Acta Genet Med Gemellol 39:307-316 (1990)
©1990 by The Mendel Institute, Rome

Sixth International Congress
on Twin Studies

A Mathematical Model for Recurrent Twinning

J.O. Fellman[1], A.W. Eriksson[2]

[1]Folkhälsan Institute of Genetics, Population Genetics Unit, Helsinki, Finland; [2]Institute of Human Genetics, Free University of Amsterdam, The Netherlands

Abstract. In an attempt to improve our understanding of the factors that affect human twinning, we further developed the models given by Hellin (1895) and Peller (1946). The connection between these models and our own model (" Fellman's law ") were studied. These attempts have resulted in a more general model, which was then applied to data from Åland Islands (1750-1939), Nîmes (1790-1875), Stuttgart (about 1790-1900) and Utah (1850-1900). The product of the mean sibship size and the total twinning rate can be considered as a crude estimate of the expected number of sets of twins in a sibship. The same can be said about the twinning parameter in our model. These estimates are in good agreement. If we consider twinning data only, we obtain the geometric distribution, and log (N_k), where N_k is the number of mothers with k twin maternities, is a linear function of the number of recurrences. Graphically, this property can easily be checked. For sibships containing three or more sets of twins, all four populations show higher values than expected, particularly the populations from Stuttgart and Utah, which data also show poor agreement according to a χ^2-test. A more exact model would demand more detailed demographic information, such as distribution of sibship sizes, age-specific twinning rates and temporal variations in twinning.

The osberved number of mothers in Åland with several recurrences of multiple maternities shows a considerable excess over the expected number as predicted by Peller's rule. The parameters in our model can be estimated by the maximum likelihood method and the obtained model fits the data better then Peller's model.

Key words: Mathematical models, Hellin's law, Peller's law, Fellman's law, Recurrence of multiple maternities, Twinning, Åland Islands

INTRODUCTION

The frequency of recurrent twinning is the repeat frequency of twinning among women who have already had one set of twins and it can be used to estimate the variability of

https://doi.org/10.1017/S0001566000005213 Published online by Cambridge University Press

twinning proneness in the population, regardless of the cause (genetic or nongenetic) of this variability [1]. In the majority of populations, 4-5% (extreme values 3-7%) of fertile women have been mothers of twins. In the period 1750-1949, about 8.5% of the mothers on the Åland Islands in the Northern Baltic Sea had twin or triplet maternities and 8.2% of the mothers with multiple maternities had a recurrent twinning. The fecundity in mothers in Åland with recurrent twin maternities was high, 7.7 maternities, that is, almost 10 children per mother, and a high proportion (about 50%) of unlike-sexed twins. This indicates that polyovulation is the chief cause of recurrent twinning. Previously, we presented and analysed data on recurrent twinning in Åland and a mathematical model was applied to the data in order to describe them in a simple way [3,4].

In this paper we will discuss a generalization of this model and relate it to Hellin's and Peller's laws. Furthermore, we relate it to fertility measurements in the corresponding populations. In this connection the total twinning rate and the mean size of the sibships are of special interest.

MATERIAL AND METHODS

Our data [3] and the data presented by Carmelli et al [2] are included in Table 1. Of course, the data do not include sibships without twin sets. Carmelli applied the same mathematical model as we did, but her parameter estimate differs from the maximum likelihood estimate we obtain in this study.

The data (Table 1) are from the Åland archipelago in southwestern Finland [3,4], from Nîmes in southern France [7], from Stuttgart in Baden-Württemberg in southern West-Germany [8], and from the Utah Mormon genealogical data base in Salt Lake City, USA [2].

Our generalization of the model used by Eriksson [3] and Carmelli et al [2] is:

(1) $$P_{st} = w^s r^t \, (1-w) \, (1-r) \qquad s, \, t = 0,$$

where P_{st} is the probability that an average sibship contains s twin sets and t triplet sets, w is the probability of a twin set, and r is the probability of a triplet set.

However, the available data set is usually truncated. We have no information about sibships without twin sets and triplet sets for which the probability is $(1-w)(1-r)$. Therefore, we have to cope with the conditional probability:

(2) $$P' = P \, (s,t \mid s \text{ or } t > 0) = \frac{w^s r^t \, (1-w) \, (1-r)}{r + w - rw}$$

The truncation complicates the model and we can only obtain approximate numerical maximum likelihood solutions. These can be obtained from the following formulas (cf Appendix):

(3a) $$w = \frac{S-N}{S} + r \, \frac{N}{S} \, \frac{1-w}{r+w-rw}$$

https://doi.org/10.1017/S0001566000005213 Published online by Cambridge University Press

Table 1 - Comparison of recurrent twinning (observed and expected) in Åland Islands (Finland), Nîmes (France), Stuttgart (W. Germany) and Utah (USA)

	k	Åland 1750-1939		Nîmes 1790-1875		Stuttgart c. 1788-1900		Utah 1850-1900	
		O	E	O	E	O	E	O	E
Number of twin sets	1	1515	1511.0	1156	1154.3	1493	1485.8	7982	7956.9
	2	121	128.1	48	50.3	81	93.9	649	694.6
	3	13	10.9	2	2.2	10	5.9	76	60.6
	4	2	1.0	1	0.2	2	0.4	11	5.9
Mean sibship size		4.6		4.7		6.5		8.1	
Total twinning		19.6		10.1		10.9		12.0	
$\chi^2_{(1)}$		1.21		0.26		6.96		9.39	
$\chi^2_{(2)}$		1.81		3.32		11.06		11.39	

Note: χ^2 pertains to the goodness of fit. In the upper line, $\chi^2_{(1)}$, mothers with 3 and 4 maternities have been pooled.

(3b)
$$r = \frac{Tw}{N - T + 2Tw + Tr - Trw}$$

where S = the total number of twin sets, T = the total number of triplet sets, N = the total number of sibships.

However, equations 3a and 3b do not give the explicit solutions of w and r, but they can be used for obtaining numerical solutions after an iterative process.

If the data contain no sibships with triplet sets, then T = 0, r = 0 and w = (S − N)/S. This formula is exact and it is the same as the one that was used by Eriksson [3].

RESULTS

Tabel 2 considers only twinning data and gives the estimate and its standard error. The estimator of w is given above and the standard error can be obtained as a special case of the Appendix, and is

$$w\,(1-w)^2/N$$

We may expect that in populations with a high total twinning rate and/or with high average sibship size, the parameter w shows high values. In Table 2 we give data for such a comparison. The product, mean sibship size times total twinning rate, is a crude esti-

https://doi.org/10.1017/S0001566000005213 Published online by Cambridge University Press

mate of the expected number of twin sets in a sibship. The same interpretation may be given to the parameter w in our model. If the average sibship size is c and the total twinning rate is tw, we obtain the approximate equation:

(4) $$w = tw \cdot c$$

We observe in Table 2 that there is a good agreement between our parameter estimates and the products. A similar connection can be assumed to hold for the parameter r and the total triplet rate. For Åland, 1750-1939, we observe $r = 0.001628$, $tr = 0.000354$, and $c = 4.6$. Hence, $c \cdot tr = 0.001628$.

Table 2 - Estimates according to the model (5) of recurrence of twinning in series from Åland, Nîmes, Stuttgart and Utah (see fext)

Population	w (1)	s_w (2)	Mean sibship size (3)	Total twinning rate (4)	(3) × (4)
Åland	0.0848	0.00656	4.6	0.0196	0.0902
Nîmes	0.0436	0.00575	4.7	0.0101	0.0475
Stuttgart	0.0632	0.00591	6.5	0.0109	0.0708
Utah	0.0873	0.00302	8.12	0.0120[a]	0.0974

[a] Crude estimates.

If we consider only twinning data we have the simple model:

(5) $$N_s = N\, w^s\, (1-w)$$

where N_s is the number of sibships with s twin sets.

A quick graphical check of the model may be obtained in the following way. We take the logarithms of both sides in formula 5. Now we get:

(6) $$\log (N_s) = s \log w + \log [N (1-w)].$$

Hence, $\log (N_s)$ is a linear function of the number of recurrences. The Figure shows the goodness of fit. We observe that, for $s > 3$, all data sets show higher values than expected. The Stuttgart data and the Utah data show the most marked discrepancies. This finding is in good agreement with the χ^2-tests in Table 1.

If we assume that our model and Hellin's law hold, then we obtain Peller's law in the following way. For triplets, we obtain an analogous formula to equation 4. Hence,

(7) $$r = c \cdot tr$$

where tr is the total triplet rate. According to Hellin's law,

(8) $$tr = (tw)^2$$

https://doi.org/10.1017/S0001566000005213 Published online by Cambridge University Press

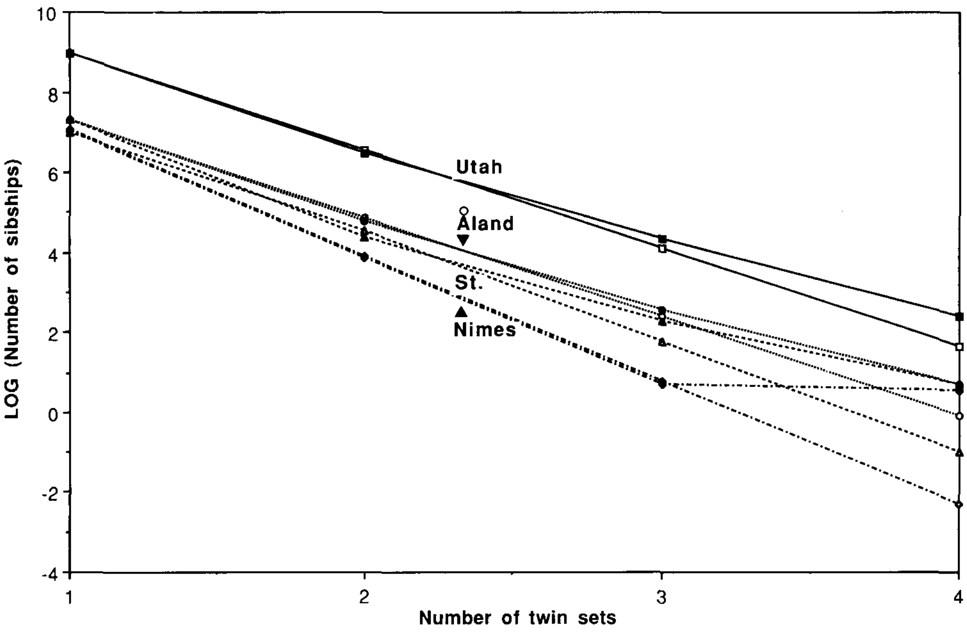

Figure. Comparison of recurrent twinning in series from Utah (1800-1950), Åland (1750-1939), ST., Stuttgart (c. 1790-1900) and Nîmes (1790-1875). The expected values are calculated according to formula (5) [3, p. 52].

If we combine equations 4, 7 and 8, we obtain:

$$\frac{r}{c} = tr = (tw)^2 = \left(\frac{w}{c}\right)^2, \text{ and}$$

(9) $$r = \frac{w^2}{c}$$

However, Peller uses $m = w^{-1}$ instead of w. If this variable is introduced, then:

(10) $$r = \frac{1}{m^2 c}$$

which is Peller's formula for the probability of a triplet set.

DISCUSSION

The most doubtful part of the deductions above is the use of Hellin's law. In our opinion, this is the most approximate of the laws discussed in this paper. It has to be taken into consideration that there is no biological justification for Hellin's law, but it serves as an approximate rule of thumb. From the 1750s, and until the last part of the 19th

https://doi.org/10.1017/S0001566000005213 Published online by Cambridge University Press

century, there is a fairly good agreement in Sweden between the actual frequencies of the different multiple maternities and the values estimated according to Hellin's law. However, all through the 20th century, there has been a deficiency in the frequency of triplet maternities both in Sweden and Finland. In the majority of populations there seems to be a much steeper downward secular trend of triplet than twin maternities. In spite of its high twinning rate, the Åland archipelago is the only known region in the North which, when Hellin's law is applied, has displayed an excess of triplets [3].

The data are too aggregated for more exact models. Such models should demand at least:

— information about the distribution of the sibship sizes;

— the effect of maternal age on the twinning rate (the age specific twinning rates). For a specific mother the chance of a twin maternity varies during her life;

— the temporal changes in the incidence of twinning. A useful set of data must cover a long period of time.

It must be taken into consideration that our model implies a long (infinite) fertile period. But in reality, the fertile period is limited and child bearing is confined to a short period (in our sample of mothers in Åland with at least two multiple maternities, about 15 years). Thus, the model is chiefly applicable to one or two multiple maternities. The main merit of our model is that it takes into account in a simple way the recurrence of twin sets and triplet sets when the probability of a multiple maternity in a sibship is estimated.

This mathematical model demonstrates that twinning is a reiteration (repetition) without memory. It can be assumed that there is a relation between the parameter w (and r) in our model and the general incidences for twin sets and triplet sets in the population and the size of the sibships. The parameter w can be considered as the probability of recurrence. When Peller's rule is applied, the observed number of mothers with several recurrences of multiple maternities in the Åland Islands exceeds the expected number (Table 3). Peller suggested that mc is close to Hellin's law unit n (our formula 4). For our data from Åland, $mc = 68.3$, but the observed $n = 51$ (twinning rate 19.6/1000 for the period 1750-1939). According to Eriksson [3] an average 8.5% of the Åland mothers should have multiple (twin or triplet) maternities once or several times. From this, we arrive at the value $m = 11.8$. The mean sibship size in Åland is $c = 4.6$, thus $mc = 54.3$, which better confirms the observed value of Hellin's law, $n = 51$. This is mainly due to the fact that all recurrences of twin maternities were taken into consideration when the model was used.

If we use our general model, we obtain the estimate:

$$\begin{aligned} w &= 0.08376 \pm 0.02053 \\ r &= 0.001628 \pm 0.000494 \end{aligned} \tag{11}$$

If we transform these estimates to estimates of Peller's parameters, m and c, we obtain:

$$m = 11.94$$

$$c = 4.31 \text{ (compare mean sibship size 4.6)}$$

$$mc = 51.45 \text{ (c.f. } (tw)^{-1} = 51).$$

https://doi.org/10.1017/S0001566000005213 Published online by Cambridge University Press

Table 3 - Occurrence of multiple maternities in Åland, 1740-1939, according to Peller's and to our rules

Maternities	Peller's formulae	Frequencies O	Frequencies E^a	Our formulae ^{b}	Frequencies $E^{c,d}$
Twins once	1:m	1515	1515.0	w	1503.6
Twins twice	$1{:}m^2$	113	113.0	w^2	125.9
Triplets once	$1{:}m^2c$	22	22.0	r	29.2
Twins three times	$1{:}m^3$	12	8.4	w^3	10.5
Twins and triplets	$1{:}m^3c$	8	1.6	wr	2.45
Twins four times	$1{:}m^4$	1	0.6	w^4	0.88
Twins twice and triplets once	$1{:}m^4c$	1	0.12	w^2r	0.20
Twins three times and triplets once	$1{:}m^5c$	1	0.01	w^3r	0.02

^{a} The values in the first three rows have been used to estimate m and c; m = 1515:113 = 13.4; c = 113:22 = 5.1.
^{b} The formula is complete after multiplication by $[(1-w)\ (1-r)]/(w+r-r \cdot w)$
^{c} $w = 0.08376 \pm 0.000494$; $r = 0.001628 \pm 0.000494$.
^{d} $\chi^2 = 18.49$ with 4 df. The last three rows are pooled together with all combinations of multiple maternities not observed.

In Table 3 we give the observed values from Åland, the estimates given by Eriksson [3] and our new estimates. The obtained $\chi^2 = 18.49$ with 4 df indicates a bad fit between the model and the data. Eriksson's estimates cannot be tested, since the expected numbers do not add to the observed total (= 1673).

The fecundity of mothers in Åland with repeated twinning was high, about 10 children per mother. It is noteworthy that, of the 32 sets of triplets born in Åland during 1750-1939, no less than 10 (almost 1/3) of the triplet sets were born to mothers with repeated multiple maternities ([3] and Table 3). In the series of 1586 mothers of twins from Stuttgart with 107 recurrent maternities, there were only two triplet maternities [8].

The 279 pairs of twins in Åland showed a high frequency of unlike-sexed twins (male + female is 50.0 ± 3%). This high rate of unlike-sexed twin maternities (definite DZ twin maternities) among mothers with recurrences in Åland indicates that polyovulation was the chief cause of repetition [3].

CONCLUSION

The tendency for repetition of multiple maternities (recurrence of twinning) is in rather good agreement with our general model which assumes that the chance of one further multiple maternity is approximately constant and independent of the number of previous multiple maternities. The model is in good agreement with the values observed in Åland and Nîmes, but in Weinberg's series from Stuttgart and in the Utah Mormon

https://doi.org/10.1017/S0001566000005213 Published online by Cambridge University Press

genealogy it shows a deficit of mothers with two multiple maternities and a surplus of mothers with three or more multiple maternities.

For all population, the model parameter w is in good agreement with the product of the total twinning rate and the average sibship size, formula 4. This agreement can be explained theoretically.

A more exact model would demand more demographic information, such as the distribution of sibship sizes, the age specific twinning rates and the temporal variations in twinning rates.

REFERENCES

1. Bulmer MG (1970): The Biology of Twinning in Man. Oxford: Clarendon Press.
2. Carmelli D, Hasstedt S, Anderson S (1981): Demography and genetics of human twinning in the Utah Mormon genealogy. In Gedda L, Parisi P, Nance WE (eds): Twin Research 3. Part A: Twin Biology and Multiple Pregnancy. New York: Alan R Liss, pp 81-93.
3. Eriksson AW (1973): Human twinning in and around the Åland Islands. Comment Biol 64: 1-159.
4. Eriksson AW, Fellman J, Forsius H (1973): The value of genealogical data in population studies in Sweden and Finland. In NE Morton (ed): Genetic Structure of Populations. Honolulu: Univ of Hawaii Press, pp 102-118.
5. Hellin D (1895): Die Ursache der Multiparität der uniparen Tiere überhaupt und der Zwillingsschwangerschaft beim Menschen insbesondere. München: Seitz & Schauer, p 70.
6. Peller S (1946): A new rule for predicting the occurrence of multiple births. Amer J Phys Anthropol 4:99-105.
7. Puech A (1877): De la répétition des accouchements multiples. Ann Gynecol Obstet 2:264-282.
8. Weinberg W (1901): Beiträge zur Physiologie und Pathologie der Mehrlingsgeburten beim Menschen. Arch Ges Physiol 88:346-430.

Correspondence: Prof. Johan Fellman, Swedish School of Economics, Arkadiagatan 22, SF-00100 Helsinki 10, Finland.

https://doi.org/10.1017/S0001566000005213 Published online by Cambridge University Press

Appendix

Deduction of the Maximum Likelihood Estimates

Consider the probability P_{st} that an average sibship contains s twin sets an t triplet sets. Then

(A1) $$P_{st} = w^s r^t (1-w)(1-r);\ s, t \geq 0$$

where w is the probability of a twin set in a sibship of average size and r is the probability of a triplet set.

However, the available data set is usually truncated. We have no information about sibships without twin sets and triplet sets. The probability of this is s=0, t=0, $P_{00} = (1-w)(1-r)$. Therefore, we have to cope with the conditional probability

(A2) $$P'_{st} = P(s,t \text{ at least } s \text{ or } t>0) = \frac{w^s r^t (1-w)(1-r)}{r+w-r \cdot w}$$

The truncation complicates the model and we can only obtain approximate numerical maximum likelihood solutions.

If we assume that we observe n_{st} sibships with s twin sets and t triplet sets then the likelihood function is

(A3) $$L(w,r) = \prod_{s,t} \left[\frac{w^s r^t (1-w)(1-r)}{(r+w-r \cdot w)}\right]^{n_{st}} = \frac{(1-w)^N (1-r)^N}{(r+w-r \cdot w)^N} \cdot w^S r^T$$

where $N = \Sigma_s \Sigma_t n_{st}$, $T = \Sigma_s \Sigma_t t \cdot n_{st}$, $S = \Sigma_s \Sigma_t s \cdot n_{st}$.

In order to simplify the calculations we study and maximize the log likelihood function

(A4) $$l(w,r) = S \ln w + T \ln r + N \ln (1-w) + N \ln (1-r) - N \ln (r+w-r \cdot w)$$

We obtain the partial derivatives

(A5a) $$\frac{\delta l}{\delta w} = \frac{S}{w} - \frac{N}{1-w} - \frac{N(1-r)}{(r+w-r \cdot w)}$$

(A5b) $$\frac{\delta l}{\delta r} = \frac{T}{r} - \frac{N}{1-r} - \frac{N(1-w)}{(r+w-r \cdot w)}$$

https://doi.org/10.1017/S0001566000005213 Published online by Cambridge University Press

Maximum is obtained if these derivatives are zero. Hence, after some calculations we get the equations

(A6a) $$w = \frac{S-N}{S} + r \cdot \frac{N}{S} \frac{1-w}{r+w-r \cdot w}$$

(A6b) $$r = \frac{Tw}{N-T+2Tw+Tr-Trw}$$

The equations (A6a) and (A6b) do not give explicite solutions of w and r. We observe that on the right hand side are the unknown parameters w and r still. The numerical solution can be obtained, so that we start from initial values of w and r (say w_0 and r_0). After some iterations, our estimates converge. If the data contain no triplet sets, then T = 0, r = 0 and w = (S−N)/S. This formula for w is the same used by Eriksson [3].

In order to estimate the standard error of the estimates we derive once more

$$\frac{\delta^2 l}{\delta w^2} = -\frac{S}{w^2} - \frac{N}{(1-w)^2} + \frac{N(1-r)^2}{(r+w-rw)^2}$$

$$\frac{\delta^2 l}{\delta r^2} = -\frac{T}{r^2} - \frac{N}{(1-r)^2} + \frac{N(1-w)^2}{(r+w-rw)^2}$$

$$\frac{\delta^2 l}{\delta w \, \delta r} = \frac{N}{(r+w-rw)^2}$$

(A7a) $$E\left(\frac{\delta^2 l}{\delta w^2}\right) = -\frac{Nr\,(1+w-wr)}{w\,(r+w-rw)^2} - \frac{N}{1-w^2}$$

(A7b) $$E\left(\frac{\delta^2 l}{\delta r^2}\right) = -\frac{Nw\,(1+r-rw)}{r\,(r+w-rw)^2} - \frac{N}{(1-r)^2}$$

(A7c) $$E\left(\frac{\delta^2 l}{\delta w \delta r}\right) = \frac{N}{(r+w-rw)^2}$$

If this method is applied to the occurrence of multiple maternities in the Åland Islands [3, p 63], we obtain the estimates w = 0.08376 and r = 0.001628.

The equations (A7a), (A7b) and (A7c) give the information matrix and its inverse gives the estimated variances.

For the estimated w, r, and N = 1673 we obtain σ_w = 0.020533 and σ_r = 0.0004936.

https://doi.org/10.1017/S0001566000005213 Published online by Cambridge University Press