Hum Genet (1993) 92:598–604

human genetics

Logistic regression model to estimate the risk of unbalanced offspring in reciprocal translocations

Christine Cans[2], Olivier Cohen[1], Christian Lavergne[3], Marie-Ange Mermet[1], Jacques Demongeot[2], Pierre Jalbert[1]

[1] Genetics Laboratory, Medical School, Joseph Fourier University Grenoble, F-38706 La Tronche Cedex, France
[2] Department of Statistics and Computer Sciences, Medical School, Joseph Fourier University Grenoble, F-38706 La Tronche Cedex, France
[3] Modelisation and Computation Laboratory, Joseph Fourier University Grenoble, BP 53X, F-38041 Grenoble Cedex, France

Received: 25 February 1993 / Revised: 1 May 1933

Abstract. The aim of this study was to estimate the risk of viable unbalanced offspring for a parental carrier of reciprocal translocation. On a large computerized database of reciprocal translocations we used logistic regression to model this risk. The status of the progeny is the outcome variable. Explanatory covariates are cytogenetic characteristics of the translocation, age and sex of the parental carrier, and potential viability of the gametes. The results obtained by the logistic model demonstrate the important role of certain variables such as the sex of the parental carrier and the R band length of the translocated segments. Within the group of lower risk (risk of viable unbalanced offspring less than 5%), 97% of the individuals are correctly classified with this model. For this group, the choice prenatal diagnosis can be best discussed by considering both the risk for viable unbalanced offspring and the risk of induced abortion following prenatal diagnosis.

Introduction

The relative frequency of reciprocal translocation is well known: 1 couple in 600 is a carrier (Crandall et al. 1980). Over the past 20 years, several authors have investigated the different types of offspring in reciprocal translocations. They have shown, either from small or large studies (Daniel 1979; Boué and Gallano 1984; Petrosky and Bergaonkar 1984; Davis et al. 1985; Stengel-Rutkowski et al. 1988), that many factors influence the occurrence probability of the different meiotic segregation modes. As the number of different reciprocal translocations (different chromosomes involved and different breakpoints) is extremely high, pedigrees are numerous and the estimation of the risk based on the description of one pedigree gives very inaccurate and incomplete results. This risk ranges from 0 to 40% for most translocations. To our knowledge, Thomas's group (personal communication 1987) has been the only one to use inferential methods and a model to estimate the risk of different meiotic segregation modes in reciprocal translocations.

The aim of the present investigation is to emphasize the advantage of generating, a model to estimate the risk of viable unbalanced offspring for a parental carrier of reciprocal translocation. This model may be helpful to geneticists in genetic counselling, which involves advising parents carrying reciprocal translocation about the possibilities of prenatal diagnosis. When a choriocentesis is proposed account must be taken of the risk of induced abortion for such a couple (with sterility history and/or no living progeny), as it can be higher than the risk of unbalanced offspring. On the other hand, proposing no, or further, prenatal diagnosis, on the assumption that either the fetus will be unviable or the child will be either a normal or a healthy carrier, can lead to serious problems if this hypothesis is not confirmed. Even if the prenatal diagnosis policy differs from one medical center to another, geneticists still need accurate estimations of the risk of viable unbalanced offspring.

Materials and methods

Database

Since 1989 a large computerized database has been available in the cytogenetic laboratory at Grenoble. The data on reciprocal translocations originates from various sources: cytogenetic laboratories, the literature, and prenatal diagnosis consultations. Only translocations between autosomes, karyotyped and inherited translocations have been selected for this study (Table 1). The breakpoints and the dimensions of the chromosomic segments were determined to an accuracy of 0.5 mm on each of the 400 bands with reference to the international system (ISCN 1981). The offspring of a healthy carrier is the statistical unit. This offspring can be either a normal individual or a healthy carrier in the case of alternate meiotic segregation mode, or either an unbalanced child or an unviable fetus (spontaneous abortion) in the case of 2:2 or 3:1 unbalanced meiotic

Correspondence to: C. Cans

Table 1. Cases selected from the database

	Cytogenetic laboratories 1975–1986 (Jalbert 1988)	Literature 1971–1990	Prenatal diagnosis consultations 1977–1981 (Boué 1984)	Total
Normal or unviable	1177	3458	1561	6196
MSMR	127	889	175	1191
(16.1%)	*(9.7%)*	*(20.5%)*	*(10.1%)*	
Total	1304	4347	1736	7387

These 7387 units issue from 1582 different pedigrees, and represent 1252 reciprocal translocations

	Xme 2	Xme 3	Xme 4	Xme 5	Xme 6	Xme 7	Xme 8	Xme 9	Xme 10	Xme 11	Xme 12	Xme 13	Xme 14	Xme 15	Xme 16	Xme 17	Xme 18	Xme 19	Xme 20	Xme 21	Xme 22
Xme 1	30	39	146	52	26	50	48	**39**	48	19	11	*16*	*23*	*11*	54	49	49	1	62	17	52
Xme 2		17	18	67	52	55	71	**26**	53	4	56	*18*	*39*	*27*	22	29	41	3	16	10	9
Xme 3			41	16	42	38	51	**22**	10	16	22	*23*	*23*	*52*	26	14	24	8	7	18	23
Xme 4				36	30	35	98	**72**	56	36	46	*64*	*12*	*57*	51	15	108	0	29	46	50
Xme 5					20	47	25	**50**	88	74	22	*57*	*44*	*101*	26	29	60	13	3	41	19
Xme 6						15	25	**36**	18	37	12	*33*	*8*	*21*	3	13	24	5	54	16	27
Xme 7							51	**71**	29	21	49	*39*	*28*	*18*	5	3	29	5	17	41	28
Xme 8								**85**	24	21	32	*65*	*35*	*35*	10	9	24	0	4	0	32
Xme 9									**40**	**19**	**31**	**51**	**42**	**89**	**18**	**5**	**55**	**28**	**50**	**30**	**101**
Xme 10										21	36	*23*	*55*	*23*	8	51	86	0	15	49	42
Xme 11											13	*36*	*10*	*4*	38	12	14	0	6	36	***675***
Xme 12												*3*	*30*	*4*	12	6	23	0	8	59	6
Xme 13													*23*	*70*	*6*	*25*	*80*	*0*	*11*	*23*	*33*
Xme 14														*9*	*2*	*4*	*36*	*0*	*7*	*4*	*9*
Xme 15															*6*	*7*	*3*	*6*	*4*	*26*	*12*
Xme 16																14	21	7	7	92	8
Xme 17																	11	7	0	14	28
Xme 18																		2	24	23	8
Xme 19																			20	2	10
Xme 20																				21	26
Xme 21																					2

Fig. 1. Chromosomes involved in the reciprocal translocation. 1) Translocations with Xmes 1 to 8, 10 to 12, 16 to 20, $n = 3452$; 2) translocations with Xme 9, $\boldsymbol{n = 960}$; 3) translocations with Xmes 13, 14, 15, $n = 1336$; 4) translocations with Xmes 21, 22, $n = 964$; 5) translocations t(11;22), $\boldsymbol{n = 675}$; $N = 7387$

segregation modes. For the response outcome variable, we did not use the meiotic segregation mode but, instead, the status of the progeny as the dichotomic variable (1 if the unbalanced child was viable, 0 otherwise). As the phenotypic expression of the great majority of these unbalanced children takes the form of multiple malformations associated with mental retardation we refer to their condition as Malformation syndrome mental retardation (MSMR). Pregnancies ending in spontaneous miscarriage, therapeutic abortion, or stillbirth have been grouped with normal or healthy carrier individuals (modality 0 called 'normal or unviable'). The two main reasons for this are that: (1) for numerous spontaneous abortions there are no data in the database, creating a significant selection bias for the meiotic segregation mode and (2) geneticists are basically more interested in the probability of viable unbalance outcome than in the mechanisms of meiotic segregation modes.

On the basis of previous studies, explanatory variables (covariates) have been selected from the database. As preliminary studies have established the relationship between the segregation mode and the R bands contained within the chromosomic segments (Cohen et al. 1992; Cans et al. 1993), the length of the centric and translocated segments were taken as their R band lengths. Chromosomes involved in the reciprocal translocation were grouped into five different patterns (Fig. 1). According to the results obtained in recent studies (Daniel et al. 1988; Jalbert et al. 1988; Stengel-Rutkowski et al. 1988) the sex and the age of the carrier parent seem to influence the segregation mode, and these were therefore also selected. Details of sex and age of patients were available for 75% and 25% of cases, respectively. Three different patterns were considered for the breakpoint locus of the two chromosomes involved: both on short arms, both on long arms, one on a short and one on a long arm. To take the viability of the potential chromosomal imbalance into account (De Arce et al. 1986), we used the number of potential viable unbalanced gametes according to a recently defined viability criterion (Jalbert et al. 1992). The choice of this criterion instead of Daniel's (Daniel 1979) was determined by the fact that it may avoid the risk of false-negatives (viable unbalanced child although outside Daniel's criteria). A list and the distribution of these covariates are shown in Table 2.

Statistical analysis

In order to investigate the effect of the covariates on the response probablility, logistic regression appears to be the most appropriate choice for these binary data. First, a univariate analysis is performed for each explicative variable. The Wald test is used to choose the best pattern for each qualitative variable. When the Score test ist not significant variables are not retained for the subsequent steps. If two variables are highly correlated, the one leading to the greatest change in deviance is maintained. Then, multivariate analysis is performed starting with the full model. Variables are excluded from the model either if the Wald test is not significant or if the Score test is not significative when variables are omitted separately. However, two exceptions are allowed: a variable may be retained (1) if enclosed further in an interaction or (2) at a geneticist's request (possible confounding influence). Only second-order interactions have been considered, because third-or-

Table 2. Covariates in the sample

			7387 cases	OR (descriptive analysis)
Centrics and translocated chromosome segments	v1	CS1R bands	0 to 71.6 Mean = 38.1	OR (CS1R > 38) = 1.05 [0.92–1.20]
	v2	CS2R bands	0 to 69.3 Mean = 20.4	OR (CS2R > 20) = 0.90 [0.79–1.02]
	v3	TS1R bands	0 to 40.6 Mean = 11.2	OR (TS1R > 10) = 0.67 [0.59–0.76]
	v4	TS2R bands	0 to 37.0 Mean = 9.5	OR (TS2R > 10) = 0.71 [0.62–0.81]
Breakpoint locus	v5	Arm 'qq'	n = 3441 (47%)	
		Arm 'pp'	n = 1079 (14%)	OR (arm 'pp') = 1.05 [0.88–1.25]
		Arm 'pq'	n = 2867 (39%)	
Carrier parent's sex	v6	Maternal origin	n = 3540 (48%)	OR (maternal) = 1.49 [1.30–1.71]
		Paternal origin	n = 2024 (27%)	
		Unknown origin	n = 1823 (25%)	
Carrier parent's age	v7	Maternal age	15 to 51 Mean = 28.0	
	v8	Paternal age	16 to 72 Mean = 29.4	
Chromosomes involved	$v9_{(1)}$	Xmes 1 to 8, 10 to 12, 16 to 20	n = 3452 (48%)	OR ($v9_{(1)}$) = 0.66 [0.58–0.75]
	$v9_{(2)}$	Xme 9	n = 960 (13%)	OR ($v9_{(2)}$) = 1.33 [1.11–1.59]
	$v9_{(3)}$	Xmes 13, 14, 15	n = 1336 (18%)	OR ($v9_{(3)}$) = 1.04 [0.88–1.23]
	$v9_{(4)}$	Xmes 21, 22	n = 964 (13%)	OR ($v9_{(4)}$) = 1.72 [1.45–2.05]
	$v9_{(5)}$	t(11;12)	n = 675 (8%)	OR ($v9_{(5)}$) = 0.93 [0.74–1.17]
Viability	v10	Number of viable gametes	0 to 12 Mean = 5.6	OR (viab < 3) = 0.56 [0.44–0.72]
				OR (viab > 8) = 1.23 [1.02–1.49]

All the OR refer to 5564 units with known origin and measure the association between one of the covariates and the outcome response (MSMR vs normal or unviable). Their IC have been calculated with maximum likelihood point estimation

Table 3. Wald test for each covariate (logistic regression on 5564 units)

		Univariate analysis	Multivariate analysis (without interaction terms)	Retained the model
v1	CS1R bands	NS	Not included	No
v2	CS2R bands	S	Not included	No
v3	TS2R bands	S	S	Yes
v4	TS2R bands	S	S	Yes
v5	Arm qq	S	NS	Yes
	Arm pp[a]	S	NS	Caused to remain
	Arm pq	NS	Not included	
v6	Maternal origin[b]	S	S	Yes
	Paternal origin	S	NS	Caused to remain
v7	Maternal age[c]	S	S	No
v8	Paternal age	NS	Not included	No
v9	Xmes 1 to 8, 10 to 12, 16 to 20	S	S	
	Xme 9	S	S	
	Xme 13, 14, 15	S	S	Yes
	Xme 21, 22	S	S	
	t(11;22)	S	S	
v10	Viability[d]	S	S	No

[a] As included in an interaction term this modality has been caused to remain in the model
[b] Origin was caused to remain in the model on request of cytogeneticians
[c] For the age of the carrier parent, logistic regression has been realised on the 2099 units with information on age available
[d] No longer significant when interaction terms are added

der interactions did not improve the Score test and are quite difficult to explain. With this strategy, we expected to obtain a model which includes the most influential variables with the smallest number of parameters (Greenland 1989; MacCullagh and Nelder 1989). The data were analyzed with GLIM (Payne 1987) using a binomial link function and a denominator equal to 1.

To test the adequacy of the model, we used graphical methods on Pearson residuals (Landwehr et al. 1984). The influence of outliers was tested by the deletion method (Pregibon 1981). The apparent error rate is calculated on the same sample as was used to describe the quality of the model, on the basis of how many cases were correctly classified (Hosmer and Lemeshow 1989). Also, a real error rate is calculated using 30 random validation samples, each comprising a third of the units. We applied the suggested model to 30 random learning samples (two other thirds of the units) and tested the classification results on each validation sample.

Interpretation of the model's coefficients is based on the odds ratio, which measures the association between a predictor covariate and the response outcome. This measurement is adjusted statistically for the other variables included in the model.

Results

Preliminary analysis

A link exists between the modality 'unknown' of the sex of the carrier parent and the outcome. It can be explained by a bias in the mode of ascertainment for some cases; for example, when there is no adverse outcome, research to determine which parent in the family is the carrier is less frequently performed. We thought it more advisable to exclude these units in order to avoid the bias. Previously, we confirmed that this exclusion does not lead to any change in the estimation of coefficients for the two modalities, namely maternal and paternal origin. Hence, to generate the logistic model we used only 5564 units, and the results for each covariate are shown in Table 3.

As paternal age was strongly correlated with maternal age ($r = 0.73$), and not significant in univariate analysis, we retained only maternal age for the analysis. The results with maternal age show a reverse influence; the younger the carrier parent, the higher the probability of unbalanced offspring. This is not what we expected (Stengel-Rutkowski et al. 1988). A closer examiniation of the data showed that this reverse influence was not constant throughout the data; no such influence was detected in the data from the literature, whereas the analysis of cases with systematic ascertainment mode showed the expected influence. In fact, the database contains two different selected groups: one issuing from young carrier parents, with a family history of affected offspring, seeking genetic counseling, and the other consisting of older carrier parents using genetic counseling for systematic identification of possible chromosomal abnormalities. As the unbiased sample of cases with systematic ascertainment mode was too small (1260 cases), we decided to exclude this variable age from the analysis (at least twice this number would be necessary to avoid cells with an insufficient number of cases).

The choice for the five different patterns of chromosomes involved (Fig. 1) was supported by a minimum size for each group, by common cytogenetic chromosome characteristics, such as length and centromeric position, and from results on the association between the meiotic segregation mode and the different chromosome subgroups. For the subsequent analysis, this variable with five modalities was transformed into five dummy variables by reference cell coding, the reference cell sending the most numerous one (chromosomes 1–8, 10–12, and 16–20).

Centric and translocated segments were correlated. Since centric segments were no longer significant when combined with the translocated segments, we retained only the translocated segments as quantitative chromosomic predictive variables.

The model

Five main effects and five second-order interaction terms were retained for the model. Among the ten available main effect variables, four were excluded after the univariate analysis (age and centric segments), one was excluded when interaction terms were added to the model (viability), and two were caused to remain in the model (origin and locus breakpoint). Among the 25 parameter coefficients, 5 are not significant.

$fit
linear predictor = v9+v3+v4+v5+v6+v4*v6+v9*v3++v9*v4+v9*v5+v9*v6

	Value	SE	*t*-value	
(Intercept)	0.45	0.17	2.6	
v9(2)	− 1.27	0.31	−4.1	Xme 9
v9(3)	− 1.10	0.28	−3.9	Xme 13, 14, 15
v9(4)	− 1.22	0.31	−4.0	Xme 21, 22
v9(5)	− 3.40	0.62	−5.5	t(11;22)
v3	− 0.10	0.01	−10.8	TS1R bands
v4	− 0.12	0.01	−8.9	TS2R bands
v5	− 0.43	0.14	−3.2	breakpoints: short arms
v6	− 0.26	0.14	−1.9[a]	maternal origin
v4*v6	0.04	0.01	3.2	
v9(2)*v3	0.08	0.01	5.4	
v9(3)*v3	0.08	0.01	5.6	
v9(4)*v3	0.09	0.02	5.7	
v9(5)*v3	0.05	0.06	0.8[a]	
v9(2)*v4	0.07	0.02	4.2	
v9(3)*v4	0.07	0.01	4.8	
v9(4)*v4	0.05	0.03	1.9[a]	
v9(5)*v4	0.12	0.05	2.5	
v9(2)*v5	0.77	0.29	2.7	
v9(3)*v5	0.07	0.31	0.2[a]	
v9(4)*v5	0.67	0.32	2.1	
v9(2)*v6	0.55	0.23	2.4	
v9(3)*v6	0.12	0.21	0.6[a]	
v9(4)*v6	0.86	0.21	4.0	
v9(5)*v6	1.99	0.51	3.9	

(Dispersion parameter for binomial family taken to be 1)
Null deviance: 5718 on 5563 *d.f.*
Residual deviance: 5337 on 5539 *d.f.*
Number of Fisher scoring iterations: 4
[a] Parameter coefficient not significant

Assessing the goodness of fit

In all, 94 outliers (1.7% of the data) were identified as they showed standardized Pearson residuals greater than 2.5 (range 2.51 to 9.78). Looking carefully at the 16 highest values (greater than 4), we notice that they belong to chromosome patterns 1, 2, or 5 (see Table 4). As the fitted value was low for all of these units (less than 0.06), a normal or unviable child would be expected. In fact, they all led to an unbalanced child at birth. Despite large translocated segments, in particular for the TS2R bands, cytogenetic characteristics of these units were very similar to those of units leading to a 'normal' or 'unviable' outcome. For each of these outliers, we tested the change in deviance and parameter coefficients when these observations were excluded. The change in deviance was less than 1%, and the result of the Wald test for the parameters remained identical. Therefore, we decided to keep these outlier units in the model.

In order to obtain a referential distribution for these Pearson residuals (Jennings 1986), we used simulated data: 45 binomial distributions with a mean equal to the fitted mean on the observed data. The aim was to estimate the distribution that these residuals would have if the fitted model were correct (Landwehr et al. 1984). The results shown in Fig. 2 did not suggest substantial deviation of the data from the model. The comparison between the Pearson residuals and the deviance residuals showed that the deviance residuals were lower than the Pearson residuals for the values above 2 (MacCullagh and Nelder 1989).

Interpretation

The influence of the variables measured by adjusted odds ratio show the classical strong association with the sex of the carrier parent and with the chromosomes involved [9, acrocentrics and translocation 11;22 (Jalbert et al. 1988; Daniel et al. 1989; Gardner and Sutherland 1989)].

The adjusted risk of unbalanced offspring is twice as high when the origin is maternal as when the origin is paternal [(1.3–3.0)] increasing to four times when chromosomes 21 and 22 are involved [(3.0–7.2)]. Concerning the well-known t(11;22)(q23;q11.2), the risk is eight times as high [(3.1–11.1)] when the origin is maternal. Whatever the sex of the carrier parent, the risk of unbalanced offspring decreases as translocated segments increase until a

Table 4. Outliers: Pearson standardized residuals (>4)

Pedigree number	CS1R bands	CS2R bands	TS1R bands	TS2R bands	Origin	v9	Chromosomes involved	Value of residual
205	35.6	2.75	8.6	14.0	Pat.	(5)	11–22	5.20
333	69.25	25.9	2.25	25.13	Pat.	(1)	2– 4	4.03
434	42.5	5.5	1.75	11.25	Pat.	(5)	11–22	4.30
440	42.5	11.1	1.75	5.6	Pat.	(5)	11–22	4.32
499 (2)	30.5	20.0	20.5	23.0	Mat.	(1)	5–10	6.42
566	33.0	12.25	18.0	26.75	Pat.	(1)	4–12	9.78
714	39.25	12.25	3.5	26.75	Pat.	(1)	8–12	4.73
793	10.6	3.5	23.6	26.25	Pat.	(2)	9–13	4.02
798	47.0	35.25	26.25	3.75	Mat.	(1)	1–12	4.05
805	35.6	2.75	8.6	14.0	Pat.	(5)	11–22	5.20
821	50.5	32.75	21.0	10.25	Pat.	(1)	2–10	4.31
963	10.6	0.88	23.6	27.9	Pat.	(2)	9–15	4.20
965 (2)	10.6	0.88	23.6	27.9	Pat.	(2)	9–15	4.20
1149	40.13	17.5	10.88	25.5	Mat.	(1)	5–10	4.64
	$m = 31.0$	$m = 12.8$	$m = 14.9$	$m = 18.3$				$m = 5.00$

All cases originate from the literature except case 205 which issued from genetic consultation and case 1149 which issued from prenatal diagnosis consultation. All these cases have shown an outcome MSMR

Literature: 333, Biederman B, Bowen P (1976) Partial trisomy 4q due to familial 2/4 translocation. Hum Genet 33:147–153; 434–440, Fraccaro M, Lindsten J, Ford CE, Iselius L (1980) The 11q;22q translocation: a European collaborative analysis of 43 cases. Hum Genet 56:21–51; 499, Gilgenkrantz S, Bugnon C, Bresson JL, Gouget A, Dulucq P (1980) Maternal translocation 5;10) identified only after amniocentesis in a second pregnancy with 10q+ fetus (partial trisomy 5q). Clin Genet 17:68; 566, Hirschhorn K, Lucas M, Wallace I (1973) Precise identification of various chromosomal abnormalities. Ann Hum Genet 36:375–379; 714, Nielsen J, Vetner M, Holm V, Askjaer SA, Reske-Nielsen E (1977) A newborn child with karyotype 47,XX,+der(12)(12pter→12q12::8q24→8qter,t(8;12)(q24;q12)pat. Hum Genet 35:357–362; 793, Schinzel A, Hayashi K, Schmid W (1975) Trisomy 9p due to paternal translocation, t(9;13)(q13;q12). Hum Genet 30:307–316; 798, van den Berghe H, van Eygen M, Fryns JP, Tanghe W, Verresen H (1973) Partial trisomy 1, karyotype 46,XY,12–,t(1q,12p)+. Hum Genet 18:225–230; 805, Schinzel A, Schmid W, Auf Der Maur P, Moser H, Degenhardt KH, Geisler M, Grubisic A (1981) Incomplete trisomy 22. I. Familial 11/22 translocation with 3:1 meiotic disjunction. Delineation of a common clinical picture and report of nine new cases from six families. Hum Genet 56:249–262; 821, Sills JA, Buckton KE, Raeburn JA (1976) Severe mental retardation in a body with partial trisomy 10q and partial monosomy 2q. J Med Genet 13:507–510; 963, Rodewald A, Stengel-Rutkowski S, Zankl M (1979) The dermatoglyphic pattern of the trisomy 9p syndrome. Clin Genet 16:405–417; 965, Zaremba J, Zdienicka E, Glogowska I, Abramowicz T, Taracha B (1974) Four cases of 9p trisomy resulting from a balanced familial translocation (9;15)(q13;q11). Clinical picture and cytogenetic findings. J Ment Defic Res 18:153–190

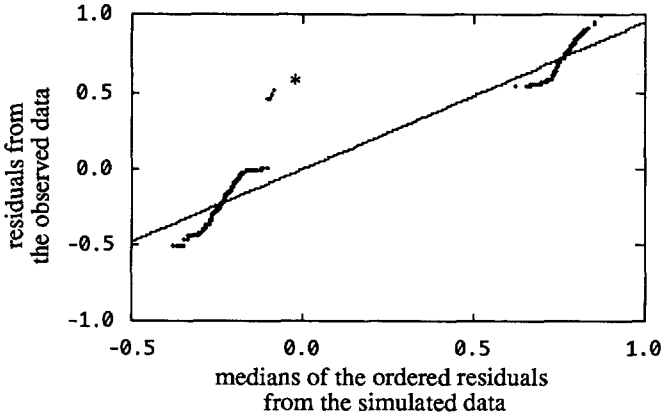

Fig. 2. Empirical probability plot for the model. (*) these points near the gap are due to the unequal number of 0 and 1 between observed and simulated data

length of 10. When the length of translocated segments exceeds 10, this risk still decreases if the origin is paternal, but remains steady if the origin is maternal.

For any chromosome involved, we note that the adjusted odds ratio increases as the length of the translocated segments increases, revealing the interaction. On the other hand, for a mean length of translocated segments of 10, the risk is two- to fivefold higher when the chromosomes involved are 9 or acrocentrics [(1.9–8.3)]. For these chromosomes, the risk increases even further if the breakpoints are both on the short arms.

To calculate the error rates, two cutpoints, 0.05 and 0.30, are defined from the estimated probability. This approach provides classification results as shown in Table 5. For the group at lower risk 97% of the units are correctly classified, while for the group of higher risk only 34% are correctly classified.

Discussion

The study of the outliers raises the question of how exhaustive the data collection has been. The outcome of the pregnancy is influenced by many different factors, including intercurrent maternal pathology, which can lead to miscarriage whether the unbalanced fetus would have been viable in other circumstances or not. Many spontaneous abortions are not in our database and the reasons for miscarriages are often unknown. This bias is unavoidable, as it is not possible to collect information on all abortions (Alberman 1973; Boué et al. 1975).

It is noteworthy that by using this model for the risk of unbalanced offspring we have demonstrated the important role of certain veriables, already identified by other investigators (Stene 1989), as having some influence on the different unbalanced meiotic segregation modes (2:2 or 3:1). This probably represents, in clinical terms, the fact that the potential gametic viability influences the different unbalanced meiotic segregation modes.

The strong influence of maternal origin observed in the case of t(11;22) can be explained by the recognized influence of maternal origin on the 3:1 tertiary segregation mode, which is the preferential unbalanced mode for these translocations (Daniel et al. 1988; Jalbert et al. 1988). In Table 4, we notice that in t(11;22) all the outliers are of paternal origin.

Comparison of the results of the logistic model with the descriptive analysis clearly shows that information obtained by logistic regression is much more detailed, allowing adjusted odds ratio and expression of interaction phenomena. Usually for unbalanced segregation modes, the largest translocated segments lead more frequently to nondisjunction (Thomas et al. 1987). Here we note that the larger the translocated segments, the less frequently viable unbalance occurs and multivariate analysis provides possible eyplanations. The risk of viable unbalanced offspring decreases as the length of the translocated segments increases; this is probably because, in the case of nondisjunction with large translocated segments, there is preferential formation of an inapparent unviable fetus with high chromosomal unbalance (mark of natural selection). At the same time, depending on other variables, an increase of TSR bands will increase the risk of unbalanced offspring (reflecting the expression of interactions between TSR bands and the other variables). For exam-

Table 5. Classification tables using the model

a Apparent error rate (data used to built the model)

Outcome response observed	Cutpoints of the estimated probability fitted by the model			Total
	Lower risk $P < 0.05$	Moderate risk $P = 0.05$ to 0.30	Higher risk $P > 0.30$	
0 (normal or unviable)	**422** (98%)	3354	620	4396
1 (MSMR)	10	817	**341** (35%)	1168
Total	432	4171	961	5564

b Real error rate (mean on 30 randomised validation samples)

Lower risk ($P < 0.05$)	97.0% Correctly classified
Higher risk ($P > 0.30$)	34.1% Correctly classified

P = estimated probability fitted by the model
As we were more interested by the extreme risks, we define two cutpoints with a group of moderate risk for whom prediction is not available

ple, the risk of unbalanced offspring due to chromosome 21 or 22 is fourtimes as high when the length of TSR bands is 15 than when it is 5. This illustrates the fact that an increase of TS is not necessarily linked with an increase of quantity of the chromosomal imbalance. Also, the influence of this increase may differ according to the sex of the carrier parent.

The choice of logistic link function was satisfactory (Hastie and Tibshirani 1990). Although we emphasize, with inferential methods, some specific effects on the risk of viable unbalanced offspring, the applicability of the model remains limited since classification results are correct only for a small number of cases, e.g., those with very low risk of unbalanced offspring. This result is, nevertheless, interesting in the sense that it answers one of the questions posed at the outset, namely whether the model can detect translocations giving a very low risk of unbalanced offspring. In this case, the choice of prenatal diagnosis can be best discussed by considering both the risk estimation for unbalanced offspring and the risk of induced abortion following prenatal diagnosis (e.g. choriocentesis). The bias due to the fact that non-karyotyped units (more frequently balanced individuals) are not included in our database can lead to an overestimation of the risk of viable unbalanced offspring (as do the use of Jalbert's criteria for viability); however, this will not lead to adverse conclusions for the lower risk group.

As in other structural chromosomal abnormalities, we suspect that the influence of age really does exist, and we realize that for this to be confirmed we should need a larger unbiased database, and an increase in the number of cases with both age details and systematic ascertainment mode.

Acknowledgements. This work was supported by research grant 91-69-001 from Institut National de la Santé et de la Recherche Médicale, Caisse Nationale de l'Assurance Maladie des Travailleurs Salariés (INSERM-CNAMTS).

References

Alberman E (1973) Epidemiology of spontaneous abortions and their chromosome constitution. In: Boué A, Thibault C (eds) Chromosomal errors in relation to reproductive failure. INSERM, Paris, pp 305–316

Boué J, Boué A, Lazar P (1975) Retrospective and prospective epidemiological studies of 1500 karyotyped spontaneous human abortions. Teratology 12:11–26

Boué A, Gallano P (1984) A collaborative study of the segregation of inherited chromosome structural rearrangements in 1356 prenatal diagnoses. Prenat Diagn 4:45–67

Cans C, Cohen O, Mermet MA, Demongeot J, Jalbert P (1993) Human reciprocal translocations: is the unbalance mode at birth predictable? Hum Genet 91:228–232

Cohen O, Cans C, Mermet MA, Demongeot J, Jalbert P (1992) Human reciprocal translocations: a new computer system for genetic counseling. Ann Génét 35:193–201

Crandall BF, Lebherz TB, Rubinstein L, Robertson RD, Sample WF, Sarti D, Howard J (1980) Chromosomes findings in 2500 second trimester amniocentesis. Am J Med Genet 5:345–356

Daniel A (1979) Structural differences in reciprocal translocations. Potential for a model of risk in reciprocal translocations. Hum Genet 51:171–182

Daniel A, Hook EB, Wulf G (1988) Collaborative USA data on prenatal diagnosis for parental carriers of chromosome rearrangements: risk of unbalanced progeny. In: Daniel A (ed) The cytogenetics of mammalian autosomal rearrangements, Liss, New York, pp 73–162

Davis JR, Rogers BB, Hagaman RM, Thies CA, Veomett IC (1985) Balanced reciprocal translocations: risk factors for aneuploid segregant viability. Clin Genet 27:1–19

De Arce MA, Grace PM, McManus S (1986) A computer model for the study of segregation in reciprocal translocation carriers: application to 20 new cases. Am J Hum Genet 24:519–525

Gardner RJM, Sutherland GS (1989) The simple reciprocal translocation. In: Motulsky AG, Bobrow M, Harper PS, Scriver C (eds) Oxford monographs on medical genetics no. 17. Chromosome abnormalities and genetic counseling. Oxfort University Press, Oxford, pp 29–46

Greenland S (1989) Modeling and variable selection in epidemiologic analysis. Am J Public Health 79:340–349

Hastie T, Tibshirani R (1990) Generalized additive models. Chapman & Hall, London

Hosmer DW, Lemeshow S (1989) Applied logistic regression. Wiley, New York

ISCN (1981) An international system for human cytogenetic nomenclature. High resolution banding. Cytogenet Cell Genet 11:88–94

Jalbert P, Jalbert H, Sele B (1988) Types of imbalances in human reciprocal translocations: risks at birth. In: Daniel A (ed) The cytogenetics of mammalian autosomal rearrangements. Liss, New York, pp 267–291

Jalbert P, Cohen O, Cans C, Mermet MA, Demongeot J (1992) Unbalances of reciprocal translocations (rcp): maximum thresholds of viability. Am J Hum Genet 51:A1240

Jennings DE (1986) Outliers and residual distributions in logistic regression. J Am Statist Assoc 81:987–990

Landwehr J, Pregibon D, Shoemaker AC (1984) Graphical methods for assessing logistic regression models. J Am Statist Assoc 79:61–71

McCullagh P, Nelder JA (1989) Generalized linear models. Chapman & Hall, London

Payne CD (ed) (1987) GLIM system release 3.77 manual. Oxford, UK

Petrosky DL, Bergaonkar DS (1984) Segregation analysis in reciprocal translocation carriers. Am J Med Genet 19:137–159

Pregibon D (1981) Logistic regression diagnostics. Ann Statistics 9:705–724

Stene J (1989) Comments on estimations of risks to translocation carriers. Am J Med Genet 33:54–57

Stengel Rutkowski S, Stene J, Gallano P (1988) Risk estimates in balanced reciprocal translocations. (Monographie des annales de génétique), Expansion Scientifique Française, Paris