Journal of Biometrics & Biostatistics
ISSN: 2155-6180

Larsen et al., J Biom Biostat 2015, 6:4
DOI: 10.4172/2155-6180.1000255

Research Article Open Access

The Concepts of Fractal Ellipses and ISO Likelihood Ratio Curves in Two-dimensional Screening Procedures with Applications in Screening for Down Syndrome

Severin Olesen Larsen[1], Paula L Hedley[1,2] and Michael Christiansen[1]*

[1]*Department of Clinical Biochemistry, Immunology and Genetics, Statens Serum Institut, Copenhagen, Denmark*
[2]*Department of Biomedical Sciences, University of Stellenbosch, Cape Town, South Africa*

Abstract

Background: Prenatal screening combines biochemical and biometric markers into a risk estimate for a particular adverse outcome, e.g. the birth of a child with Down syndrome. The statistical calculations are complicated. We describe a simple graphical method to perform risk estimation in the case of two biochemical markers, to assess the consequences of changes in gestational dating of the pregnancy and to perform quality control.

Materials and methods: We used the formulae for the Normal distribution to establish the expression for fractal ellipses, i.e. contour ellipses describing a certain fractile of the total distribution. This expression was used to establish mathematical expressions for curves describing two-dimensional pairs of analytes giving the same likelihood-ratio, i.e. iso likelihood-ratio curves.

Results: The fractal ellipses provide an overview of marker distributions that allow for an easy control of empirical marker distributions. The iso likelihood-ratio curves provide a relation between likelihood-ratio and marker values. They can be used for assessment of the consequences of changes in gestational age and introduction of truncation limits on markers.

Conclusions: Fractal ellipses and iso likelihood-ratio curves can be used to make software-independent calculations and modifications of risk in prenatal screening and quality control of an ongoing screening program.

Keywords: Prenatal screening; Normal distribution; Quality control; Diagnostic biochemistry

Abbreviations: βhCG: Free β-Form of Human Chorionic Gonadotrophin; LR: Likelihood Ratio; MoM: Multiples of the Median; PAPP-A: Pregnancy-associated Plasma Protein A; SD: Standard Deviation

Introduction

Screening is an activity where unaffected persons are assessed with respect to the risk of a particular condition with the purpose of evaluating whether diagnostic procedures for this condition are warranted [1]. A typical example is first trimester screening for Down syndrome involving the use of the biochemical parameters PAPP-A and free hCGβ in combination with the thickness of the nuchal translucency [2]. The risk calculation is based on the association of the biochemical and biometric markers to a likelihood of the fetus being affected with Down syndrome. Thus, the distribution of parameter values, normalised for gestational age and maternal weight, in unaffected and affected pregnancies are determined and the distribution of likelihood values as a function of parameter values is established [1]. The likelihood ratio calculation is based on the assumption that the distributions of the parameter in question in affected and unaffected cases are compatible with a Gaussian distribution [3].

The statistical calculations in the two- or three-dimensional Gaussian distribution are very complicated and involve the use of statistical software. This is the case both for actual individual risk calculations and quality control of screening programmes. This makes the performance of a screening program entirely dependent upon the correct performance of statistical software and thus very sensitive towards bugs in such programmes. There is, in conclusion, a need for a simple way to monitor the correspondence between marker distributions and likelihood ratio (or risk) distributions and a way to analyse the consequences of manipulation with risk calculation parameters, e.g. truncation limits, or gestational age in individual cases.

Here the concepts of fractal ellipses and of iso likelihood-ratio curves are presented. Fractal ellipses are shown to be useful in defining outliers and thus the empirical marker distributions of a screening program. Iso likelihood-ratio curves are shown to be useful when understanding the effects of introducing truncation limits and changing gestational age calculations.

Material and Methods

Statistical models of two-dimensional Gaussian screening – Assumptions and notation

We consider two markers x and y being measured so that the mean values for the normal population are=0. This is the case if the values are log MoM values, i.e. the log of the ratio between observed and expected values as in Down's syndrome screening, but the measurements can always be expressed so that this is true. Furthermore we consider a

***Corresponding author:** Michael Christiansen, Research Director, Chief physician, Department of Clinical Biochemistry. Immunology and Genetics Statens Serum Institute, 5 Artillerivej DK 2300S, Denamark, Tel: +4532683657; Fax: +4532683878; E-mail: mic@ssi.dk

Received October 09, 2015; **Accepted** October 17, 2015; **Published** October 26, 2015

Citation: Larsen SO, Hedley PL, Christiansen M (2015) The Concepts of Fractal Ellipses and ISO Likelihood Ratio Curves in Two-dimensional Screening Procedures with Applications in Screening for Down Syndrome. J Biom Biostat 6: 255. doi:10.4172/2155-6180.1000255

background population with normal values and a population affected by some disease or syndrome e.g. pregnant women with a foetus with Down's syndrome. For both populations we assume that the (x,y) values follow a two-dimensional normal (Gaussian) distribution. In such screening the likelihood ratio is used in combination with an a priori risk value to calculate an improved estimate of the risk of being affected. The likelihood ratio is the probability of (x,y) in the distribution for affected cases divided by the probability for the normal cases. The final risk estimate is the a priori risk value multiplied by the likelihood ratio.

For the normal population the mean values will be=0, whereas the mean values for the affected population are M{x}=μ and M{y}=λ. For the normal population the standard deviations are SD{x}=σ and SD{y}=ω, whereas for the affected population these values are SD{x}=γ and SD{y}=δ. There may be a correlation between x and y and the correlation may differ between normal and affected pregnancies. For the normal population the correlation coefficient is ρ and for the affected population it is φ.

Calculations

All calculations were performed using algorithms programmed in S-PLUS Vers.6. (Insigthful, Seattle,USA). The scripts are available upon request.

Results

Fractal ellipses

Fractal ellipses are contour ellipses that comprise a certain fractile of a distribution of parameters. In the two-dimensional Gaussian (normal) distribution, fractal ellipses are defined by expression (i), where C is a constant, x and y are parameters in the distribution, ρ is the correlation between the parameters, and σ and ω are the standard deviations of x and y, respectively.

(i) $C=((x/\sigma)^2-2\rho(x/\sigma)(y/\omega)+(y/\omega)^2)/(1-\rho^2)$

C follows a χ^2 distribution with two degrees of freedom [4]. For a constant value of C, corresponding to a particular P-fractile in the distribution, the expression defines a contour ellipse inside which P % of the x, y parameter pairs should be located. C values corresponding to particular P %-fractiles are given in Table 1.

As seen in Table 1, only 1 % of the observations will lie outside the fractellipse defined by C=9.210. For a certain C value the parameters x and y are located within the intervals $-\sigma\sqrt{C}$ to $\sigma\sqrt{C}$ and $-\omega\sqrt{C}$ to $\omega\sqrt{C}$, respectively. Corresponding x and y pairs located on a certain fractellipse, with a constant C, can be obtained by expressing (i) as a second-order relation between y and x (ii).

(ii) $y=\acute{\omega}(\rho x/\sigma \pm \sqrt{((1-\rho^2)(C-x^2/\sigma^2))})$

In a situation where two populations, e.g. a healthy normal population and an affected population, are described we need two groups of fractellipses; one defining the normal population (i) and a similar defining the fractellipses in the affected population (iii).

Fractiles in the χ^2 distribution (f=2)	
P	**C**
95%	5.991
99%	9.21
99.50%	10.597
99.90%	13.816

Table 1: Fractiles in the χ^2 distribution (f=2).

	Normals	Affected
(x)	Mean=0 σ=0.30	μ=0.4 γ=0.4
(y)	Mean=0 ω=0.25	λ=-0.35 δ=0.35
Correlation (A) Correlation (B)	ρ=0.15 ρ= 0.70	φ=0.15 φ=0.70

Table 2: Parameters used in constructed example.

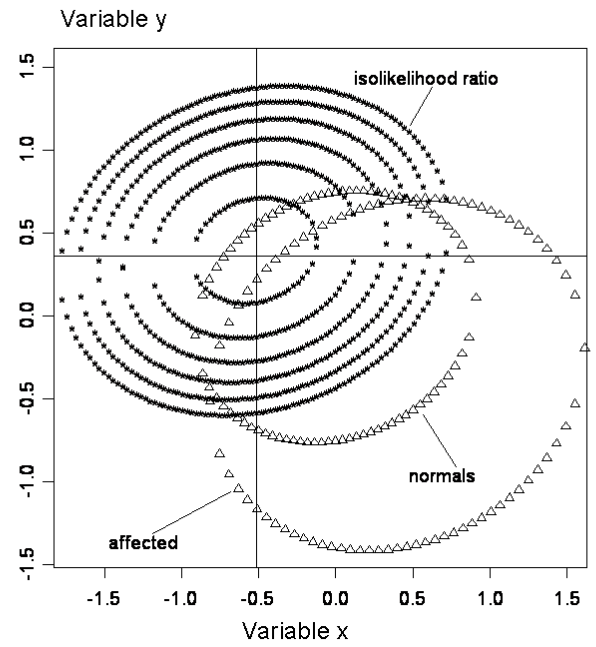

Figure 1: Case A in Table 2 (low correlation). 99% fractalellipses (ΔΔΔ) for normal and for affected persons. Iso likelihood ratio curves (...), likelihood ratio (affected (normals) from inside out=0.0625, 0.125, 0.25, 0.5, 1, and 2. Straight lines=one-dimensional truncation limits to point of risk reversal (see text).

(iii) $D=(((x-\mu)/\gamma)^2-2\varphi(x-\mu)(y-\lambda)/\gamma\delta+((y-\lambda)/\delta)^2)/(1-\varphi^2)$

Here, μ and λ are the mean values for the distribution of x and y, respectively. The correlation between x and y in the affected population is denoted φ. The standard deviations in the affected populations of x and y are γ and δ, respectively. Based on parameters from a constructed example, Table 2, Figures 1 (case A) and 2 (case B) show fractal ellipses from an affected and unaffected population. Figure 1 shows a case with a low correlation between x and y, whereas Figure 2 shows a case with the same mean values and standard deviations but now with a high correlation between x and y.

Fractal ellipses can be used to monitor the performance of a screening programme, simply by plotting x and y values in a fractal ellipse diagram. If the distribution of empirical x-y pairs differ from the expected – defined by fractellipses – a revision of performance is advisable. The shape of the empirical distribution of x-y pairs may also suggest the underlying deviation from normal performance. Furthermore, in the single case, the fractellipses may aid in giving an impression of the deviation from the normal. This may be of importance when advising the individual woman.

Iso likelihood-ratio curves

Using (i) and (iii) and the general formula for the two-dimensional

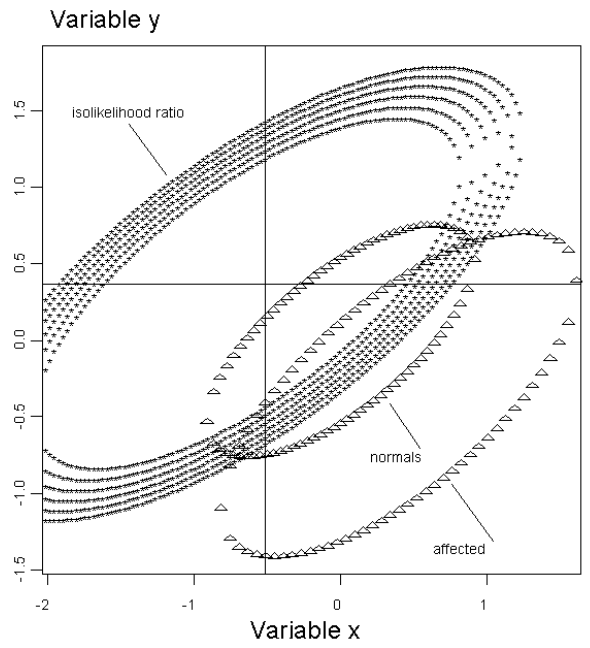

Figure 2: Case B in Table 2 (high correlation). 99 % fractal ellipses (ΔΔΔ) for normal and for affected persons. Iso likelihood ratio curves (…), likelihood ratio (affected/normals) from inside out=0.0625, 0.125, 0.25, 0.5, 1, and 2. Straight lines=one-dimensional truncation limits to point of risk reversal (see text).

Gaussian distribution, it is possible to define the likelihood ratio of being affected as compared to normal, L, as:

(iv) $L=(\sigma\omega\sqrt{(1-\rho^2)}/(\gamma\delta\sqrt{(1-\varphi^2)}) \times \exp((C-D)/2)$

where exp denotes the exponential distribution. For a constant L, expression (iv) defines an iso likelihood-ratio curve when plotted in an x,y diagram. This relation is a second order equation in x and y, see (i) and (iii), but the number of parameters, nine including the constant L, precludes an explicit solution, like that given for the fractal ellipses (iii), with respect to y as a function of x. However, for a predefined set of parameters a simple second order equation in x and y can be obtained. The S-Plus scripts necessary to define pairs of x and y can be obtained from the authors.

The iso likelihood-ratio curve is an ellipse if the standard deviation differences σ-γ and ω-δ have the same sign. If this is not the case the iso likelihood-ratio curve will be a hyperbola, except when both the standard deviations and correlations are identical in the two populations. In the latter case, the iso likelihood-ratio curve will be a straight line.

It should be noted that not all x and y pairs occur with equal frequency. The values of x and y close to the mean are much more frequent than the values farther apart. So, for practical purposes only a relatively small part of the iso likelihood-ratio curve is of interest.

Use of iso likelihood-ratio curves in the analysis of effect of truncation

Truncation of extreme marker values in biochemical screening is often performed by adjusting a value to a value ("The truncation limit") closer to the mean value in the normal population. Truncation is carried out primarily for two reasons: 1) If the if it is obvious that some values do not fit the assumed Gaussian distribution but represent a "tail" of outlying observations and 2) If the standard deviations for affected cases are considerably larger than for normal cases [5-7]. In the latter case observations far from the expected but to the normal side may have a larger likelihood for being affected than for being normal. This is because the distribution of affected cases with its larger standard deviation can easier "reach out" to such values. This can be easily illustrated by graphs (being ellipses) of a series of iso likelihood ratio curves in cases with considerably larger standard deviations for the affected cases. The center for these ellipses will be the point with the lowest two-dimensional likelihood ratio (affected/normal). Based on one-dimensional thinking, treating each marker separately, the markers should be truncated separately to the points of risk reversal i.e. for the observed x values to the value $-\sigma^2\mu/(\gamma^2-\sigma^2)$ and for y values to the value $-\omega^2\lambda/(\delta^2-\omega^2)$. At these points the one-dimensional likelihood ratios are at their minimum value [5-7].

Using the constructed data in Table 2 a number of iso likelihood ratio curves based on the equation L=K are plotted in Figures 1 (A) and 2 (B) for various values of K. In both cases the mean values and the standard deviations are the same, but in case B there is a much higher correlation between the x and y values. The fixed likelihood ratio values used are from outside K=2, 1, 0.5, 0.25, 0.125 and 0.0626. The 99% fractal ellipses for the normal and for the affected population are also shown. It is seen that because of the higher correlation case B gives a much better discrimination than case A with almost no correlation. Combinations of x and y lying far from the mean values have of course a very low probability. Using one-dimensional truncations to the points (in one dimension) of risk reversal would in both cases change all high values of y down to 0.365 and all low values of x up to a value of -0.514. These truncation limits are indicated by a vertical and a horizontal line. It is seen that truncation to these lines would perform rather well if there is only a very weak correlation between the markers (Figure 1, case A). A similar good performance of combining two one-dimensional truncations is obviously not possible if there is a strong correlation between the two markers (Figure 2 case B). In fact many observations inside the 99% contour ellipse for the normal population will in this case be truncated into considerably higher likelihood ratio values (affected/normal) if the two one-dimensional truncations are used.

Use of iso likelihood-ratio curves in screening for Down syndrome

A very useful application of iso likelihood-ratio curves can be illustrated by the case of biochemical screening for Down's syndrome using the markers PAPP-A (=y) and free beta hCG (=x) in the first trimester (the double test). The curves are based on published distribution parameters (Table 3) [8]. When inserting these parameters in the natural logarithm of L the following expression is obtained:

(v) $\ln(L)=0.09847\ x^2+0.02474\ xy+1.4469\ y^2+4.1256\ x-4.1468\ y-1.5617$

ln denotes the natural logarithm. The equation can be used to

	Normals	Down's
Free hCGβ (x)	σ=0.287	μ=0.296
		γ=0.290
PAPP-A (y)	ω=0.285	λ=-0.39
		δ=0.326
Correlation	ρ=0.111	φ=0.130

Table 3: Parameters used in the double test for Down syndrome.

find x, y values corresponding to a fixed value of L (the iso likelihood-ratio values), and to find values of L for a fixed value of y and varying values of x. A step from one curve in the figure to the next represents a doubling of the likelihood ratio. The curve corresponding to L=1 is specially marked (***). For a constant value of y (log MoM PAPP-A) there is, in this case, an almost perfect linear relation between the log likelihood ratio and x (log MoM beta hCG). This is because in this case σ is almost equal to γ and ρ and φ are both small and almost equal. It is seen that the graph offers good opportunities for a graphical determination of the likelihood ratio. In the figure an arrow has been inserted indicating the change in log MoM values that would result from a change in the estimated gestation age. This is possible because for the double test markers there is a linear relation between the log values and the gestation age in days. The following relations were used for the construction of the arrow:

(vi) log PAPP-A~0.6799+0.0321 × gestation age (days) and

(vii) log betahCG~2.5595-0.0108 × gestation age (days)

For an assessment of the change of the likelihood ratio as a function of a change in the gestational age (a situation that may occur following the revision of a gestational dating based on LMP by later ultrasound determination) the starting point of the arrow is placed in the point (x, y-point) representing the log MoM values for PAPP-A and hCGβ corresponding to the first gestational age. The end point of the arrow will then be placed on the new iso likelihood-ratio curve. The length of the arrow is directly proportional to the change in gestational age. The arrow in Figure 3 has a length corresponding to a change of 10 days in gestational age. The blue arrow in Figure 1 represents an example where the change in gestational age is 10 days and the resulting change in LR is a factor 8. The red arrow is half the length of the blue arrow and it represents a change of 5 days in gestational age. The resulting change in LR is a factor 4. It is important to note that the length as well as the direction of the arrow, both defined by (vi) and (vii), are necessary for its use.

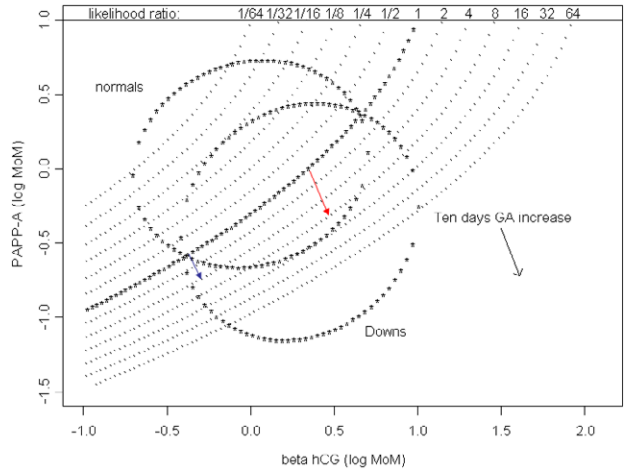

Figure 3: Isolikelihood ratio curves for Down's syndrome screening using the markers beta hCG and PAPP-A. The likelihood ratio (Downs/normals) indicated as 1:64, 1:32 etc. 95% fractal ellipses for normals and for Downs. The arrow indicates the change in log MoM values that would be the result of a ten-day increase in the estimated gestation age.The red arrow shows the change in likelihood ratio for a woman where the gestational age was changed by five days and the blue for a gestational age change of ten days.

Discussion

We have shown that fractal ellipses allow for a simple way to control empirical marker distributions in prenatal screening. Without the use of statistical software it is possible to define outliers in the two-dimensional distribution. Defining a normal range for the (x,y) values based on fractal ellipses is important in screening for Down's syndrome for two reasons. Firstly, an increased number of outliers, and their distribution with respect to the theoretical marker distribution may call attention to an inappropriate risk algorithm and also point towards the irregularity causing the problem. A χ^2-test of the difference between the expected and observed distribution around a fractal ellipse is easily performed and more relevant than monitoring individual marker distributions. Secondly, it is well documented that outliers comprise a variety of pregnancy complications, e.g. pre-eclampsia, some of which will benefit from either prophylactic measures or clinical follow-up. The fractal ellipses may aid in quickly identifying such individual outliers.

Another aspect of the use of fractal ellipses is their value in communicating risk results. To see a graphical depiction of the result may help the pregnant woman to understand where her risk result is compared to that of other women. If it is reported and even illustrated that the screening result is well within the normal range but lying where there are relatively many cases with Down's syndrome she may more easily assured, if Down's syndrome is ruled out, that there is no reason for anxiety. We suggest that fractal ellipses are introduced as means of communicating risk.

The concept of iso likelihood ratio curves is useful in studying what is going on in the screening including possible truncations of the markers used. It is evident from the case shown in Figure 2 (high correlation between the markers) that a very careful study of the iso likelihood ratio curves should be recommended before truncation limits are decided upon. It seems that in some cases each marker should not be truncated separately but be combined in an adjustment to values nearer the expected value.

In practice the possibility of a graphical determination of the likelihood ratio may in many case not seem to be very useful since the likelihood ratio will often be calculated using computer software and reported together with the marker values. However, in screening for Down's syndrome the iso likelihood-ratio curves could be combined with the age dependent a priori risk values allowing construction of a nomogram for graphical determination of the final risk. This can be done simply by linking the iso likelihood ratio curves to a graph of isorisk curves (not shown).

Conclusions

The use of fractal ellipses and iso likelihood-ratio curves can thus aid in 1) identifying screening programs with unsatisfactory performance; 2) calculation of risk and risk modifications and 3) the information of screened women. It should also be noted that the application of fractal ellipses and iso likelihood-ratio curves is described here in relation to prenatal biochemical screening, but the principle is applicable in all cases where several markers, complying with the Gaussian distribution, are combined into a diagnostic or prognostic index.

References

1. Wald NJ (1984) Antenatal and neonatal screening (1st.ed.) Oxford University Press, Oxford and New York.

2. Avigdou K, Papageorghiou A, Bindra R, Spencer K, Nicolaides KH (2005) Prospective first-trimester screening for trisomy 21 in 30,564 pregnancies. Am J Obstet Gynecol 192: 1761-1767.

3. Larsen SO, Christiansen M, Norgaard-Pedersen B (1998) Calculation of ROC curves in multidimensional likelihood ratio based screening with Down's syndrome as a special case. J Med Screen 5: 57-62.

4. Hald A (1952) Statistical Theory with Engineering Applications (1stedn.) Academic Press, New York, London.

5. Larsen SO (2006) Another aspect of truncation limits. Down's Screening News 13: 24.

6. Morris JK, Wald NJ (2005) Graphical presentation of distributions of risk in screening. Journal of Medical Screening 12: 155-160.

7. Barnes IM, Bestwick JP, Larsen SO (2007) Setting truncation limits for marker values in medical screening. Commentar. Journal of Medical Screening 14: 103-106.

8. Cuckle HS, van Lith JM (1999) Appropriate biochemical parameters in the first-trimester screening for Down syndrome. Prenat Diagn 19: 505-512.