Brazilian Journal of Genetics, 19, 4, 659-663 (1996)
(Revista Brasileira de Genética)

METHODOLOGY

A simple model for the estimation of congenital malformation frequency in racially mixed populations

J.S. Lopez-Camelo, P.H. Cabello and M.G. Dutra

ABSTRACT

A simple model is proposed, using the method of maximum likelihood to estimate malformation frequencies in racial groups based on data obtained from hospital services. This model uses the proportions of racial admixture, and the observed malformation frequency. It was applied to two defects: postaxial polydactyly and cleft lip, the frequencies of which are recognizedly heterogeneous among racial groups. The frequencies estimated in each racial group were those expected for these malformations, which proves the applicability of the method.

INTRODUCTION

One of the principal goals of epidemiological genetic studies of congenital malformations is to determine the influence of genetic and environmental factors on their etiology. In this context, several authors have studied the relationships between the incidence of congenital defects and racial antecedents in populations (Altemus and Ferguson, 1965; Erickson, 1976; Christianson *et al.*, 1981). A correlation between the frequency of congenital malformations and racial composition suggests genetic control. Nevertheless, interpretation is often difficult, due to the multiple factors that may change the rate of malformations, such as the different criteria used to describe the malformations, maternal age, parity or socioeconomic level. Chakraborty and Weiss (1986) argue that increased environmental contributions to the etiology of multifactorial diseases obscures the relationship between racial admixture and the frequency of the disease.

The aim of this paper is to derive a general and simple model to estimate malformation frequency among racial groups using hospital-base data of admixed populations. Krieger *et al.* (1965) designed a method to estimate the racial proportions in tri-hybrid populations using genetic polymorphisms. These authors estimated such proportions using previously known gene frequency of the ancestral populations. On the other hand, they suggest that if the proportion of racial admixture is known, one can easily estimate the gene frequency in each ancestral population.

MATERIAL AND METHODS

Sample

A sample of 724,149 births was obtained during 1982-1991 from 35 hospitals of eight South American countries which participated in the Latin-American Collaborative Study of Congenital Malformations (ECLAMC) (Castilla and Orioli, 1983).

The 35 hospitals were grouped into 19 different South American regions. Table I shows the number of births, number of hospitals in each region, and the

Departamento de Genética, Instituto Oswaldo Cruz, Av. Brasil 4365, 21045-900 Rio de Janeiro, Brasil. Phone: 55-21-5984358; Fax: 55-21-5903495. Send correspondence to J.S.L.-C.

Table I - Proportion of racial admixture observed in Latin American populations.

Country region	Code	Total births	Number of hospitals	Racial admixture proportion		
				Black	Native	White
Argentina						
Buenos Aires	BUE	105307	6	0.1070	0.3740	0.5190
Cordoba	CDB	74448	2	0.0000	0.3348	0.6652
Mendoza	MDZ	12695	1	0.0000	0.3615	0.6385
Pampa	PAM	38346	3	0.0902	0.2886	0.6212
Patagonia	PAT	5823	3	0.1692	0.5397	0.2911
Bolivia						
La Paz	LPZ	29960	1	0.0009	0.8259	0.1732
Brazil						
São Paulo	SAO	72678	4	0.3031	0.1335	0.5634
Rio de Janeiro	RIO	5795	1	0.5206	0.0776	0.4018
Florianópolis	FLP	33710	1	0.0000	0.1733	0.8267
Porto Alegre	POA	25836	2	0.0933	0.1675	0.7392
Chile						
Santiago	STG	47130	2	0.2290	0.3669	0.4041
Colombia						
Barranquilla	BRQ	4093	1	0.8434	0.1565	0.0001
Bogota	BGT	19303	2	0.1146	0.4658	0.4196
Peru						
Lima	LIM	36613	1	0.0000	0.5584	0.4416
Uruguay						
Montevideo	MVD	64023	1	0.1540	0.2124	0.6336
Venezuela						
Maracaibo	MCB	51218	1	0.0000	0.4595	0.5405
Coro	COR	30621	1	0.1793	0.4111	0.4096
Ciudad Ojeda	OJE	8211	1	0.0000	0.3912	0.6088
Ciudad Bolivar	BLV	58339	1	0.1891	0.4607	0.3502

proportions of racial admixture. We used ABO and Rh blood systems from 32,945 controls to estimate the proportion of racial admixture through the maximum likelihood method proposed by Krieger *et al.* (1965).

To test the model, two malformations were used: postaxial polydactyly (N = 1058) and cleft lip (N = 533), which show known racial differences. Cases without other associated malformations were selected since they are more etiologically homogeneous.

The model

Consider a set of data collected from m regions. Let Q_i be the frequency of a given malformation in the ith region which can be expressed as

$$Q_i = B_i\alpha_1 + I_i\alpha_2 + W_i\alpha_3 \qquad (B_i + I_i + W_i = 1)$$

where, B, I and W are the proportions of the black, native (Indian) and white ethnic groups, respectively, and α_1, α_2 and α_3 are the parameters that represent the frequency of the malformation in blacks, natives and whites, respectively.

Considering that the observations correspond to independent samples from a binomial distribution, the total probability, as a function of the α parameters, can be expressed as

$$P = \prod_{i=1}^{m} \left\{ \begin{matrix} N_i \\ r_i \end{matrix} \right\} Q_i^{r_i} \ (1-Q_i)^{N_i - r_i}$$

where N_i and r_i represent the total number of individuals and the number of affected children, respectively.

The α parameter estimates may be obtained using the maximum likelihood (ML) method. The solution of the four equations is accomplished by the iteration process of Newton-Raphson, modified (Krieger *et al.*, 1965; Elston, 1969). In this way estimation of the α values consists in the correction of given initial values, through the following iteration process:

$$A_j = A_{j-1} + K^{-1}U$$

where A_j is a vector constituted by α_i (i = 1, 2, 3) values evaluated at the jth iteration, and U is a vector of ML-scores whose uth element is calculated as

$$\Sigma \frac{r_i}{Q_i}\left[\frac{\partial Q_i}{\partial \alpha_u}\right] + \frac{(N_i - r_i)}{(1-Q_i)}\left[\frac{\partial (1-Q_i)}{\partial \alpha_u}\right]$$

and K is a 3 x 3 symetric matrix, information matrix, whose (u, v)th element is

$$\Sigma \frac{r_i}{Q_i^2}\left[\frac{\partial Q_i}{\partial \alpha_u}\right]\left[\frac{\partial Q_i}{\partial \alpha_v}\right] + \frac{(N_i - r_i)}{(1-Q_i)^2}\left[\frac{\partial (1-Q_i)}{\partial \alpha_u}\right]\left[\frac{\partial (1-Q_i)}{\partial \alpha_v}\right]$$

The information inverse matrix corresponds to the αs' variance-covariance matrix. The fit between the estimate values and the real parameter values can be tested by:

$$\chi^2 = U^t K^{-1} U, \text{ with three degrees of freedom}$$

where, U^t is the transpose of the U vector scores.

To test the fit of the model to the observed data in each region, the χ^2 test was used. The fit was measured by summation of the χ^2 values obtained for each region, with as many degrees of freedom as the number of regions minus the number of estimated parameters minus one.

RESULTS AND DISCUSSION

Table II shows the estimates and standard deviation of the malformation frequency for each race. The postaxial polydactyly frequency in blacks is reported to be about 100 among each 10,000 newborns (Chung and Myrianthopoulos, 1968; Erickson, 1976; Kromberg and Jenkins, 1982). We found a lower frequency (Table II). Probably black Latin American populations are more admixed than American blacks.

A high frequency of cleft lip, with or without cleft palate, has been described among oriental populations (21/10,000), an intermediate one in Caucasians (13/10,000) and a low frequency in blacks (4/10,000) (Neel, 1958; Chung *et al.*, 1974; Chung and Kau, 1985; Marazita *et al.*, 1986; Melnick *et al.*, 1986). We also found a high frequency in natives (Table II), a group supposedly presenting a strong component of oriental ancestry. Similar results have been observed for Mexican populations (12/10,000) which have a high Amerindian component (Lisker *et al.*, 1986; C.M.W., 1991). These observations reinforce the phylogenetic relationship between Amerindians and Oriental populations.

In some geographic regions, there was little agreement between the observed and the expected number of cases by the model (Tables III-IV). Regional differences in the definition and ascertainment of cases or environmental heterogeneity could have influenced the results. The model is deficient since it does not consider such heterogeneity, which is very important when the defect is determined by multifactorial causes. In this sense, a possible geographic cluster of cleft lip in La Paz (Orioli and Grisolia, 1992) could explain the excess of observed cases. Also, the model does not distinguish between familial and sporadic cases, an important fact since there is evidence that over 50% of cleft lip cases, with no other malformations associated, are sporadic (Marazita *et al.*, 1986).

Another source of variation is the definition of the native group, which is very heterogeneous. In Argentina, Bolivia, Peru and Colombia this group is an admixture of Latin Europeans and Amerindians, but in the different regions of Brazil and Venezuela the characterization of the native depends on the participation of blacks. The high proportion of black genes in Santiago (Chile) is not expected; this observation had been previously related in the

Table II - Estimated frequencies and standard errors (per 10,000) of the malformations in three racial groups from Latin America.

Malformation	N	Black	Native	White	χ^2
Postaxial Polydactyly	1058	60.99 ± 3.79	8.58 ± 1.68	8.23 ± 1.26	0.062
Cleft lip	533	2.73 ± 2.02	15.02 ± 1.40	3.29 ± 0.96	0.229

$\chi^2 = U^t K^{-1} U$ (DF = 3).

Table III - Goodness of fit test on postaxial polydactyly with and without application of the model.

Region*	Observed	Applying the model		Simple binomial	
		Expected	Chi square	Expected	Chi square
BUE	85	147.48	26.47	153.86	30.82
CDB	96	62.14	18.44	108.77	1.50
MDZ	13	10.61	0.54	18.55	1.66
PAM	21	50.20	16.98	56.02	21.89
PAT	5	10.10	2.58	8.51	1.45
LPZ	27	25.68	0.07	43.77	6.43
SAO	232	176.39	17.53	106.18	149.09
RIO	28	20.70	2.57	8.47	45.03
FLP	22	27.94	1.26	49.25	15.08
POA	40	34.13	1.01	37.75	0.13
STG	29	96.34	47.06	68.86	23.07
BRQ	9	21.61	7.36	5.98	1.53
BGT	19	26.94	2.34	28.20	3.00
LIM	14	30.85	9.21	53.49	29.15
MVD	76	105.19	8.12	93.54	3.29
MCB	55	42.98	3.36	74.83	5.25
COR	98	54.61	34.47	44.74	63.40
OJE	10	6.87	1.42	12.00	0.33
BLV	179	107.18	48.12	85.23	103.17
Total	1058	1058.00	248.91	1058.00	505.27

*See Table I.

Table IV - Goodness of fit test on cleft lip with and without application of the model.

Region*	Observed	Applying the model		Simple binomial	
		Expected	Chi square	Expected	Chi square
BUE	68	80.27	1.87	77.51	1.17
CDB	69	53.77	4.32	54.80	3.68
MDZ	13	9.58	1.23	9.34	1.43
PAM	30	25.43	0.83	28.22	0.11
PAT	12	5.56	7.50	4.29	13.86
LPZ	69	38.89	23.31	22.05	99.97
SAO	36	34.08	0.11	53.49	5.72
RIO	3	2.27	0.24	4.27	0.38
FLP	16	17.97	0.21	24.81	3.13
POA	17	13.46	0.94	19.02	0.21
STG	47	35.21	3.96	34.69	4.37
BRQ	2	1.90	0.00	3.01	0.34
BGT	15	16.35	0.11	14.21	0.04
LIM	24	36.05	4.02	26.95	0.32
MVD	39	36.49	0.17	47.12	1.40
MCB	28	44.48	6.11	37.70	2.50
COR	18	24.55	1.74	22.54	0.91
OJE	3	6.47	1.86	6.04	1.53
BLV	24	50.12	13.61	42.94	8.35
Total	533	533.00	72.15	533.00	149.42

*See Table I.

ECLAMC study (Dutra, 1989). This probable mistyping of racial admixture estimated in Chile could be the cause of deviations from the model.

The significance of the model was evaluated by subtraction of χ^2 of goodness of fit without the model from the χ^2 obtained by application of the model, with two degrees of freedom. The expected numbers, without application of the model, were obtained from the binomial simple distribution, with parameter p calculated as the total number of affected children divided by the total number of newborns.

The χ^2 values of goodness of fit were strongly diminished when the model was applied (Tables III-IV). In cleft lip the χ^2 of significance of the model was 77.27 (with 2 d.f.); whereas in postaxial polydactyly the significance was more evident ($\chi^2 = 256.36$; 2 d.f.). This was expected since postaxial polydactyly probably has an oligogenic etiology, involving a few genes (Paz and Castilla, 1976).

This model may be applied using the proportion of racial admixture obtained from any source of information, e.g. census data, genetic markers, etc. The estimates depend on the precision of these racial proportions.

RESUMO

É proposto um modelo simples, fazendo uso do método da máxima verossimilhança, para estimar as freqüências de malformações em grupos raciais, baseado em dados obtidos em serviços hospitalares. Este modelo usa as proporções de mistura racial e a freqüência observada da malformação. O método foi aplicado a dois defeitos: polidactilia pós-axial e lábio leporino, cujas freqüências são reconhecidamente heterogêneas entre grupos raciais. As estimativas obtidas em cada grupo racial foram as esperadas para estas malformações, o que prova a aplicabilidade do método.

REFERENCES

Altemus, L.A. and **Ferguson, A.D.** (1965). Comparative incidence of birth defects in negro and white children. *Pediatrics 36*: 56-61.

Castilla, E.E. and **Orioli, I.M.** (1983). El Estudio Colaborativo Latinoamericano de Malformaciones Congenitas: ECLAMC/MONITOR. *Interciencia 8*: 271-278.

Chakraborty, R. and **Weiss, K.M.** (1986). Frequencies of complex diseases in hybrid populations. *Am. J. Phys. Anthropol. 70*: 489-503.

Christianson, R.E., Van Den Berg, B.J., Milkovich, L. and **Oechsli, F.** (1981). Incidence of congenital anomalies among White and Black live births with long-term follow-up. *Am. J. Public Health 71*: 1333-1341.

Chung, C.S. and **Myrianthopoulos, N.C.** (1968). Racial and prenatal factors in major congenital malformations. *Am. J. Hum. Genet. 20*: 44-60.

Chung, C.S. and **Kau, M.C.** (1985). Racial differences in cefalometric measurements and incidence of cleft lip with or without cleft palate. *J. Craniofac. Genet. Dev. Biol. 5*: 341-349.

Chung, C.S., Ching, G.H.S. and **Morton, N.E.** (1974). A genetic study of cleft lip and palate in Hawaii. II. Complex segregation analysis and genetic risk. *Am. J. Hum. Genet. 26*: 177-188.

C.M.W. (1991). *Congenital Malformations Worldwide: A Report from the International Clearinghouse for Birth Defects Monitoring Systems*. Elsevier Science Publishers B.V., Amsterdam, pp. 89-94.

Dutra, M.G. (1989). Análise da etnicidade em populações urbanas da América do Sul. Doctoral thesis, Universidade Federal do Rio de Janeiro, Rio de Janeiro, RJ.

Elston, R.C. (1969). Iteration problems. In: *Computer Applications in Genetics* (Morton, N.E., ed.). University of Hawaii Press, Hawaii, pp. 30-37.

Erickson, J.D. (1976). Racial variations in the incidence of congenital malformations. *Ann. Hum. Genet. 39*: 315-320.

Krieger, H., Morton, N.E., Mi, M.P., Azevedo, E., Freire-Maia, A. and **Yasuda, N.** (1965). Racial admixture in northeastern Brazil. *Ann. Hum. Genet. 29*: 113-125.

Kromberg, J.G.R. and **Jenkins, T.** (1982). Common birth defects in South African Blacks. *S. Af. Med. J. 62*: 599-602.

Lisker, R., Perez-Briceno, R., Granados, J., Babinsky, V., De Rubens, J., Armendares, S. and **Buentello, L.** (1986). Gene frequencies and admixture estimates in a Mexico City population. *Am. J. Phys. Anthropol. 71*: 203-207.

Marazita, M.L., Spence, M.A. and **Melnick, M.** (1986). Major gene determination of liability to cleft lip with or without cleft palate: multiracial view. *J. Craniof. Genet. Dev. Biol. 2* (Suppl.): 89-97.

Melnick, M., Marazita, M.L. and **Hu, D.N.** (1986). Genetic analysis of cleft lip with or without cleft palate in Chinese kindreds. *Am. J. Med. Genet. 2* (Suppl.): 183-190.

Neel, J.V. (1958). A study of major congenital defects in Japanese infants. *Am. J. Hum. Genet. 10*: 398-445.

Orioli, I.M. and **Grisolia, S.V.** (1992). Geographic variation in the frequency of selected birth defects in Latin America. *Genet. Reprod. Epidemiol. Res. Soc. 1*: 83-96.

Paz, J.E. and **Castilla, E.E.** (1976). Minimal loci number estimate for postaxial polydactyly in humans. *Excerpta Médica, International Congress Series 397*: 192.

(Received May 30, 1994)