particular ranges. In the theoretical reasoning, we first utilize the equation of a straight line (y = ax + b). Next, we consider the relationship between relative amounts of the variables and their sum. The theory and computer experiments suggest that relative amounts (percentages) should correlate positively or negatively, because of their particular ranges.

Unit systems with three variables [16]

Applying the equation of a straight line (y = ax + b)

In the current context, we have A + B + C = S, i.e. %A + %B + %C = 100, or %B = -%A + (100 - %C). This equation resembles that of a straight line (y = ax + b), however involving relative amounts of the three variables (A, B, C), each of which having a particular range. We will consider this equation in three particular situations: 1) if the expression (100 - %C) is approaching zero, i.e. %C being close to 100, 2) if %C is approaching zero, and 3) if %C approaches a constant. As shown below, in these conditions, we may obtain very strong correlations between relative amounts. Presumably, we should expect strong correlations, if they reflect biological regulation.

% C Approaching 100

The %A vs. %B Association

If %C consists of high values (close to 100) and (100 - %C) > %A, then the equation appears to approach **%B = %A**, apparently showing a linear positive association between %A and %B. The requirement (100 - %C) > %A is indeed satisfied, since the remaining value when calculating (100 - %C) would have to be divided between %A and %B. Hence, the slope of the %A vs. %B regression line should be positive. Additionally, to obtain %C values close to 100, the numbers of A and B should be very low.

Finding the %A vs. %B Slope

We may estimate the slope by utilizing maximum and minimum values of %B and %A, i.e. by the ratio $(\%B_{max} - \%B_{min})/(\%A_{max} - \%A_{min})$. A more general equation would therefore be:

$$\%B_{(p-q)} = [(\%B_{max} - \%B_{min})/(\%A_{max} - \%A_{min})] * \%A_{(r-s)} + z$$

The subscript parentheses indicate ranges of %A and %B, and z = 100 - %C. Thus, ***z*** becomes increasingly small as %C increases. The approximated slope value would accordingly be:

$$\Delta Y/\Delta X = (100 \cdot B_{max}/S_{min} - 100 \cdot B_{min}/S_{max})/(100 \cdot A_{max}/S_{min} - 100 \cdot A_{min}/S_{max})$$

Since *ranges of A and B are very narrow in this case*, we may do the following approximations: $A_{max} = A_{min} = \boldsymbol{A}$, and $B_{max} = B_{min} = \boldsymbol{B}$. Thus,

$$\Delta Y/\Delta X = (\mathbf{B} \cdot S_{max} - \mathbf{B} \cdot S_{min})/(\mathbf{A} \cdot S_{max} - \mathbf{A} \cdot S_{min}) = \boldsymbol{B/A}.$$

Accordingly, the slope may be estimated by the ***B/A ratio***, and should approach +1 only if A approaches B. Conceivably, the slope value computed manually based upon the approximated values may deviate somewhat from the corresponding one found by the computer. This deviation should increase in response to increasing the A and/or B ranges.

Thus, *if **ranges** of A and B are very narrow relative to the C – range*, the slope estimate of the %A vs. %B regression line should be little influenced by the *magnitudes* of the variables. Furthermore, with these restrictions laid upon A, B, and C *variabilities*, we suggest that the slope estimate of the %A vs. %B regression line (i.e. the B/A *ratio*) should apply to any C value on the positive scale, and to any *sizes* of the A and B numbers. Furthermore, with very narrow ranges of A(B), the scatterplot of the %A vs. %B association should be close to a line, since one particular value of %A (and of %B) corresponds closely to one S-value only. Hence, %A and %B should show a strong positive association. Thus, the %A vs. %B scatterplot should improve (be poorer) in response to narrowing (broadening) the A and/or B ranges, and also improving (be poorer) when increasing (decreasing) the C-range.

The %C vs. %A (%B) Association

Since %C values are very high, it follows that A(%A) and B(%B) should be low. We rewrite the equation %A + %B + %C = 100, to be *%B = -%C + (100 - %A).* With very low %A - values, the equation would approach ***%B = -%C + 100***, suggesting that %C and %B are *inversely* related. Similarly, the approximation **%A = -%C + 100**, suggests an inverse %C vs. %A association. Hence, in the current case, we should expect a *negative* %C vs. %A (%B) association.

Finding the slope of %C vs. %A (%B) when A (B) Ranges are Narrow Relative to the C – Range [26]

Slope of the *%C* vs. *%A* (%B) regression line may be roughly estimated using maximum and minimum values of %C and %A, i.e. $\Delta Y/\Delta X = -(\%C_{max} - \%C_{min})/(\%A_{max} - \%A_{min})$. Similarly, slope of the *%C* vs. *%B* regression line may be estimated by $\Delta Y/\Delta X = -(\%C_{max} - \%C_{min})/(\%B_{max} - \%B_{min})$. In these cases, the simplification above does not work, due to high C (%C) - variability.

Computer Test: To obtain very high %C values relative to %A and %B, we arbitrarily chose A 1.0 - 1.04; B 2.0 - 2.05; C 1 - 10. As shown in Fig. 2, there was a strong positive association between %A and %B (rho = 0.999, p<0.01, n =200. Equation of the regression line was %B = 1.985 (0.003)*%A + 0.001 (0.014), SE is shown in parentheses. The B/A ratio estimated the slope well. As expected, there was a strong *negative* relationship between %C and %A (%B), rho = -1.000 (-1.000), p<0.01 for both, n = 200. Quartiles of %A, %B and %C were 1.2, *1.8, 4.1; 2.4, 3.5, 8.0; and 87.9, 94.6, 96.3,* respectively. Thus, %C had high values relative to %A and %B. Skewness of %A, %B, and %C was 2.56, 2.57, and -2.57, respectively (SD 0.17 for all). Below, we will explain this skewness outcome.

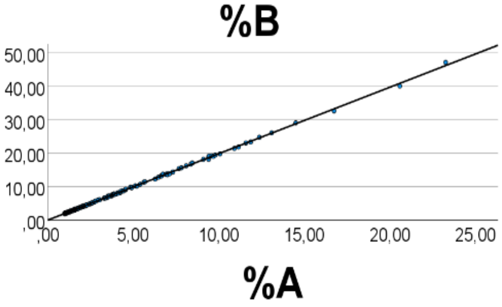

Figure 2: *Association between %A and %B. The figure relates to the equation %A + %B + %C = 100, see text. Random numbers (n = 200) with uniform distribution were generated. Ranges were A 1.0 - 1.04; B 2.0 - 2.05; C 1 - 100. %A vs. %B: rho = 0.999; %C vs. %A (%B): rho = - 0.999 (-1.000); p<0.01 for all.*

These results seem to be in line with the reasoning above: with high %C values relative to %A and %B values, we should expect a positive association between %A and %B, and a negative relationship between %C and %A (%B). Additional computer experiments with a large number of varying ranges of the variables, however always keeping the above restrictions, showed results in keeping with the reasoning above (results not shown).

% C Approaching Zero [16, 19]

If %C in the equation %B = -%A + (100 - %C) consists of very *low* values relative to %A (%B), we would expect a *negative* %A vs. %B association, since the equation then would approach ***%B = - %A + 100.*** However, in this case we should probably not expect that a decrease in %C would suffice to compensate a major increase in %A or %B. Hence, we should probably expect a poor correlation between %C and %A (%B).

Computer Test: To obtain very low values of %C relative to %A and %B, we arbitrarily chose A 10 - 50, B 20 - 67, C 0.10 - 0.13. Spearman’s rho = -1.000 for %A vs. %B, p<0.01, n =200; rho = 0.044 (-0.048), p = 0.532 (0.502) for %C vs. %A (%B). Quartiles of %A, %B and %C were *33.2, 40.8, 50.0; 49.9, 59.1, 66.6; 0.12, 0.15, and 0.18,* respectively. Thus, values of %C were small relative to those of %A and %B.

%C Approaching a Fixed Number

If %C is close to a constant value, **k**, then the equation %A + %B + %C = 100 may be written ***%B = -%A + k***, where k = 100 - %C. In this case, %A and %B should be inversely related. To achieve that %C is near a constant, the C-range should be very narrow.

Computer Test: To make %C close to a constant, we arbitrarily chose A 2 - 24; B 4 - 40, and **C 3.0 - 3.3.** As expected, %A correlated negatively with %B (Fig. 3, left panel), rho = **-0.975**, $p<0.01$, n =200. Quartiles of the %A, %B, and %C histograms were 23.2, 34.8, 45.2; 45.9, 57.0, 68.4; 6.5, 8.1, 10.4, respectively. The %C vs. %A (%B) scatterplots were poor (not shown). We found strong positive skewness (1.62) of the % C distribution, and weak skewness of %A and %B, i.e. 0.29 and -0.37, respectively.

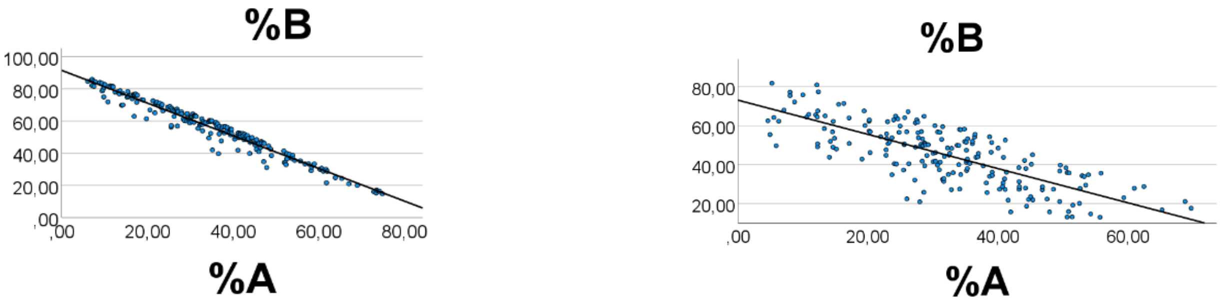

Figure 3: *Association between %A and %B. The figure relates to the equation %A + %B + %C = 100, see text. Random numbers (n = 200) with uniform distribution were generated. Ranges in left panel were A 2 - 24; B 4 - 40;* ***C 3.0 – 3.3****, and in* ***right*** *panel A 2 - 24; B 4 - 40;* ***C 3 – 18.*** *Left panel: rho = -0.975. Right panel: rho = -0.764, p<0.01 in both.*

According to the reasoning above, we should expect a poorer scatterplot if broadening the C range, thereby moving away from a situation where C is close to a constant. To obtain this condition, we broadened the C range to 3 - 18, while keeping ranges of A and B. As shown in Fig., right panel), the %A vs %B scatterplot did become poorer, and so did the correlation coefficient (rho = **-0.764**, $p<0.01$, n =200. Quartiles of the %A, %B, and %C histograms were 23.7, 31.4, 40.9; 34.3, 46.0, 56.5; 14.7, 22.6, 30.8, respectively. The positive skewness of the % C distribution was attenuated to 0.34. %A and %B distributions had normalized, skewness being 0.18, and -0.06, respectively. Thus, broadening the C-range had moved the %C distribution towards higher values, and made %C deviate appreciably from a fixed number. Both mechanism should work in favour of attenuating the negative %A vs. %B association.

To illustrate a mathematical point, we studied correlations between percentages of variables with *very narrow ranges*. First, we made the following ranges: A 4.0 - 4.1; B 7.0 - 7.1; **C 5.00 - 5.01**, i.e. C should be closer than A and B to be a fixed number. We should, accordingly, expect a negative %A vs. %B association, and this was verified in a computer test (Fig. 4). %A vs. %B, rho = -0.805, $p<0.01$, n =200.

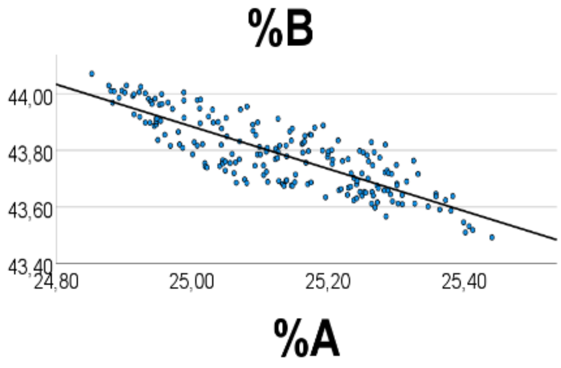

Figure 4: *Association between %A and %B when range of C << ranges of A and B. The figure relates to the equation %A + %B + %C = 100, see text. Random numbers (n = 200) with uniform distribution were generated. A 4.0 - 4.1, B 7.0 - 7.1,* ***C 5.00 - 5.01*** *(%B vs. %A, rho= -0.805, p<0.01).*

Considering Associations between Fractions and Sum (S)

Accordingly, if %C is approaching 100, obtained e.g. when *A and B have low numbers and narrow ranges* relative to C, we might expect a positive association between %A and %B, and a negative relationship between %C and %A (%B). Below, we consider further how ranges might govern associations between relative amounts of A, B, and C.

Two Positive Scale Variables (A and B) with Narrow Ranges Relative to a Third One (C) with High Variability [25, 26]

We consider the relationship between sum (S = A + B + C) of the variables and the A (B, C) *fractions* of S. If %A, as well as %B, relate negatively to S, we should expect a positive correlation between percent A and percent B. Furthermore, %C vs. %A (%B) should be inversely related, because we would expect %C to increase with increasing S. To explain this outcome in more detail, we omit ranges of the variables. The A, B, and C *fractions* of S are Af =A/S, Bf = B/S, and Cf = C/S, respectively. By definition, $Af = A/(A + B + C) = 1/(1 + B/A + C/A)$. However, since we - in the current context - define *ranges* of A and B to be very narrow, the B/A ratio is *close to be a fixed number*. Therefore, *Af* would approach $Af = 1/(k + C/A)$ where **k** approaches a constant, k = 1 + B/A. Similarly, the B-fraction of S, $Bf = B/(A + B + C) = 1/(1 + A/B + C/B)$, i.e. $Bf = 1/(t + C/B)$, where **t** is close to be a constant, t = (1 + A/B).

This means that C will largely govern the A (B) *fractions of S*. Thus, when C and S (being mainly composed of C) go from lowest to highest value, then *Af = 1/(k + C/A)*, and also *Bf = 1/(t + C/B)*, will go from the highest to the lowest value. Hence, S should relate inversely to the A- and B-fractions (percentages). This way of reasoning should apply to any positive values of A, B, and C, *if ranges of A and B are very narrow relative to that of C*. Accordingly, with this restriction, we should expect percent A to be positively associated with %B, *wherever we place A, B, and C on the positive scale.* However, *increasing the A- and/or B-ranges* (variabilities), and/or *decreasing the C-range*, would cause deviations from the above restrictions, and accordingly attenuate the %A vs. %B association, suggested to be reflected in poorer scatterplots and correlation coefficients.

The C-fraction of S is Cf = C/*S = C/ (A + B + C*), i.e. C*f = 1/ (1 + z/C), where z* is close to a constant, *z =A + B*. Therefore, the C fraction (and percentage) of S should *increase* with increasing C (from lowest to highest value), and accordingly also with increasing S, because C is the main contributor to S. Thus, S should be *positively* associated with %C, irrespective of where on the positive scale we place A, B, and C. It follows that %C should be negatively associated with %A and %B. In summary, from the relationships between S and A (B, C) percentages (fractions) of S, when putting the current restrictions on the ranges, we would anticipate a positive %A vs. %B association, and an inverse relationships between %C and %A (%B), wherever we encounter A, B, and C on the positive scale.

Computer Tests

A (B) with low variability and C with high variability

We arbitrarily used the following ranges A 3.0 - 3.3, B 8.0 - 8.8, C 0 - 100, i.e. A (B) with low variability relative to C. %A correlated positively with %B, rho = 1.000, p<0.01, n = 200.

%C vs. %A (%B), rho = -0.999 (-1.000), p<0.01 for all, n = 200. Quartiles of %A, %B, %C were *3.6, 5.0, 8.4; 9.1, 12.7, 21.5; 70.0, 82.3, 87.2*, respectively. Skewness of %A, %B, and %C were 2.07, 2.05, and -2.06, respectively, i.e. a strong positive skewness of %A and %B, and a strong negative skewness of the %C distribution (Fig. 5).

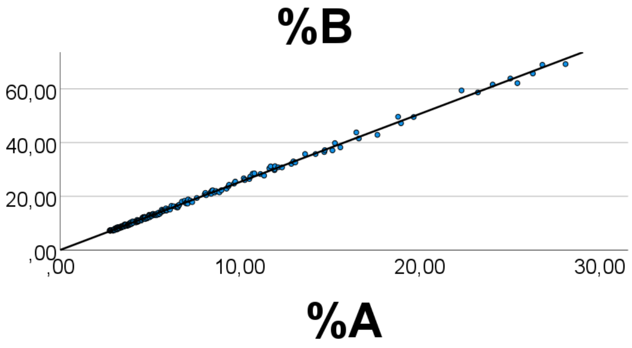

Figure 5: *Association between %A and %B (left panel); the figure relates to the equation %A + %B + %C = 100, see text. Random numbers (n = 200) with uniform distribution were generated. Ranges were A= 3.0 - 3.3, B= 8.0 - 8.8, and C 0 - 100. %A vs. %B: rho = 1.000; p<0.01, p<0.01, n=200. Eq. of the regression line %B = 2.533 (0.008)*%A + 0.041 (0.070)*

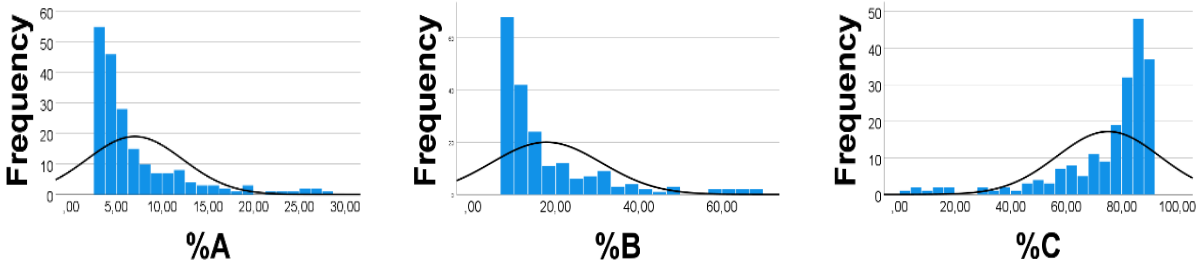

Figure 6: *Histograms of the distributions of %A, %B, and %C; the figure relates to the equation %A + %B + %C = 100, see text. Random numbers (n = 200) with uniform distribution were generated. Ranges were A= 3.0 - 3.3, B = 8.0 - 8.8, and C 0 - 100. Skewness of %A, %B, and %C were 2.07, 2.05, and -2.06, respectively; SE of skewness was 0.17 for all.*

As expected, C correlated positively with %C, and negatively with %A (%B), not shown.

All of the three variables (A, B, C) have very narrow ranges, but one of them (C) is closer to a constant than the remaining ones

Above, we showed that, if choosing **C** to be close to a fixed number, the equation %B = -%A + (100 - %C) may be approximated to *%B = -%A + k*, where k is near a constant equal to (100 - %C). Accordingly, %A should relate *negatively* to %B, irrespective of the sizes and ranges of A and B.

All of three variables (A, B, C) have very narrow ranges, but two of them (A, B) are closer to a constant than the remaining one (C)

If A and B are even closer than C to be fixed numbers, then %A and %B should be *positively* associated. In this latter case, Af = A/(A + B + C), i.e. Af =1/(1 + B/A + C/A), which may be written ***Af = 1/(t + C/A)*** where t = 1 + B/A, being close to a constant. Similarly, ***Bf = 1/(k + C/B)*** where **k** is near a constant, 1 + A/B. Accordingly, Af, as well as Bf, are not far from being *dependent upon C only*. Both fractions should decrease as C increases from lowest to highest value. Hence, Af (%A) and Bf (%B) should correlate *positively*.

Computer Test: First, we arbitrarily chose the following ranges: A 3.00 – 3.01; B 7.00 – 7.01; C 3.000 – 3.001, i.e. all are very near constants; however C being closer to a fixed number than A (B). As predicted, %A correlated *negatively* with %B (Fig. 7, left panel), rho = -0.906, p<0.01, n =200. Eq. of the regression line was %B= -0.819 (0.025) * %A + 72.76 (0.59). We next made the ranges to be A 3.000 - 3.001, B 7.000 - 7.001, and C 3.0 - 3.1, i.e. A (B) were made closer to fixed numbers than C. As shown in Fig. 7, right panel, we then obtained a strong *positive* correlation between %A and %B (rho =1.000, p<0.01, n =200.

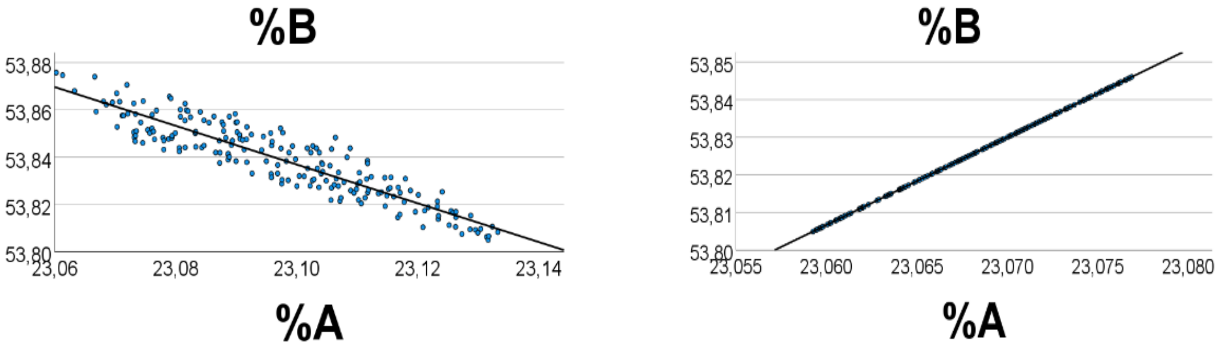

Figure 7: *Association between %A and %B, when ranges of A and B differ appreciably. The figure relates to the equation %A + %B + %C = 100, see text. Random numbers with uniform distribution.* ***Left panel: Range*** *of A 3.00 - 3.01, B 7.00 - 7.01, and C 3.000 - 3.001. %A vs. %B: rho = -0.906; p<0.01, n =200.* ***Right panel:*** *A 3.000 - 3.001, B 7.000 - 7.001, and C 3.00 - 3.10. %A vs. %B: rho =1.000; p<0.01, n =200.*

We finally used these ranges: A 200.0 - *200.2; B 30.0 - 30.3; and C 20.00 - 20.02*. The relative variability was for A (0.2·100)/200 = **0.1%;** for B (0.3·100)/30 =**1.0%;** and for C (0.02·100)/20 = **0.1%.** This means that A and C were closer to fixed numbers than B. In this latter case, Af = 1/(1 + B/A + C/A). However, since A and C are closer than B to fixed numbers, we may write ***Af = 1/(q + B/A),*** where q = 1 + C/A is close to be a constant. Similarly, Bf = B/(A+B+C) = 1/(1 + A/B + C/B) may be written ***Bf = 1/(u + A/B)*** where **u** is close to be a constant, (1 + C/B). Therefore, Af - as well as Bf - are mainly dependent upon changes in **B**. However, Af should *decrease* as B runs from lowest to highest value, whereas Bf should *increase*. Accordingly, *Af and Bf should relate inversely*. Cf = C/(A + B + C)= 1/(1 + A/C + B/C). Since A and C are closer than B to be constants, we may write ***Cf = 1/(v + B/C)*** where v = 1 + A/C , i.e. near a constant. Thus, Cf is also mainly dependent upon **B**, and Cf should *decrease* as B increases from lowest to highest value. Since Af as well as and Cf are expected to decrease as B increases, Af and Cf should correlate positively.

Computer Test: As expected, we found a positive correlation between %A and %C (rho = 0.570), and a negative association between %B and %A(%C), rho = -0.994 (-0.648); p<0.01 for all, n = 200. Furthermore, B correlated negatively with %A (%C): rho = -0.985 (-0.683), and positively with %B (rho =0.996), p<0.01 for all, n = 200. The scatterplot of %A vs. %B is shown in Fig. 8.

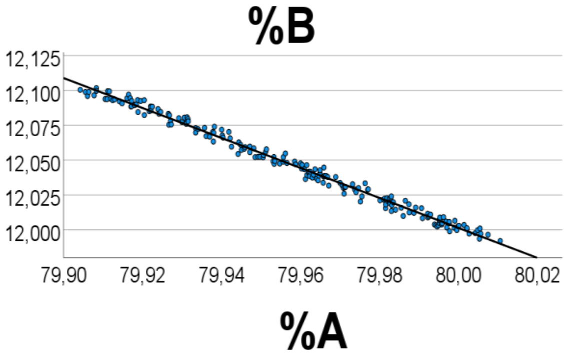

Figure 8: *Association between %A and %B). The figure relates to the equation %A + %B + %C = 100, see text. Ranges were A 200.0 - 200.2; B 30.0 - 30.3; and C 20.00 - 20.02; rho = -0.994, p<0.001, n = 200.*

These experiments illustrate that, if all of the three variables are close to be fixed numbers, then *closeness to be constants (zero variability*) appears to govern the correlation outcome. Anyhow, distributions (ranges) of the variables seem crucial for making strong positive or negative *Distribution Dependent Correlations.*

Turning Point

The experiments above show that we may achieve that a positive (negative) correlation between relative amounts changes to become negative (positive), in response to altering ranges of the variables. Below, we consider further how ranges may govern direction and strength of correlations.

Based upon the equation %A + %B +%C = 100, i.e. %B = -%A + (100 - %C), we suggested above that %A should correlate positively with %B if %C values were very high, and negatively if %C values were very low. This reasoning implies that somewhere in-between these extreme conditions, we should expect to find e.g. that a positive (negative) %A vs. %B association would turn to become negative (positive), in response to decreasing (increasing the values of %C. We define this condition the ***Turning Point***.

Computer Test: To possibly find a Turning Point from positive to negative correlations, through gradually *decreasing* %C, we started with a condition giving a positive %A vs. %B association, i.e. A 3.0 - 3.3, B 8.0 - 8.8, and **C 0 - 100** (Fig.9, left panel). To decrease C gradually, we narrowed the C range towards the lower limit, while keeping the ranges of A and B. With ***C 0 - 10*** (Fig. 9, middle panel) **,** we found *rho = 0.976* for %A vs. %B; %C vs. %A (%B), rho = -0.987 (-0.998); S correlated negatively with %A and %B; rho =-0.989 and -0.988, respectively, and positively with % C (rho = 0.993), p<0.01 for all, n = 200. The scatterplots for the %A vs. %B association became poorer with ***C range 0 - 10*** as compared with **C range 0 - 100** (Fig 9. middle and left panels). With C 0 - 10, *skewness* of the %A, %B, and %C distributions was *0.20, 0.32, and - 0.30*, respectively (SE of skewness was 0.17 for all). Quartiles of the %A, %B, and %C distributions were *16.9, 19.8, 22.4; 45.3, 52.5, 61.1;* ***16.4, 27.8, 38.1***. Thus, skewness of relative amounts had decreased appreciably, i.e. from *1.79, 1.79, and -1.78*, for %A, %B, and %C, respectively, observed with C-range 0 - 100. Furthermore, the %C distribution had moved towards appreciably lower values i.e. quartiles of the %C histogram were ***69.2, 81.8, and 87.0,*** respectively when the C - range was 0 - 100. These experiments show that the positive association between %A and %B prevailed in spite of moving %C far away from 100%.

We next decreased %C further by narrowing the ***C range to be 0 - 1***. This narrowing was accompanied by a very poor %A vs. %B scatterplot (Fig. 9, right panel) and correlation coefficient poor (rho = 0.268, p<0.01, n =200). In this case, the %C distribution had moved even more towards low values, quartiles of %C being ***2.1, 3.9, and 6.0***, respectively.

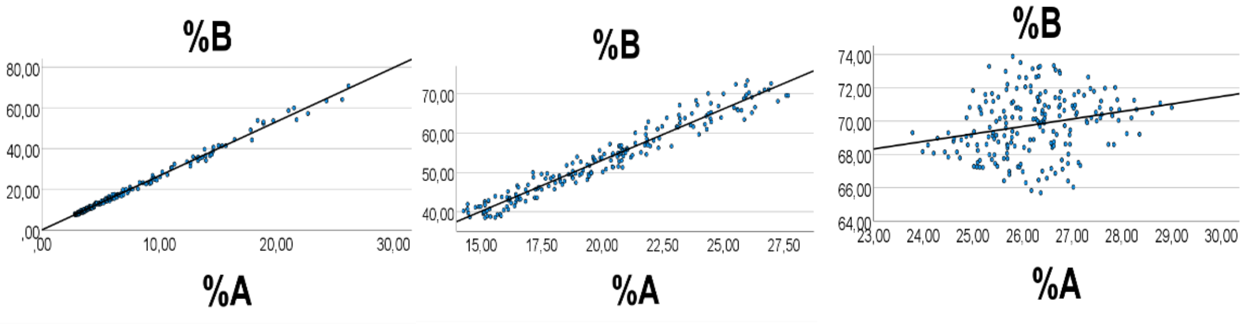

Figure 9: *Association between %A and %B (left panel); the figure relates to the equation %A + %B + %C = 100, see text. Random numbers (n = 200) with uniform distribution were generated.* ***Left panel:*** *ranges were A 3 - 3.3; B 8- 8.8,* **C 0 - 100. Middle panel** *A 3.0 - 3.3; B 8.0 - 8.8,* **C 0 - 10.** ***Right panel:*** *A 3.0 – 3.3; B 8.0 - 8.8, C 0 - 1. %A vs. %B (left/middle/right): rho = 0.994/0.975/0.268 p<0.01 for all.*

We continued narrowing the ***C range to be 0 - 0.1***, while still keeping A 3.0 - 3.3, and B 8.0 - 8.8. As shown in Fig. 10, we then obtained a strong ***negative*** association (rho = -0.946, p<0.01) between %A and %B, showing that we had passed through the *Turning Point*. Thus, by gradually decreasing the %C - values, the %A vs. %B association had turned from being strongly positive to becoming strongly negative. Conceivably, alterations in the C range caused changes in histograms of %A and %B as well. As predicted, skewness of the %A (%B, %C) histograms changed appreciably in response to narrowing the C range. With C 0 - 0.1, we found close to a normal distribution, i.e. skewness of %A, %B, and %C being *-0.060, -0.086, and 0.110*, respectively. In this case, %C quartiles had decreased to ***0.2, 0.4, and 0.6***, respectively, i.e. approaching a situation where we might roughly approximate the equation %B = -%A + (100 - %C) to %B = -%A + 100, showing a negative %A vs. %B association.

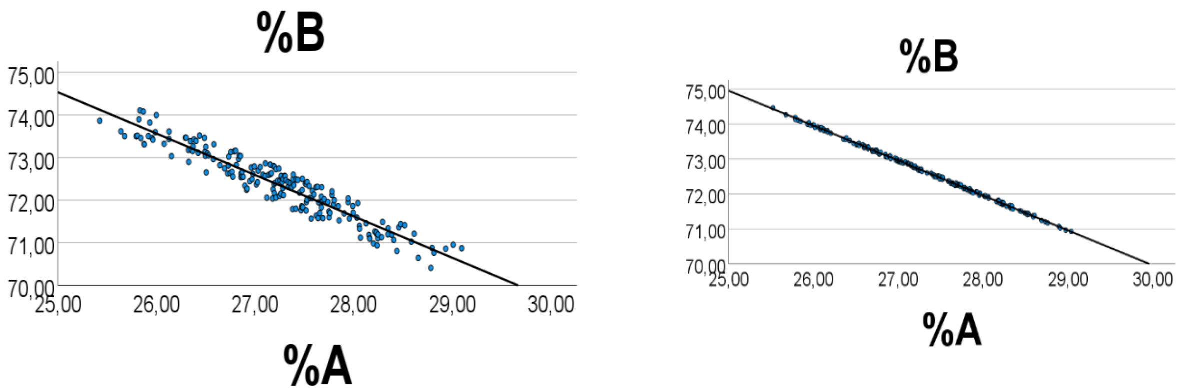

Figure 10: *Association between %A and %B; the figure relates to the equation %A + %B + %C = 100, when ranges of C is varied (see text). Random numbers (n = 200) with uniform distribution were generated; left panel: A 3.0 - 3.3; B 8-– 8.8,* ***C 0 - 0.1.*** *%A vs. %B, rho = -0.946, p<0.01; right panel: A 3.0 - 3.3; B 8.0 - 8.8,*

C 0 - 0.01. *%A vs. %B, rho = -0.999, p<0.01.*

We finally narrowed the C range to **0 - 0.01** (Fig. 10, right panel). In this case, we observed that the scatterplot improved further; rho = -0.999, $p<0.01$, n=200. Skewness of %A, %B, and %C were *-0.074, 0.064, and -0.069,* respectively. %C quartiles had decreased to ***0.02, 0.04, and 0.06.***

These examples seem to support the idea that there should be a *Turning Point* where a positive correlation between A- and B percentages of S turns to become negative, in response to progressively decreasing the C - range. Furthermore, the experiments indicate that we may reach the *Turning Point* as skewness of the %C distribution approaches zero. Moreover, the correlations attenuate when approaching the *Turning Point.*

Skewness of the distributions of relative amounts

In many of the previous examples, we observed that distributions of A- and B- percentages of S had positive skewness, and %C had negative skewness, raising the question of how to explain skewness of the relative amounts of A, B, and C.

We consider again $S = A + B + C$, where *A and B have narrow ranges* relative to the C-range. The A-percentage of S, $\%A = 100A/(A + B + C) = 100/(1 + B/A + C/A) = 100/(t + C/A)$, where $t = 1 + B/A$ is close to a constant. Thus, C is the governor of %A, irrespective of where we place A and B on the positive scale. Furthermore, %A should relate *inversely* to C, since the denominator increases as C increases from lowest to highest value.

Additionally, we should expect *a low number* of C-values to be associated with each % A- unit decrease in the *upper* end of the %A scale. In contrast, there should be a *high* number of C values associated with each %A–unit decrease in the lower end of the %A– scale. We should, accordingly, anticipate a positively skewed histogram of %A, keeping in mind that the number of C values is an estimate of the number of cases (points on the scatterplot). The same reasoning goes for %B, which should have a positively skewed distribution as well. The following example might serve to illustrate this general suggestion, using the ranges: *A 1.0 - 1.1, B 2.0 - 2.1, and C 1 - 100*. Thus, with $C = 1$, $\%A = 100\cdot A/(A + B + C) = 100\cdot 1/(1 + 2 + 1) = 100/4 =$ ***25.0%.*** If increasing C one unit, $\%A = 100/(1 + 2 + 2) =$ ***20.0%.*** However, a similar one-unit increase in C at the *upper* end of the C-range results in a much smaller decrease in %A, i.e. from $100/(1 + 2 + 99) =$ ***0.98%*** to $100/(1 + 2 + 100) =$ ***0.97%.*** Accordingly, the curvilinear negative association between %A and C should have *the concave upwards*. Similar considerations should apply to the negative %B vs. C association. It is beyond the scope of this article to discuss this outcome in more detail, mathematically.

The C-*percentage* of S is $\%C = 100\cdot C/(A + B + C) = 100/(1 + z/C)$ where $z = (A + B)$ is close to a fixed number. Thus, %C should increase when increasing C from lowest to highest value. However, this effect should attenuate with increasing C-values, showing a positive curvilinear relationship between percent C and C, with the *concave downwards*. For example, when C goes from 1 to 2, then %C increases from approximately $100\cdot 1/(1 + 2 + 1) =$ ***25%*** to $100\cdot 2/(1 + 2 + 2) =$ ***40%.*** A similar one-unit increase in C at the upper end of the C-range, i.e. from $C = 99$ to $C = 100$, is associated with a very small increase in %C, i.e. from $100\cdot 99/(1 + 2 + 99) =$ ***97.06%*** to $100\cdot 100/(1 + 2 + 100) =$ ***97.09%.*** This reasoning suggests that the concave should be downwards for the positive relationship between %C and C. The finding that C is negatively associated with %A and %B explains that these percentages of S are positively associated. Furthermore, since %C is positively associated with C, percent C should be negatively related to %A and %B.

Accordingly, we should expect a low number of C-values to be associated with each %C - unit increase in the *lower* end of the %C scale. In contrast, there should be a high number of C values associated with each %C – unit increase in the upper end of the %C – scale. Hence, the number of C values (each of which representing a case) *increases* for each unit increase in %C. Thus, there should be *a negatively skewed histogram of %C* (Fig. 11, lower panels). Since the C - range is the governor of skewness, an increase (decrease) in this range should increase (decrease) skewness of the %A(%B, %C) histograms.

Thus, on the condition that *ranges of A and B are very narrow relative to that of C*, the C-range will determine *skewness* of the %A-, %B-, and %C- distributions, as well as *correlations* between the percentages. Indeed, we may consider skewness as a marker of the current correlations [22].

When making a scatterplot of e.g. the association between C and %A, the independent variable (C) regularly appears on the abscissa. However, to improve the illustration of how the C-range governs skewness of the distributions of relative amounts, in the next example we have switched axes, so that percentages appear on the abscissa, and C values on the ordinate (Fig. 11). We used the following ranges: *A 2.00 - 2.02, B 3.00 - 3.03, and C 1 - 100*. Thus, with %A = 2, and using the definition %A =

100·A/(A + B + C), the equation would be 2 = *100·2/(2 + 3 + C); i.e.* **C = 95.0.** If increasing %A by one unit, we have 3 = 200*/(5 + C), i.e.* **C = 61.7.** A similar one-unit increase at higher %A values is associated with a lower decrease in C. For example, if %A goes from e.g. 20% to 21%, C decreases, from **5.0 to 4.5**, indicating that the curvilinear negative association between %A (%B) and C should have *the concave upwards* (Fig. 11 , *top panels).*

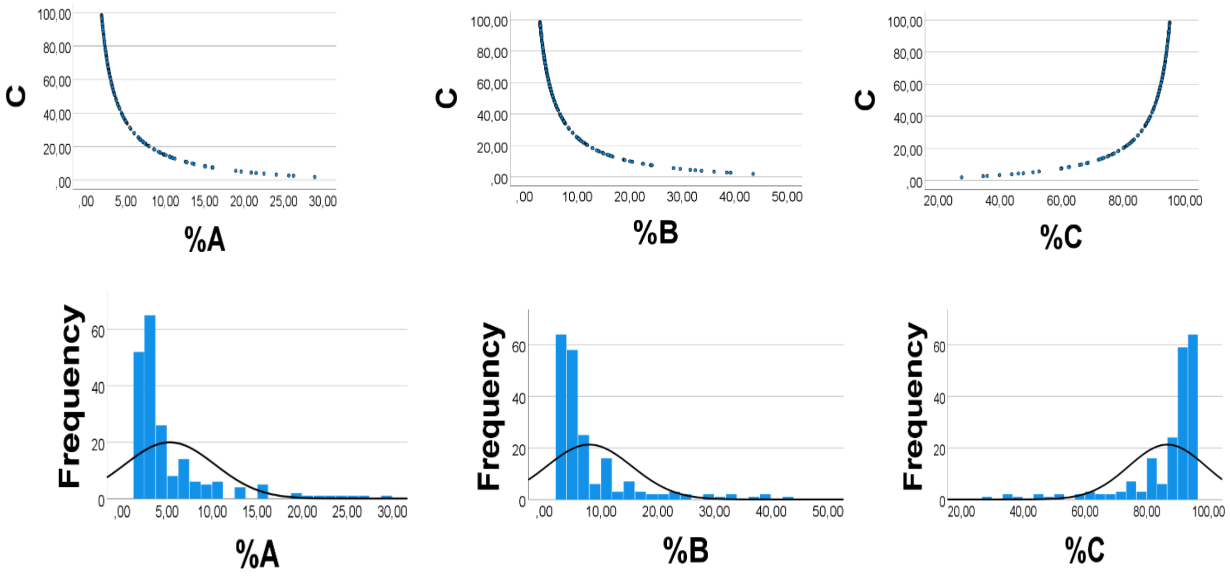

Figure 11: *Association between C and %A(%B, %C),* **upper panels;** *and histograms of %A, %B, and %C* **(lower panels).** *The figure relates to the equation %A + %B + %C = 100, see text. The ranges were, A 2.00 – 2.02; B 3.00 – 3.03; and C 1 – 100, n = 200 (uniformly distributed random numbers).*

As shown in Fig. 11 (top panels), C *decreases as* %A (%B) increase, but the C - decrease *per unit* increase in %A (%B) attenuates with increasing %A (%B). In contrast, increasing values of C is associated with *increasing* values of %C (Fig. 11, top panel, right). Additionally, the C - increase *per unit increase in %C* rises with increasing %C values. Accordingly, the %A and %B distributions are positively skewed, whereas the %C histogram has negative skewness (Fig.11, lower panels).

This example shows that, in the current context, alterations in C values are crucial to explain skewness of the relative amounts, i.e. a broad (narrow) range of C gives high (low) skewness of %C (%A, %B). Furthermore, high (low) variability of C also promotes strong (weak) correlations between %A and %B. In the current example, we may apply both the equation of a straight line, and the relationship between S and percentages of S, to explain correlations between the relative amounts.

Ranges, Skewness, and Correlations [23]

It would appear, accordingly, that *ranges* of the variables are crucial for obtaining strong correlations between percentages of three variables, and for making skewness of the relative amounts. Skewness of the percentages may decrease appreciably if gradually *narrowing* the C-range, and so will strength of the %A vs. %B correlation. We here show the outcome with narrowing to ***C 1 - 2.*** First, we arbitrarily choose r*anges of A to be 2.00 - 2.02, B 3.00 - 3.03, and C* ***1 - 100****. We found skewness of %A, %B, and %C to be* ***2.10, 2.10, and -2.10***, respectively, i.e. high positive skewness of %A and %B, and high negative skewness of %C. There was a perfect positive correlation (rho =1.000) between %A and %B, and a perfect negative (rho = -1.000) association between %C and %A (%B). However, with *C range narrowed to 1-2*, skewness of %A, %B, and %C had attenuated to *0.21, 0.20, and -0.20*. However, the strong positive %A vs. %B correlation prevailed (rho = 1.000, Fig. 12, left panel), and also the strong negative association between %C and %A (%B), rho = -1.000; -1.000), p<0.01 for all, n = 200. Quartiles of the %A, %B, and %C distributions were *29.56, 30.60, 31.96; 44.32, 45.96, 47.86; 20.18, 23.38, 26.13*, respectively. Thus, in this case, the strong positive %A vs. %B correlation is not well explained by very high %C values, with reference to the equation %B = -%A + (100 - %C), see above. Alternatively, we may consider fractions of S: *Af = A/(A + B + C) = 1/(1 + B/A + C/A).* Since A and B are close to be constants, we may write Af = ***1/(k + C/A)*** where k = 1 + B/A, which is close to a constant. A similar reasoning goes for *Bf = B/(A + B + C) = 1/(1 + A/B + C/B) = 1/(t + C/B)*, where t = 1 + A/B. Thus, Af (and Bf) should decrease as C increases, thereby explaining the positive association between %A and %B. Furthermore, Cf = C/(A + B + C) = 1/[1 + (A + B)/C], which may be approximated to Cf = **1/(1 + u/C**] where u = A + B, i.e. near a constant. Thus, Cf increases as C goes from lowest to highest value. The relationships between C and %A (%B, %C) explain the correlations between the percentages.

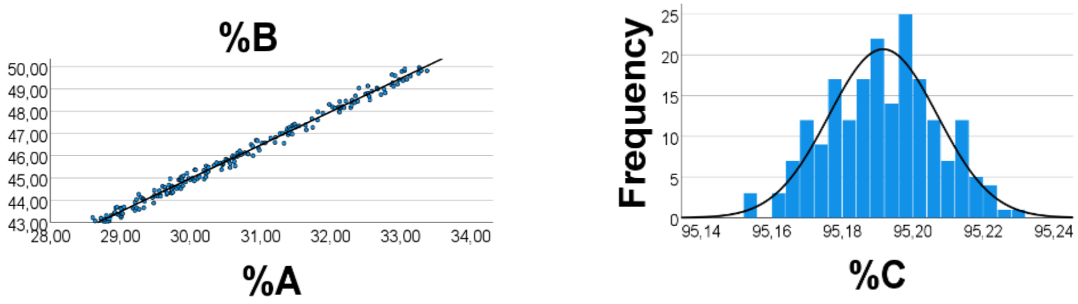

Figure 12: *Association between %A and %B), and histogram of %C. The figure relates to the equation %A + %B + %C = 100, see text. Ranges were A 2.00 – 2.02; B 3.00 – 3.03; and C 1 – 2; rho = -1.000, p<0.01, n = 200 (uniformly distributed random numbers).*

In brief: With three positive scale variables (A, B, C), two of which *(A, B) having narrow ranges relative to the C-range*, we might expect high negative skewness of %C, high positive skewness of %A and % B, strong positive %A vs. %B correlation, and a strong negative correlation between %C and %A (%B). With decreasing C - variability, we should find attenuated skewness of %A, %B, and %C, and decreased strength of the correlation between %A and %B.

We carried out 10 repeats [23] of a condition expected to give a near-symmetrical distribution of percentages of A, B, and C: i.e. A and B 0.10 - 0.15; C 9 - 10 (n = 200 in each repeat). Coefficients of variation (CV) were for rho (%A vs. %B): *119.6%;* for Skewness of %C: *- 352.5%;* for Q3 of %C: *0.02%.* Thus, Q3 of the %C histogram had low CV, as compared with CV of *skewness* of the %C histogram.

We did some experiments to examine how skewness relate to the correlation between %A, %B, and %C, as obtained by changing A; B, and C ranges in many ways [23]. The results are summarized in Fig. 13, where skewness of %C is plotted against Spearman's rho for the association between %A and %B. Each of the 49 points represents 200 random number "cases", computed with particular ranges for A, B, and C. The relationship between skewness of %C and correlation between %A and %B seemed like a mirror image of a sigmoidal scatter of points (Fig. 13). Thus, with increasing negative (positive) skewness we observed a progressive improvement of the positive (negative) correlation between percentages of A and B. Similar relationships were obtained when skewness of the distribution of % A (%B) was plotted against rho for the correlation between %B vs. %C (%A vs. %C), not shown. *We emphasize that we had to use narrow ranges of A and B, and broad C-range to obtain the high positive %A vs. %B correlations (Fig.13, left part of the scatterplot). Similarly, to make strong negative correlations between %A and %B (Fig.13, right part), we used narrow ranges of A and C (or of B and C), and broad range of B (A). The strong negative %A vs. %B association, found with a very narrow C-range, was not, however, accompanied with high positive %C skewness, as discussed above. In this latter case, we have a situation where one of the variables approaches a fixed number. Thus, we have close to a two-variable condition, making the remaining two variables relating inversely.*

The figure illustrates that we may achieve the *Turning Point* when skewness of the %C distribution approaches zero. It would appear, that when skewness of the %C distribution approaches zero (symmetrical histogram), then rho (%A vs. %B) varies greatly in response to minor changes in skewness of %C. Thus, close to a symmetrical distribution of the histogram of %C, the correlation between percentages of the two remaining variables (A and B) is very sensitive to changes in skewness of %C. On the other hand, with very high (positive or negative) skewness of the %C distribution, only small changes in the size of Spearman's rho for the %A vs. %B correlation are allowed. *Thus, in particular cases, skewness of the %C distribution seems to relate to the correlation between A (B) percentages of S. However, as pointed out above, we may achieve strong negative correlations without any accompanying major skewness. Indeed, ranges of the variables seem to be the crucial force both for making skewness and correlations.*

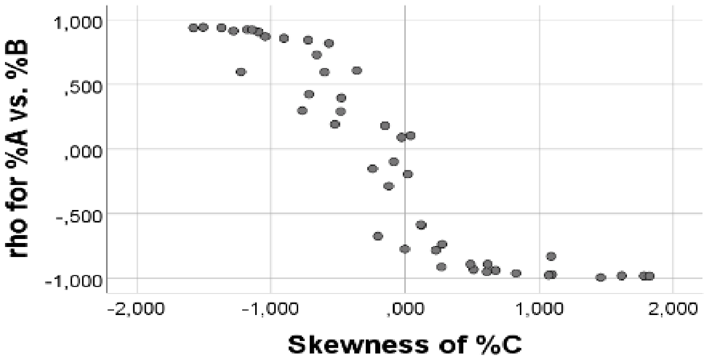

Figure 13: *Association between skewness of the %C histogram and Spearman's rho for the correlation between percentages of the remaining two variables (A and B). The figure relates to the equation %A + %B + %C =100, or %B = - %A + (100 - %C), see text. We made the figure using uniformly distributed random numbers of A, B, and C. Each of the 49 points represents 200 random number "cases", computed with particular ranges for A, B, and C. From Høstmark 2019, JNFP.*

More than three variables

When computing fractions (percentages) in the present work, we used sum (S) of two or three variables. We may raise the question of what happens to the association between e.g. %A and % B if the denominator (S) includes more than 3 variables. In general, with a large number of positive scale variables (A, B, C……..), their sum is S = A + B + C…., giving the equation: %A + %B + %C + ……= 100, This equation may be simplified to apparently involve 3 variables: %A + %B + %R = 100, if R is the sum of all variables, except A and B. With reference to the outcome presented above, we should expect that %A and %B correlate positively, if both of these latter variables have very narrow ranges relative to R. In fact, the association should prevail if broadening the R-range. Previous computer experiments were in favor of this reasoning [23].

Computer test: Above, we made sum of random numbers representing OA, AA, and EPA only; their ranges being 1- 9, 0.3 - 0.4, and 0.1 - 0.2, respectively. In a repeat of this experiment, we obtained %AA' vs. %EPA': rho = 0.867, p<0.001, n =200. Equation of regression line was *%AA' = 1.78 (0.07) · %EPA' + 0.98 (0.15).* We then included altogether 12 fatty acids in the denominator, i.e. the total range of the fatty acids was increased to 3 -15. The %AA' vs. %EPA' association did not change much: rho = 0.856, p<0.001, n =200. Equation of regression line changed to be *%AA' = 1.96 (0.09) · %EPA' + 0.85 (0.18).* This outcome seems in favor of using the "three-variable approach".

Intended Ranges and Distribution Dependent Correlations: Examples from Physiology

In the first part of this article, we presented some general considerations of how percentages of the same sum should relate, and showed computer experiments to support the reasoning. Some of the examples illustrated mathematical points, without necessarily relating to physiology. Below, we show some examples from physiology, where "***Intended Ranges***" could represent evolutionary selections, serving to achieve that relative amounts must become positively or negatively associated, mathematically [28].

In biology, we may encounter ranges presumably representing evolutionary selection. For example, body temperature, weight, height, heart rate, blood pressure, organ sizes, and amounts of many tissue and blood factors, such as electrolytes, glucose, lipoproteins, and fatty acids seem to exist within particular ranges. The spread of a physiological variable could mean that nature failed when trying to hit the target (e.g. a certain level of a blood or tissue variable). In laboratory medicine, we regularly define the "normal range" of a variable as the mean value ± 2SD, based upon data found in healthy subjects. A complete different view is that ranges in many contexts depend on *evolutionary selection*.

Thus, *variation* is a central concept. To assess variability, we regularly use the range of the variable, as well as the interquartile range, and standard deviation. In general, we may categorize the spread of a biological variable as true biological, pre-analytic, and analytic. Alternatively, we may divide into *common cause* variation and *assignable* variation [31], i.e. unexpected large variation; caused by for example computer crash, corona virus infection, or problems with the supply of water and food. We should try to avoid these negative types of variation.

In contrast to this, some types of biological variability could represent *wanted* variation, appearing during evolutionary selection. This suggested biologically ***intended****, advantageous variation should go from lower to upper limits,* developed through evolutionary selection. We suggest that evolution possibly might have selected the necessary regulatory mechanisms to achieve the intended limits as well, for example effected through synthesis and metabolic regulation of key enzymes. Notably, the regulatory processes governing these limits would be subject to common cause variation. We hypothesize that ranges of fatty acids in tissues and blood, as well as ranges of white blood cell counts, could be examples of intended ranges in biology (*vide infra*).

As shown mathematically in the first part of this article, fractions of variables might correlate positively or negatively, on the conditions that their concentrations exist within particular ranges. In physiology, ranges could be intended ones. Conceivably, the *measured* ranges will also reflect common cause variability.

The many causes of error could make it hard to detect and appreciate the suggested biological, *intended variability*, e.g. of particular fatty acid concentrations, and frequency distributions and scatterplots should regularly be observed, to find outliers. To detect the suggested true, *intended ranges*, we should have very low external variability. Nevertheless, DDC should exist, if the variables in question do have the required concentration ranges. We emphasize that the chicken population referred to below, was very homogeneous, genetically and environmentally, thereby offering an excellent opportunity to evaluate how the suggested *intended ranges* might influence associations between relative amounts of fatty acids.

Brief Synopsis of Previous Articles Related to Distribution Dependent Correlations

From the papers below, we present some main results related to DDC:

 # 1: Høstmark AT, Haug A (2018) The Fatty acid Distribution per se Explains Why Percentages of Eicosapentaenoic Acid (20:5 n3) and Arachidonic Acid (20:4 n6) are Positively Associated; a Novel Regulatory Mechanism? J Nutr Diet Suppl 2 (1):103

Since EPA and AA are metabolic antagonists, in this paper we raised the question of whether their *relative amounts* might correlate. It turned out that %AA and %EPA were *positively* related, as observed in chicken breast muscle lipids [14], raising the question of how to explain the association. To circumvent the numerous physiological regulatory mechanisms, I started using *random numbers* in lieu of the true values, however *sampled with the true ranges* of AA and EPA. Surprising at the time, the correlation outcomes were similar with these random numbers and with the true values [14]. Furthermore, even slight alterations of the random number ranges had major effects upon the correlation outcomes, as documented by appreciably altered scatterplots and correlation coefficients. These findings suggested the concept of *Distribution Dependent Regulation (Correlations)*. The finding that particular ranges (distributions) of variables were crucial for making strong correlations between their relative amounts seemed at the time to be a novel regulatory mechanism.

 # 2: Høstmark AT (2019) *Associations Between %AA (20:4 n6) and Percentages of EPA (20:5 n3), DPA (22:5 n3), and DHA (22:6 n3) Are Distribution Dependent in Breast Muscle Lipids of Chickens. J Nutr Diet Suppl 3(1):103*

Since the ranges of DPA (22:5 n3) and DHA (22:6 n3) are narrow, we investigated whether their relative amounts might be positively associated with %AA, and if the ranges could explain the correlation outcome. As expected, corresponding scatterplots (correlation coefficients) were generally very similar, irrespective of using true (measured) values of the fatty acids, or random numbers, *if we generated the random numbers with the true ranges*. However, altering the ranges had appreciable effects upon the associations. Thus, narrowing (broadening) the ranges improved (made poorer) the scatterplots and correlation coefficients.

To explain the correlation outcome, we first utilized the equation of a straight line *($y = ax + b$)*. Ranges of DPA, DHA, and AA were narrow compared with range of the sum of the remaining fatty acids (R). Thus, AA had range 0.25 - 0.42, DPA 0.21-0.43, and DHA 0.11-0.32 g/kg. In contrast, R had range 5 - 15 g/kg. Therefore, the equation %AA = -%DHA + (100 - %R) would approach %AA = % DHA, due to the high values of %R. Quartiles of the %R distribution were *93.1, 95.0, and 96.0%*, respectively. Hence, %AA and %DHA should correlate positively. Using random numbers with the true ranges, we found rho = 0.599 for %EPA' vs. %AA' (the substitute EPA and AA variables are named EPA' and AA'). %R' vs. %EPA' (%AA'): rho = -0.858 (-0.919), p<0.01 for all, n =200. These results seems to support our previous finding that %AA correlated positively with %EPA, *because of the particular ranges*.

 # 3: Høstmark AT (2019) Associations between Percentages of Scale Variables, as Related to Distributions. J Nutr Diet Suppl 3 (1):104.

This article summarizes some of my opinions at the time, concerning DDC. For example, I suggested this rule : " With 3 scale variables, *two of which having low-number distribution as compared with the third variable, we might expect a positive association between percentages of the low-number variables, and a negative association between percentage of the high-number variable and percentage of each of the low-number variables."* Another rule was: *"Distributions of 3 scale variables can be manipulated so as to obtain a Turning Point, i.e. a situation where a positive (negative) association between percentages of two of them turns to become negative (positive)".* Since the explanation of the rules was preliminary at the time, we tried to improve the rules in the succeeding articles

 # 4: Høstmark AT, Haug A (2019) High Variability of Oleic Acid (OA, 18:1 c9) improves the Positive Association between %EPA (20:5 n3) and %AA (20:4 n6). J Nutr Diet Suppl 3 (1):106.

 # 5: Høstmark AT, Haug A (2019) The inverse association between relative abundances of oleic acid and arachidonic acid: a case of distribution dependent regulation? Lipids in Heath and Disease 18:123

In these two papers, we discuss further, how the inverse %OA vs %AA association observed in chicken muscle, may be explained, as well as how oleic acid (OA) might influence the positive correlation between %AA vs. %EPA. We continued to use random numbers of OA (marked OA') in lieu of the measured values of the fatty acid. First, we used the true OA-range, and then gradually altered the range. Increasing (decreasing) the OA'- range, in computer experiments, towards higher (lower) values improved (made poorer) the association between relative amounts of AA' and EPA'. In general, we obtained similar results (scatterplots and correlation coefficients) with the true values of the fatty acids, and with random numbers in lieu of the true values, if the random numbers had the true ranges [17]. Furthermore, we found similar %AA vs. %EPA correlations, as well as correlations between %OA and %AA (%EPA) - irrespective of using *sum of all fatty acids* in the denominator when computing the percentages, or the sum of AA, EPA, and OA only.

We noticed that ranges of AA, EPA, and OA differed appreciably, being 0.25 - 0.42; 0.13 - 0.24; and 1.04 - 8.56 g/kg, respectively. Corresponding coefficients of variation (CV) were *9.7, 11.1, and 43.9%.* Using substitute random numbers instead of the true values, we broadened the OA' range towards higher values (Fig. 14). This increase in OA' variability greatly improved the positive correlation between %AA' and %EPA' (left panel), and was accompanied by increasingly higher values of %OA' (right panel). Interestingly, approximately at the physiological maximum value of OA (about 9 g/kg), the correlation coefficient – as well as the third quartile of the %OA distribution - levelled off rapidly.

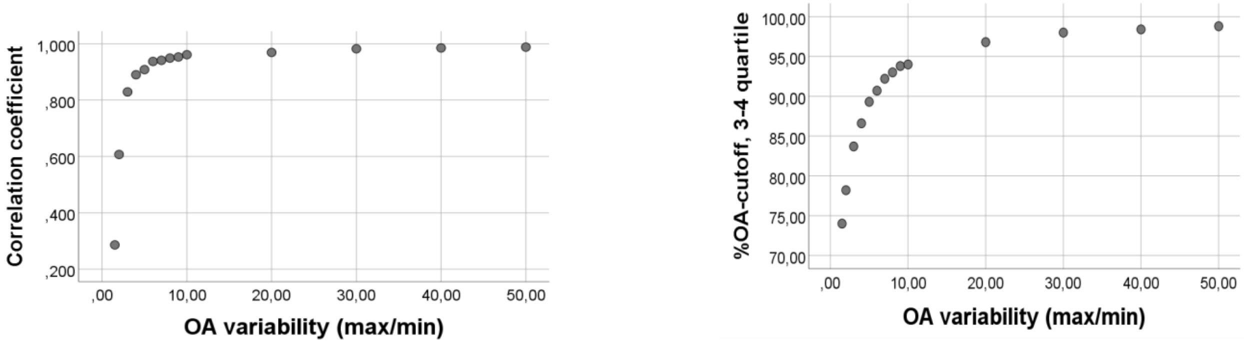

Figure 14: *Effect of increasing the OA'-range towards higher values upon the %AA' vs. %EPA' correlation coefficients (left panel), and on the third quartile of the %OA' distribution (right panel), see text. Note that we omitted the random number mark (OA') when making this figure. From Høstmark AT, Haug A (2019),*

J Nutr Diet Suppl 3:106

Below, we *extend the explanation* given previously [17] of the correlation outcome. Thus, we have a condition from physiology where two variables (i.e. AA and EPA) are closer to fixed numbers than a third variable (OA). We utilize the equation AA + EPA + OA = S, i.e. %AA + %EPA + %OA = 100, or *% AA = -%EPA + (100 -%OA*). With increasingly higher %OA' values, caused by broadening the OA'-range towards higher values, we should move towards an increasingly improved %AA' vs. % EPA' association, as discussed in the mathematical part of this article.

Rewriting the equation to *%OA' = -%AA' + (100 - %EPA'*), we might consider two approximations, to simplify the equation: 1) Thinking that the low EPA' (%EPA') *values* (as compared with OA') were approaching zero, thereby making the equation approach %OA' = -%AA' + 100, i.e. %OA' should relate negatively to %AA'. 2) The narrow EPA *range* would make (100 - % EPA) approach a constant value (**k).** In that case, the equation would be %OA = -%AA + k, which would give a negative correlation between %OA and %AA, as well. Similarly, we may rewrite the equation above to be *%OA' = -%EPA' + (100 - %AA').* Using the above simplifications, we would have *%OA' = -%EPA' + 100,* and *%OA' = -%EPA' + k.* In both cases, we should obtain a negative association between %OA' and %EPA'. In a computer test, we found that the following equations of regression lines*:1) %AA' = 1.6 · %EPA + 0.7; 2) %OA' = -2.6·%EPA' + 99.3;and 3) %OA' = -1.5·%AA' + 99.3*. For the present purpose, we omit SE values in the equations.

Alternatively, we may consider what happens to the AA'-, EPA'- , and OA'- *fractions* of S (S = OA' + AA' + EPA'), as OA' goes from the lowest to the highest value within the OA' range. The AA' fraction is AA'/(AA' + EPA' + OA') = 1/(1 + EPA'/AA' + OA'/AA'), i.e. the AA' fraction (percentage) is governed by the EPA'/AA' ratio, and by the

OA’/AA’ ratio. The EPA’/AA’ ratio has low variability as compared with the OA’/AA ‘ratio. The former one goes from 0.13/0.42 = ***0.3*** to 0.24/0.25 = **1.0**, and the latter from 1.04/0.42 = ***2.5*** to 8.56/0.25 = ***34.2***. Accordingly, the EPA’/AA’ ratio should have a much lower impact than the OA’/AA’ ratio to change the AA’ fraction. Additionally, the OA’/AA’ ratio is mainly influenced by OA’. Therefore, OA’ is the main governor of the AA’ - fraction (percentage), which should decrease as OA’ increases from the lowest to the highest value (Fig. 15). A similar reasoning goes for the EPA’ – percentage of S. Thus, %AA’ should correlate positively with %EPA’, since both fractions decrease as OA’ increases from lowest to highest value. As seen from the formulas, the decrease in %AA’ and %EPA’ per unit increase in OA’, should level off at increasing OA values, making the %AA’ (%EPA’) vs. OA’ scatterplot to have the concave upwards.

The OA’ fraction of S is OA’/(AA’ + EPA’ + OA’) = 1/[1 + (AA’ + EPA’)/OA’]. This fraction should decrease as (AA’ + EPA’) increases, and increase as OA goes from lowest to highest value. However, the denominator should have a greater impact than the numerator on the (AA’ + EPA’)/OA’ ratio, making the OA’ fraction (percentage) to *increase* as OA’ goes from lowest to highest value. In addition, the %OA’ vs. OA’ association should be curvilinear with the concave downwards, as seen from the definition of the OA’ fraction. Interestingly, when approaching the measured, upper OA’ level (i.e. about 9 g/kg in chicken breast muscle, Fig. 15), we apparently should expect only a small further decrease in %AA’ and %EPA’, and a weak further increase in %OA’, in response to increasing the OA’ range towards higher values. Computer experiments verified this suggestion (not shown). The associations with OA’ appear to be the basic explanation of the positive %AA’ vs. %EPA’ association, and the negative %OA’ vs. %AA’ (%EPA’) correlation. Since OA’ is positively related to %OA’, but negatively to %AA’ and %EPA’, it follows that %OA’ should relate negatively to %AA’ (%EPA’). %AA’ and %EPA’’ should be positively correlated, since both relate negatively to OA’. In a computer test, we found Spearman’s rho = -0.946 (-0.976) for %OA’ vs. %AA’ (%EPA); %OA’ vs. OA’, rho = 0.964, p<0.01, n =200, Fig. , right panel); p<0.01 for all, n = 200.

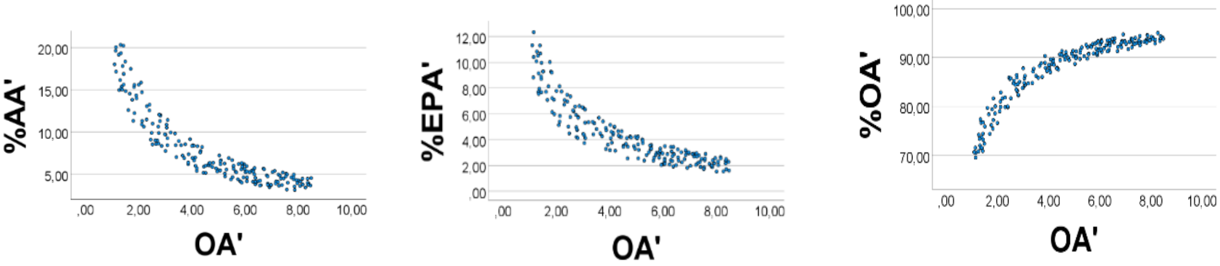

Figure 15: *Relationship between relative amounts of AA (EPA, OA) and OA, estimated by random number substitutes (marked EPA’, AA’, and OA’), see text. We generated uniformly distributed random numbers (n =200) with the true ranges, i.e. for EPA, 0.13 - 0.24 g/kg; for AA, 0.25 - 0.42 g/kg; and for OA, 1.04 - 8.56 g/kg. Spearman’s rho was -0.928 (left), -0.936 (middle), and 0.964 (right).*

Accordingly, *ranges (distributions*) of variables such as OA, AA, and EPA seem to govern associations between their relative amounts. Accordingly, some ranges, such as those of some fatty acids, could be examples of ***Intended Ranges***, i.e. ranges arising through evolutionary selection, possibly selected to ensure that associations between relative amounts must become either positive or negative, mathematically. For this phenomenon, I previously suggested the name ***Distribution Dependent Correlations (- Regulation).*** Possibly, the suggested selection serves to improve the balance between metabolites, as commented further, below.

 # 6: Høstmark AT (2019) Body Fatty Acids, Nutrition, and Health: Is Skewness of Distributions a Mediator of Correlations? JNFP 2 (1); DOI: 10.31579/2637-8914/009

In this paper, we raised the question of 1) how *skewness* of relative amounts of three scale variables (A, B, C) is brought about, and 2) whether *skewness* of the %A, %B, and %C histograms relate to correlation between the relative amounts. Particular focus was upon the relationship between %A and %B in response to altering the C - range. We succeeded in making a scatterplot of %C skewness (abscissa) against rho for %A vs %B association (ordinate), resembling a mirror image of a sigmoidal curve, as discussed in the first part of this review.

 # 7: Høstmark AT, Haug A (2019) Alpha Linolenic Acid Variability Influences the Positive Association between %Eicosapentaenoic Acid and % Arachidonic Acid in Chicken Lipids. JNFP 2 (2); DOI: 10.31579/2637-8914/016

The fact that ALA (18:3 n3) is precursor of EPA (20:5 n3) and DHA (22:6 n3) [2] could at least partly explain alleged health effects. In chicken breast muscle, we observed that the range of ALA (i.e. 0.12 - 2.40 g/kg) was appreciably broader than the ranges of AA (0.25 - 0.42 g/kg) and EPA (0.13-0.24 g/kg); coefficients of variation being *60.4, 9.4, and 11.7%,* respectively [29]. Thus, again we have two variables (AA and EPA) with narrow ranges relative to a third one (ALA), making the suggestion that %AA and %EPA should correlate positively, whereas %ALA should relate negatively to %AA (%EPA). Additionally, we should expect that ranges would influence skewness of the %ALA distribution. Indeed, with true values of the fatty acids, and with substitute random numbers, generated within the true ranges, the correlation outcomes were similar, as predicted [29].

 # 8: Høstmark AT, Haug A (2020) Relative Amounts of Eicosanoid and Docosanoid Precursor Fatty Acids Are Positively Associated: A Distribution Dependent Regulation

In this work, we investigated whether relative amounts of eicosanoid and docosanoid precursors did correlate positively. The study seemed to verify this hypothesis. Again, the ranges of the variables seemed crucial to explain the correlation outcomes. Thus, we obtained qualitatively similar corresponding results with true values, and when using substitute, random numbers in lieu of the fatty acids, on the condition that the random numbers had the true ranges. Additionally, when we hypothetically altered the ranges in computer experiments, we observed appreciable changes in the strength of the associations. Thus, there seems to be *distribution dependent positive associations* between eicosanoid and docosanoid fatty acid precursor percentages, possibly serving to ensure balance between effects of important regulatory molecules in physiology.

 # 9: Høstmark AT, Haug A (2020) Associations between %AA (20:4 n6) and Relative Amounts of Other Body Fatty Acids. JNFP 3(2) (2); DOI:10.31579/2637-8914/024

In chicken muscle, we identified *two groups* of fatty acids: Group 1) with relative amounts correlating *negatively* with %AA, and Group 2) with relative amounts correlating *positively* with %AA (Table 1). With the positive correlations, but not with the negative ones, we obtained qualitatively similar scatterplots using true and random numbers, if generated with the true ranges. We suggested that the apparent discrepancy possibly could relate to differences in *skewness* of the concentration distribution (Table 1). Interestingly, most of Group 2 fatty acids were eicosanoid or docosanoid precursors. The particular ranges of the fatty acid seemed in general to explain the correlation outcome, in support of the concept of *Distribution Dependent Correlations.*

Absolute amounts (g/kg)	Min	Max	Mean	SD	Skewness	CV (%)
Total SUM	3.04	14.69	5.21	1.67	2.29	32.1
Group 1 18:3 n3	0.12	2.40	0.53	0.32	2.50	60.4
18:1 c9	1.04	8.56	2.44	1.07	2.29	43.9
16:1 c9	0.03	0.78	0.18	0.11	2.27	61.1
18:3 n6	0.00	0.02	0.01	0.00	1.70	44.8
Sum of Group 1 fatty acids	1.20	11.77	3.16	1.50	2.33	47.5
Group 2 20:3 n6	0.06	0.11	0.08	0.01	0.70	11.0
20:3 n3	0.04	0.09	0.05	0.01	1.06	12.2
20:2 n6	0.04	0.06	0.05	0.01	0.46	13.7
18:0	0.65	1.70	0.89	0.16	1.69	18.7
22:5 n3	0.21	0.43	0.31	0.04	0.52	13.2
20:4 n6	0.25	0.42	0.31	0.03	0.66	9.4
22:6 n3	0.11	0.32	0.19	0.04	0.76	21.2
20:5 n3	0.13	0.24	0.18	0.02	-0.16	11.7
Sum of Group 2 fatty acids	1.64	2.93	2.05	0.22	0.90	10.7
Relative amounts (%)						
Group 1 18:3 n3	3.91	16.34	9.44	2.40	0.17	25.4
18:1 c9	34.06	58.28	45.58	5.00	0.02	11.0
16:1 c9	1.12	6.82	3.23	0.95	0.53	29.5
18:3 n6	0.09	0.22	0.15	0.03	0.19	17.6
Group 2 20:3 n6	0.62	2.45	1.55	0.36	0.02	23.3
20:3 n3	0.40	1.86	1.00	0.26	0.42	25.7
20:2 n6	0.40	1.53	0.92	0.21	0.26	23.1
18:0	11.60	24.68	17.73	2.71	0.02	15.2
22:5 n3	2.07	11.00	6.36	1.64	0.26	25.9
20:4 n6	2.25	11.02	6.35	1.53	0.05	24.1
22:6 n3	1.24	7.70	3.95	1.25	0.41	31.7
20:5 n3	1.32	5.70	3.73	0.91	-0.00	24.5

Table 1. Absolute (g/kg) and relative (%) amounts of fatty acids in chicken breast muscle lipids (n = 163); min (max) values, means, SD, skewness, and coefficient of variation (CV), categorized into Group 1 and Group 2, see text. Note that some values appear as zero due to the number of decimals. Standard error of Skewness: 0.19. From Høstmark AT, Haug A, JNFP, 3(2), 2020.

 # 10: Høstmark AT, Haug A (2020) Distribution Dependent and Cluster Regulation of Associations between Body Fatty Acid Percentages, as observed in Chicken. JNFP 3(2); DOI:10.31579/2637-8914/025

We here tried to explain further our previous finding in chicken muscle, i.e. the apparent existence of two groups of fatty acids: Group 1) with relative amounts correlating *negatively* with %AA (20:4 n6), and Group 2) with relative amounts correlating *positively* with %AA. Within each of the two groups, we found positive correlations between the fatty acid percentages. It follows that, percentages of Group 1 fatty acids correlated negatively with percentages of Group 2 fatty acids. Furthermore, with random numbers in lieu of the true values of Group 2 fatty acids, however using the true ranges, we were able to reproduce the positive correlations found with true values. In contrast, with random numbers we did not succeed in reproducing *all* of the negative correlations between Group 1 and Group 2 fatty acid percentages. We then observed that the *absolute* amounts (g/kg) of fatty acids in Group 1 correlated positively and strongly (r > 0.9), suggesting a coordinated regulation of these fatty acids (Table 2). Thus, Group 1 fatty acids seemed to be a *cluster* of fatty acids. "Random number cluster percentage" showed nice inverse associations with random number Group 2 fatty acid percentages, like the outcome observed with the true values, suggesting that associations between fatty acid percentages are not only caused by their concentration distributions, but also by cluster regulation.

Pearson's r - values				
Fatty acid	18:3 n3	18:1 c9	16:1 c9	18:3 n6
18:3 n3	1			
18:1 c9	0.988	1		
16:1 c9	0.932	0.953	1	
18:3 n6	0.943	0.949	0.913	1
Spearman's rho - values				
Fatty acid	18:3 n3	18:1 c9	16:1 c9	18:3 n6
18:3 n3	1			
18:1 c9	0.984	1		
16:1 c9	0.904	0.926	1	
18:3 n6	0.914	0.923	0.865	1

Table 2. *Correlations between absolute amounts (g/kg) of Group 1 fatty acids. All correlation coefficients are with $p < 0.01$, $n = 163$.*

	Group 1	Group 2
Place on the scale	Generally HIGH NUMBERS	LOW NUMBERS
Variability	HIGH	LOW
Skewness, *absolute* values	HIGH positive	LOW
Skewness, *relative* values	LOW	LOW
Correlations, within - group percentages	POSITIVE	POSITIVE
Correlations, between – group percentages	NEGATIVE	NEGATIVE
Cluster regulation	YES	NO
Eicosanoid/docosanoid precursors	NO	YES

Table 3. *Characteristics of Group 1 and Group 2 fatty acids in chicken breast muscle*

 # 11: Høstmark AT (2020) Association between Relative Amounts of White Blood Cell Counts: a Case of Distribution Dependent Correlations JNFP 3 (2); DOI:10.31579/2637-8914/028

Since counts of segmented neutrophil leukocytes (N) and lymphocytes (L) are normally much higher than sum of the remaining (R) white blood cells (WBC), we suggested that %N might possibly relate to %L. Thus, with random numbers, sampled in lieu of reported WBC subgroup values, however using the true (reported) mean ± SD values, we were able to show that %N' and %L correlated negatively, in both sexes, i.e. Spearman's rho = -0.9, $p < 0.01$, $n = 200$. We found qualitatively similar results using *within*-person data, and *between*-person data. However, altering *distributions (ranges)* of WBC subgroups changed the correlation outcome, as evaluated by scatterplots and correlation coefficients. Decreasing (increasing) values of %R improved (made poorer) the negative association between %N and %L. Accordingly, the observed negative association between %N and %L seems to be a case of *Distribution Dependent Correlations.* Thus, directing the counts of WBC subgroup to particular places on the scale could have major effects to change associations between their relative amounts.

 # 12: Høstmark at (2020) Distribution dependent correlations: a mathematical principle utilized in physiology, or correlation bias? Int J Res - GRANTHAALAYAH, 8(11), 63-75. https://doi.org/10.29121/granthaa layah.v8.i11.2020.1470

The aim of this work was to elucidate further, how correlations between percentages of the same sum arise. We extended and systematized our previous theoretical considerations, and carried out new sets of computer experiments to test the hypotheses. The results were in support of the idea that true, within-person distributions of the variables are crucial for obtaining positive or negative correlations between their relative amounts.

 # 13: Høstmark A T, Haug A (2021) Studies to Explain Associations between Relative Amounts of Body Fatty Acids JNFP 4(1); DOI:10.31579/2637-8914/039

The purpose of this work was to explore in more detail the concept of Distribution Dependent Correlations (DDC). We did many computer experiments to establish that, with three positive scale variables, two of which (A, B) having very low variability relative to a third one (R), we should expect a strong positive association between percent A and percent B, and that ***the B/A ratio*** could estimate ***slope*** of the regression line. Furthermore, we should expect a negative relationship between %R and %A (%B). This correlation outcome did occur wherever we placed A and B on the positive scale. On the other hand, if A and B have high numbers and broad ranges relative to R, then %A should relate inversely to %B. We additionally showed computer experiments (and explained) how differences in ranges might give skewness of the relative amounts. Thus, ranges of A, B, and R seem to govern associations between their relative amounts, and alterations in the ranges have appreciable effects to change the associations, suggesting that evolution might utilize DDC to regulate metabolism.

 # 14: Høstmark AT. Haug A (2021) Studies to Explain why Percentages of Eicosanoid Precursor Fatty Acids Are Positively Associated in Chicken Lipids. J Food Sci & Nutri: JFSN-112.DOI: 10.46715/jfsn2021.01.1000112

Previously, we reported that relative amounts of eicosanoid precursor fatty acids were positively associated in breast muscle lipids of chickens. The concentration ranges seemed to largely explain e.g. the positive association between % AA and % EPA. Here we tried to explain such correlations in more detail. We focused upon EPA, AA, DGLA, and DHA. Thus, EPA + AA + DGLA + DHA + R = S, where R is sum of the remaining fatty acids, and S is sum of all. Since amount of these four fatty acids is small as compared with S, we anticipated that their percentages related negatively to S, as corroborated by Spearman's rho values for S vs. %EPA, %AA, %DGLA, and %DHA, i.e. -0.867, -0.913, - 0.886, and -0.730, respectively. However, also RANDOM numbers, generated with the true ranges of the fatty acids, gave similar correlations, i.e.: - 0.919, - 0.937, -0.910, and - 0.838. Accordingly, all precursor percentages correlated positively. A narrowing (broadening) of concentration ranges strongly improved (made poorer) correlations between percentages. Thus, eicosanoid (docosanoid) fatty acid precursor percentages must be positively associated as a mathematical consequence of their particular concentration distributions. Again, the results are in favor of the concept of Distribution Dependent Correlations, possibly representing an evolutionary regulatory mechanism, where *variation in ranges* could improve the balance between eicosanoids/docosanoids, mathematically.

 # 15: Høstmark AT (2021) Intended Ranges and Correlations between Percentages of Variables Like Oleic Acid, Eicosapentaenoic Acid, and Arachidonic Acid. Foods 2021, 10, 1012. https://doi.org/10.3390/foods10051012

In chicken muscle, we previously showed that *ranges* of OA, AA, and EPA might explain why %OA was inversely related to %AA, and that %EPA correlated positively with %AA. We here tried to clarify further, how *ranges* of the fatty acids could make *strong associations* between their relative amounts, utilizing published data from chicken muscle and human sera. We generated random number variables (OA', AA', EPA') in lieu of the true variables, and studied effects of altering their ranges upon scatterplots of %OA' vs. %AA' (%EPA'), and %AA' vs. %EPA'. To explain the results, we first applied the equation OA' + AA' + EPA' = S, i.e. %OA' + %AA' + %EPA' =100. Next, we considered how the OA' (AA', EPA') *fractions* of S related to S. Increasing the OA' range towards higher values improved the positive association between %AA' and %EPA'. Thus, broadening the OA-range, presumably obtained with increased intake of OA, could improve the positive associations between relative amounts of eicosanoid precursors, raising the question of whether "*intended ranges*" of some fatty acids could represent a case of evolutionary selection, to achieve balance between eicosanoids. In this context, we also refer to the supplementary and more detailed comments, related to Paper *# 4* and *Paper # 5,* above.

Conclusions

These studies seem to suggest that the *particular distributions* (range/place on the scale/spread) of variables such as fatty acids will determine whether their relative amounts are positively or negatively associated, or not correlated at all.

Suggested rules pertaining to DDC:

1. *With 3 positive scale variables (A, B, C) , two of which (A, B) having narrow ranges as compared with the third variable (C), we might expect that*
   - there will be positive association between %A and %B, and negative association between %C and %A (%B).
   - we can use the ***B/A ratio*** *to estimate slope of the %A vs. %B regression line.*
   - a decrease (increase) in the variability of either one or both of the two narrow - range variables will improve (make poorer) the association between their relative amounts.
   - narrowing (broadening) the range of the broad - range variable will make poorer (improve) the association between percentages of the narrow-range variables.
2. *With 3 positive scale variables, two of which (A, B) having high numbers and broad ranges relative to the third variable (C), we should expect a negative association between %A and %B, and probably a poor association between %C and % A (%B).*
3. *Ranges of 3 positive scale variables can be manipulated so as to obtain a* ***Turning Point****, i.e. a situation where a positive (negative) association between percentages of two of them turns to become negative (positive).*
4. *With more than three variables, we may calculate the sum of all, minus the two under investigation; the preceding rules should in general apply to this "3-variable-modification".*

The observed positive associations between percentages of fatty acids, that are precursors of eicosanoids (docosanoids), could possibly serve to improve the balance between molecules having opposing actions. We suggest that **Intended Ranges** could be a type of evolutionary selection, to obtain ***Distribution Dependent Correlations (DDC)***, mathematically.

Conflicts of Interest: None

Acknowledgment:

I want to thank Professor emerita Anna Haug very much for allowing me to use data from her excellent diet trial in chickens.

References

1. Høstmark AT, Haug A (2014). The inverse association between relative abundances of oleic acid and arachidonic acid is related to alpha -linolenic acid. Lipids in Health and Disease 13(1):76.
2. Mayes PA (2000). Metabolism of unsaturated fatty acids and eicosanoids. In: Murray RK, Granner DK, Mayes PA, Rodwell VW, ed. Harper's Biochemistry, New York: McGraw-Hill, p 250-258.
3. Baker RR (1990). The eicosanoids: a historical overview. Clin Biochem 23:455-458.
4. Gogus U, Smith C (2010). n-3 Omega fatty acids: a review of current knowledge. Int J Food Sci Tech 45:417–436.
5. Christie W. W. The Lipid Web (2019). Eicosanoids and Related Oxylipins, an Introduction.
6. Kremer JM, Bigauoette J, Michalek AV (1985). Effects of manipulation of dietary fatty acids on manifestations of rheumatoid arthritis. Lancet: 184–187.
7. Lorenz R, Weber PC, Szimnau P (1989). Supplementation with n-3 fatty acids from fish oil in chronic inflammatory bowel disease – a randomized, placebo-controlled, doubleblind cross-over trial. J Intern Med Suppl 731: 225–732.
8. Kromhout D, (2012). Omega-3 fatty acids and coronary heart disease. The final verdict? Curr Opin Lipidol 23:554-559.
9. Watanabe T, Ando K, Daidoji H, Otaki Y, Sugawara S, et al. (2017). A randomized controlled trial of eicosapentaenoic acid in patients with coronary heart disease on statins: The CHERRY study. J Cardiol 70:537-544.
10. Nagata M, Hata J, Hirakawa Y, Mukai N, Yoshida D, et al.(2017). The ratio of serum eicosapentaenoic acid to arachidonic acid and risk of cancer death in a Japanese community: The Hisayama Study. J Epidemiol 27: 578–583.
11. Abdelhamid AS, Brown TJ, Brainard JS, Biswas P, Thorpe GC, et al. (2018). Omega-3 fatty acids for the primary and secondary prevention of cardiovascular disease. Cochrane Database of Systematic Reviews.
12. Alvheim AR, Malde MK, Osei-Hyiaman D, Lin YH, Pawlosky RJ, et al. (2012). Dietary linoleic acid elevates endogenous 2-AG and anandamide and induces obesity. Obesity 20:1984-1994.
13. Haug A, Olesen I, Christophersen OA (2010). Individual variation and intraclass correlation in arachidonic acid and eicosapentaenoic acid in chicken muscle. Lipids Health Dis 15:37.
14. Høstmark AT, Haug A (2018). The Fatty Acid Distribution per se Explains Why Percentages of Eicosapentaenoic Acid (20:5 n3) and Arachidonic Acid (20:4 n6) are Positively Associated; a Novel Regulatory Mechanism? J Nutr Diet Suppl 2(1): 103.
15. Høstmark AT, Haug A (2019). Associations Between %AA (20:4 n6) and Percentages of EPA (20:5 n3), DPA (22:5 n3), and DHA (22:6 n3) Are Distribution Dependent in Breast Muscle Lipids of Chickens. J Nutr Diet Suppl 3(1): 103.

16. Høstmark AT (2019). Association between Percentages of Scale Variables, as Related to Distributions. J Nutr Diet Suppl 3(1): 104.
17. Høstmark AT, Haug A (2019). High Variability of Oleic Acid (OA, 18:1 c9) improves the Positive Association between %EPA (20:5 n3) and %AA (20:4 n6). J Nutr Diet Suppl 3(1): 106.
18. Høstmark AT (2019). Body Fatty Acids, Nutrition, and Health: Is Skewness of Distributions a Mediator of Correlations? J Nutrition and Food Processing, 2(1).
19. Høstmark AT, Haug A (2019). The inverse association between relative abundances of oleic acid and arachidonic acid: a case of distribution dependent regulation? Lipids in Health and Disease, 18:123.
20. Høstmark AT, Haug A (2020). Relative amounts of eicosanoid and docosanoid precursor fatty acids are positively associated: a distribution dependent regulation. Journal of Nutrition and Food Processing.
21. Høstmark AT, Haug A (2020). Associations between %AA (20:4 n6) and relative amounts of other body fatty acids. Journal of Nutrition and Food Processing, 3(2).
22. Høstmark AT, Haug A. (2020). Distribution dependent and cluster regulation of associations between body fatty acid percentages, as observed in chickens. Journal of Nutrition and Food Processing, 3(2).
23. Høstmark AT (2019). Body fatty acids, nutrition, and health: Is skewness of distributions a mediator of correlations? Journal of Nutrition and Food Processing, 2(1).
24. Høstmark AT, Haug A (2019). Alpha Linolenic Acid Variability Influences the Positive Association between %Eicosapentaenoic Acid and %Arachidonic Acid in Chicken Lipids. J Nutrition and Food Processing, 2(2).
25. Høstmark AT (2020). Distribution Dependent Correlations: A Mathematical Principle utilized in Physiology, or Correlation Bias? Int J Res - Granthaalayah, 8(11), 63-75.
26. Høstmark AT, Haug A (2021). Studies to Explain why Percentages of Eicosanoid Precursor Fatty Acids Are Positively Associated in Chicken Lipids. Jou Food Sci&Nutri: JFSN-112.
27. Høstmark AT (2020). Association between relative amounts of white blood cell counts: a case of Distribution Dependent Correlations. Journal of Nutrition and Food Processing, 3(2).
28. Høstmark AT (2021). Intended Ranges and Correlations between Percentages of Variables like Oleic Acid, Eicosapentaenoic Acid, and Arachidonic Acid. Foods, 10, 1012.
29. Meng L-B, Yu Z-M, Guo P, Wang Q-Q, Qi R-M, Shan M-J, Lv J, Gong T (2018). Neutrophils and neutrophil-lymphocyte ratio: Inflammatory markers associated with intimal-media thickness of atherosclerosis. Thrombosis Res 170: 45-52.
30. Liu Y, Du X, Chen J, Luo M, Chen L, Zhao Y (2020). Neutrophil-to-lymphocyte ratio as an independent risk factor for mortality in hospitalized patients with COVID-19. J Infect 81: e6-e12.
31. Swamidass, P.M. (Ed.) (2000). Assignable Causes of Variations. In Encyclopedia of Production and Manufacturing Management; Springer: Boston, MA, USA, 2000.

DOI: 10.31579/2637-8914/077