

Estimación del riesgo de descendencia con aberraciones cromosómicas desbalanceadas en progenitores portadores de translocaciones

Joenith Aguilar¹, Jorge Bacallao-Guerra², Jorge Bacallao-Gallestey³, Estela Morales⁴

¹Departamento de Citogenética, Centro Nacional de Genética Médica, Ave. 31, esq. 146, Cubanacán, Playa, CP 11600, La Habana, Cuba

²Departamento de Asistencia Médica, Centro Nacional de Genética Médica, Calle E, No. 309, esq. 15, Plaza de la Revolución, CP 10600 La Habana, Cuba

³Departamento de Matemática, Instituto de Cibernética, Matemática y Física, Ministerio de Ciencia, Tecnología y Medio Ambiente, CITMA Calle Tulipán, esq. Panorama, Plaza de la Revolución, CP 10600, La Habana, Cuba

⁴Centro de Investigaciones y Referencia de Aterosclerosis de La Habana Ave. 31, esq. 146, Cubanacán, Playa, CP 11600, La Habana, Cuba
E-mail: bacallao@icmaf.cu

INVESTIGACIÓN

RESUMEN

Las translocaciones recíprocas y robertsonianas son aberraciones cromosómicas estructurales que durante la meiosis pueden originar gametos cromosómicamente desbalanceados. Estos desequilibrios pueden generar una progenie con múltiples malformaciones. El objetivo de este trabajo es proponer una metodología para estimar el riesgo de descendencia con aberraciones cromosómicas desbalanceadas compatibles con la vida, en progenitores portadores de alguna translocación recíproca o robertsoniana. La metodología combina los resultados de los métodos de regresión logística, regresión múltiple y regresión de Poisson. Por sus características, es teóricamente superior a otras variantes descritas en la literatura revisada. El riesgo se estima a partir de una base de datos que contiene información de 41 estudios de familias portadoras de translocaciones, de tres provincias de Cuba. Los resultados con los tres métodos aplicados son coherentes en la selección de las variables predictoras (presencia de cromosomas 9, cromosoma 21 y existencia de puntos de ruptura en los brazos cortos de los cromosomas involucrados) y en las estimaciones del riesgo. El 80% de las familias se clasificaron por estos métodos.

Palabras clave: estimación del riesgo, translocaciones, regresión de Poisson

Biotecnología Aplicada 2011;28:151-155

ABSTRACT

Estimation of risk of progeny with unbalanced chromosome aberrations in parents carriers of translocations. The reciprocal and Robertsonian translocations are structural chromosomal aberrations that during meiosis can produce unbalanced gametes. In some cases, these imbalances lead to an offspring with multiple malformations. The purpose of this study was to propose and evaluate a methodology to estimate the risk of live offsprings with unbalanced chromosome aberrations (LOUCA) when parents carry some reciprocal or Robertsonian translocation. The methodology was based on the replicability of several widely known regression methods: multiple linear, logistic and Poisson regression. Risk was estimated on a database which contains information on 41 familiar records of translocations in three Cuban provinces. The results yielded by the three models were quite consistent as regards variable selection (presence of chromosome 9, chromosome 21 and the existence of breaking points in the short arms of the chromosomes involved) and risk estimation. There was an 80% of coincidence in the classification yielded by the three methods.

Keywords: Risk estimation, translocations, Poisson regression

Introducción

Durante la meiosis, las translocaciones recíprocas y robertsonianas pueden originar gametos cromosómicamente desbalanceados, que pueden suscitar una progenie con múltiples malformaciones [1, 2]. Para las familias con translocaciones, es muy importante conocer la probabilidad o el riesgo de aparición de las consecuencias de este fenómeno en su descendencia, como los abortos espontáneos, las muertes fetales y con aberraciones cromosómicas desbalanceadas compatibles con la vida (DacDCV) [3].

Este trabajo propone una metodología basada en modelos lineales generalizados, para la estimación del riesgo de DacDCV en progenitores portadores de translocaciones. Se sustenta en bases teóricas y puede emplearse en la práctica médica durante el proceso

de asesoramiento genético. Desde finales de los años 70 han aparecido numerosas investigaciones sobre el tema: Daniel en 1979 [4], Stengel-Rutkowski y cols. en 1988 [5], Cans y cols. en el 1993 [6], y Cohen y cols. en 1992 [7] y 1994 [8].

En 1993, Cans y colaboradores propusieron el uso de la regresión logística como método de estimación del riesgo, y dos años más tarde se recomendó también el uso de modelos aditivos [9].

Algunas consideraciones teóricas y prácticas se oponen a la regresión logística como herramienta aislada para la estimación de riesgos. Cuando se usa con propósitos explicativos (elección de variables relevantes para predecir el riesgo), es muy sensible a los efectos de la multicolinealidad, como todos los mode-

1. Morel F, Douet-Guilbert N, Le Bris MJ, Herry A, Amice V, Amice J, et al. Meiotic segregation of translocations during male gametogenesis. *Int J Androl.* 2004;27(4):200-12.

2. Young ID. Risk calculation in genetic counseling. 3rd ed. Oxford: Oxford University Press; 2007.

3. Stasiewicz-Jarocka B, Haus O, Van Assche E, Kostyk E, Constantinou M, Rybalko A, et al. Genetic counseling in carriers of reciprocal chromosomal translocations involving long arm of chromosome 16. *Clin Genet.* 2004;66(3):189-207.

los lineales. Supone además la relación lineal entre la variable dependiente y los predictores.

Cuando los datos provienen de estudios familiares, hay un efecto ‘familia’, debido a que los sujetos de una misma familia no solo están más relacionados entre sí que con el resto, sino que comparten exactamente los mismos valores de las variables que describe la translocación que porta la familia. Esto contradice los supuestos básicos de la regresión logística, que necesita independencia entre los sujetos de la base.

El procedimiento que se propone para eludir estas dificultades consiste en la clasificación en grupos de riesgo, a partir de los resultados con la regresión múltiple y la regresión de Poisson, haciendo modificaciones en la estructura de los datos provenientes de estudios familiares de translocaciones. Es un compendio de técnicas estadísticas de distinta naturaleza que se complementan y permiten hacer valoraciones objetivas sobre los riesgos en sí, con vistas a su posterior utilización en el asesoramiento genético.

Materiales y métodos

Base de datos

La base de datos original recoge la información de 200 sujetos pertenecientes a 41 estudios de familias portadoras de translocaciones recíprocas o robertsonianas (en cada estudio de familia se tuvieron en cuenta los individuos con relaciones filiales con respecto al propósito de cada estudio, incluyendo este último). De estos dos tipos de aberraciones estructurales, 26 translocaciones presentan diferentes puntos de ruptura (Figura 1). Solo se incluyeron translocaciones entre cromosomas autosómicos, cariotipadas y heredadas. Los datos se obtuvieron de la colaboración de los laboratorios de citogenética de La Habana, La Habana y Pinar del Río. Los puntos de ruptura se homogenizaron con una resolución de 400 bandas, según el Sistema Internacional de Nomenclatura (ISCN 2005) [10]. La longitud de los segmentos céntricos y translo-

cados se midió desde el punto de ruptura hasta la zona terminal de los brazos corto y largo, con una exactitud de 0.5 mm (bandas G).

Se tuvo en cuenta la variable *condición de la descendencia* con seis categorías: no portador, portador balanceado, aborto espontáneo, muerte fetal, muerte neonatal y DacDCV. Esta última variable se reestructuró como binaria, y en una sola se unieron todas las categorías diferentes a DacDCV. Se crearon cinco grupos cromosómicos según los cromosomas involucrados en cada translocación. Según la ubicación de los puntos de ruptura se definieron tres categorías (pp, pq y qq). Además se tuvo en cuenta el origen parental de la translocación. Teniendo en cuenta estudios recientes que plantean que no se ha demostrado la relación entre la edad parental y el riesgo de DacDCV, y el alto porcentaje de observaciones faltantes, en la investigación no se incluyó la variable *edad del progenitor portador* [11]. La variable *viabilidad de los gametos*, no se incluyó en el análisis debido a los pocos casos con DacDCV. Los predictores tentativos en los modelos se resumen en la tabla 1. Las variables nominales *grupos cromosómicos* y *ubicación de los puntos de ruptura* con cinco y tres categorías respectivamente, se transformaron en variables de simulación que toman los valores 0 o 1 para incluirlas en los modelos de regresión.

Análisis estadístico

La estrategia metodológica consistió en la aplicación de tres modelos de regresión y en el análisis de la replicabilidad de los resultados. A partir del reconocimiento de las limitaciones de cada procedimiento, esta estrategia de triangulación, más allá del análisis estadístico, confirmó mayor veracidad a los resultados, cuando se reproducen a partir de supuestos diferentes. En la estimación del riesgo se tuvieron en cuenta los puntos de corte usados en el asesoramiento genético (riesgo bajo < 5%, riesgo moderado entre 5 y 15%, riesgo alto > 15%).

4. Daniel A. Structural differences in reciprocal translocations. Potential for a model of risk in Rcp. Hum Genet. 1979; 51(2):171-82.
5. Stene J, Stengel-Rutkowski S. Genetic risks of familial reciprocal and Robertsonian translocation carriers. In: Daniel A, editor. The Cytogenetics of Mammalian Autosomal Rearrangements. New York: Alan R Liss; 1988. p. 54-61.
6. Cans C, Cohen O, Lavergne C, Mermet MA, Demongeot J, Jalbert P. Logistic regression model to estimate the risk of unbalanced offspring in reciprocal translocations. Hum Genet. 1993;92(6):598-604.
7. Cohen O, Simonet M, Cans C, Mermet MA, Demongeot J, Amblard F, et al. Human reciprocal translocations: a new computer system for genetic counseling. Ann Genet. 1992;35(4):193-201.
8. Cohen O, Cans C, Mermet MA, Demongeot J, Jalbert P. Viability thresholds for partial trisomies and monosomies. A study of 1,159 viable unbalanced reciprocal translocations. Hum Genet. 1994; 93(2):188-94.
9. Cans C, Lavergne C. De la régression logistique vers un modele additif généralisé: un exemple d'application. Rev Stat Appl. 1995;43(2):77-90.
10. Shaffer LG, Tommerup N. ISCN: An International System for Human Cytogenetic Nomenclature. Memphis: Karger; 2005.

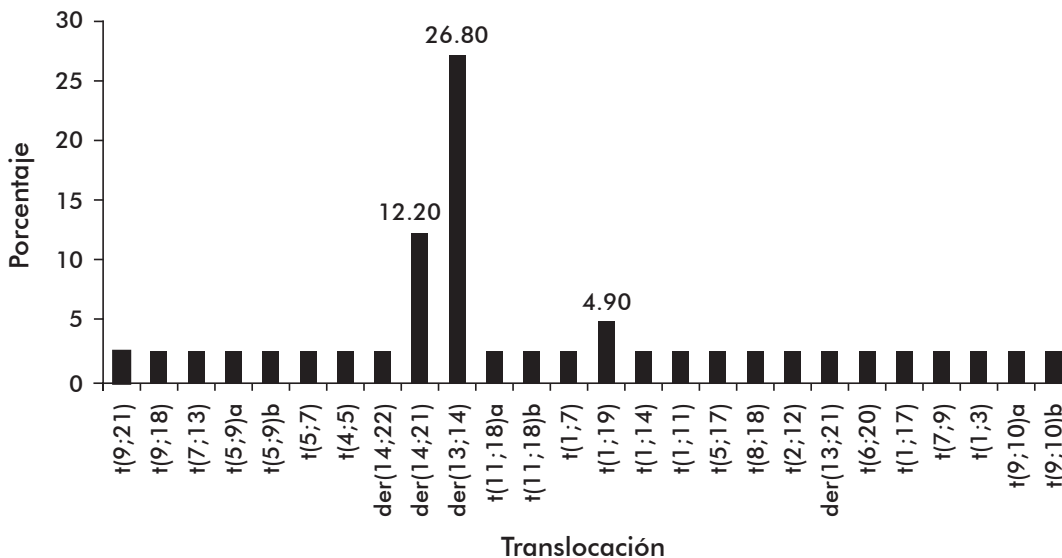


Figura 1. Distribución de las familias afectadas según el tipo de translocación (porcentaje). La translocación der(13;14) afecta 11 familias; der(14;21) afecta cinco y t(1;19) afecta dos. El resto de las translocaciones (23/26), cada una afecta a una familia (2.4%). a y b: translocaciones de iguales cromosomas y puntos de rupturas diferentes.

Se aplicaron los modelos de regresión logística y regresión lineal múltiple mediante el paquete estadístico SPSS (Statistical Package for the Social Sciences, versión 15.0), y la regresión de Poisson, utilizando el paquete estadístico STATA.

Regresión logística

Con el objetivo de comparar los resultados, se analizaron los datos a partir de un modelo de regresión logística (se utilizó la base de datos con 200 sujetos). Aunque se observaron cautelosamente las dificultades que podría presentar la aplicación de este método, se muestran los resultados.

La regresión logística [12-15] forma parte de los modelos lineales generalizados. Se utiliza para modelar una respuesta categórica a partir de predictores que indistintamente pueden ser continuos o discretos ordinales (en la mayoría de los casos, esta respuesta es binaria). En el modelo, p es la probabilidad de éxito y las X_i son los predictores o variables independientes:

$$\log \text{it} (p) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k$$

Donde:

$$\log \text{it} (p) = \log (p / (1 - p))$$

La regresión logística binaria es muy utilizada en ambientes de investigación biomédica y epidemiológica, pues la forma de la función logística es ideal para la modelación de riesgos y respuestas dicotómicas, como el padecimiento de determinada enfermedad o no [12].

Regresión lineal múltiple

El modelo tiene la forma:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k$$

Donde:

Y es la variable dependiente y X_i son las variables independientes o predictoras. Las letras griegas beta indicadas representan los parámetros del modelo. Se utilizó este modelo en una base de datos obtenida a partir de la modificación de la base de datos original para eludir el efecto 'familia' de esta última (que afecta al modelo de regresión logística anterior).

En la nueva base de datos, la unidad de análisis es la familia (las familias son independientes entre sí). La variable dependiente (definida en el intervalo [0, 1]) indica la proporción de individuos afectados en cada una. De esta manera, riesgo individual y riesgo familiar son la misma cosa.

Esta modificación se justifica porque de acuerdo con las características de los datos, el riesgo se define por factores intrínsecos a la translocación que afecta a cada familia y no a los sujetos individualmente. Las variables contempladas en este estudio y que influyen en el riesgo de aparición de DacDCV solo caracterizan a la translocación portada y no a otros fenómenos genéticos precigóticos y psicigóticos (igualmente influyentes en el riesgo) propios de cada individuo portador. Los resultados de este trabajo son coherentes con estos planteamientos.

Regresión de Poisson

La regresión de Poisson [13-15] se utiliza para la modelación de variables del tipo conteo, y en espe-

Tabla 1. Predictores tentativos

Variabes	Descripción de la variable	Símbolo	Nº
Segmentos cromosómicos céntricos y translocados	Segmento céntrico del primer cromosoma involucrado	SC1G	1.2-13.7
	Segmento céntrico del segundo cromosoma involucrado	SC2G	1.2-8.3
	Segmento translocado del primer cromosoma involucrado	ST1G	0.3-6.1
	Segmento translocado del segundo cromosoma involucrado	ST2G	0.4-6.7
Puntos de ruptura	Brazo qq	BRA 3	n = 119
	Brazo pq	BRA 2	n = 40
	Brazo pp	BRA 1	n = 41
Cromosomas involucrados	Gc1 [1-8, 10-12, 16-20]	TRAS 1	n = 53
	Gc2 [9]	TRAS 2	n = 42
	Gc3 [der(13;14)]	TRAS 3	n = 56
	Gc4 [13-15]	TRAS 4	n = 14
	Gc5 [21-22]	TRAS 5	n = 35
Origen parental	Origen maternal	Sexo F	n = 138
	Origen paterna	Sexo M	n = 62

Los valores de N para los sectores cromosómicos céntricos y translocados se expresan en % en base a la frecuencia esperada para una N de 200 sujetos.

cial para la modelación de riesgos de ocurrencia de eventos muy poco frecuentes. La naturaleza de este problema se adapta a la aplicación de la regresión de Poisson, toda vez que la cantidad de casos en una familia es una variable del tipo conteo, y a partir de ello, se puede asignar un riesgo a cada familia.

Este tipo de modelo es el más adecuado desde el punto de vista teórico (los resultados respaldan esa afirmación), por lo que al analizar los resultados de clasificación en grupos de riesgo, se da mayor importancia a los resultados a partir de él.

Una variable aleatoria tiene distribución de Poisson si su función de probabilidad se puede escribir como:

$$P(X = x) = \frac{e^{-\lambda} \lambda^x}{x!}$$

La formulación básica de la regresión de Poisson consiste en escribir la media de la variable conteo, como la exponencial de una función lineal de los predictores:

$$\lambda = \exp(\beta_0 + \beta_1 X_1 + \dots + \beta_n X_n)$$

En la base transformada se aplicó este modelo debido a la estructura resultante. Se consideró la variable "cantidad de casos con DacDCV en la familia", y se consideró el riesgo de cada individuo como la cantidad de casos en su familia dividida por los integrantes de la familia (variable *tamaño de la familia*, TMFA). Ante la duda de si la influencia del tamaño de la familia pudiera ser trascendente, se consideraron modelos adicionales que incluían esa variable y se compararon con modelos que no la contenían. El tamaño de la familia no fue significativo.

Resultados

Regresión logística

Se dieron tres pasos hasta obtener el modelo 3, que predice el riesgo de DacDCV, a partir de las variables TRAS2 (Gc2), TRAS5 (Gc5) y BRA1 (pp). Las tres variables son relevantes, y sus niveles de significación de 0.003, 0 y 0.013, respectivamente (Tabla 2). El parámetro R^2 de Nagelkerke es la medida de variabilidad explicada por cada modelo. Este parámetro se comportó para cada uno de los modelos de la siguiente forma: Modelo 1, 0.065; Modelo 2, 0.166; Modelo

11. McKinlay Gardner RJ, Sutherland GR, editors. Chromosome abnormalities and genetic counseling. 3rd ed. Oxford: Oxford University Press; 2004.

12. Kleinbaum DG, Klein M. Logistic regression. A self learning text. New York: Springer-Verlag; 2002.

13. Agresti A. Categorical Data Analysis. 1st ed. New York: John Wiley & Sons; 1990.

14. McCullagh P, Nelder JA. Generalized Linear Models. 2nd ed. London: Chapman and Hall/CRC; 1989.

15. Christensen R. Loglinear Models. 1st ed. New York: Springer-Verlag; 1990.

3, 0.228. Hay una diferencia entre el modelo 3 y los dos anteriores. Luego de este estudio, los sujetos quedan clasificados de acuerdo con el riesgo de DacDCV, estimado a partir de la regresión logística (Anexo).

Regresión lineal múltiple

Seguidamente, se ajustó un modelo de regresión múltiple que tiene la variable ‘frecuencia’ como variable dependiente. Esta se aplicó hacia adelante (*forward*) que incluye variables en el modelo bajo el criterio de maximizar la ganancia en el valor de R², y que contiene un criterio adicional para detener la incorporación de variables cuando estas no originan un incremento sustancial en la R².

El modelo que mejor elige el método *forward* es aquel que tiene las variables TRAS5 (Gc5) y BRA1 (pp) como predictores (R² ajustada = 0.27, Tabla 3). Estas son variables con efectos significativos (0.004 y 0.002, respectivamente; Tabla 4) que también están en el modelo de regresión logística analizado.

La variable que pudiera estar en el modelo actual sería TRAS2; exactamente la misma que completa la terna en el anterior modelo de regresión logística. Después de este análisis, las familias quedaron clasificadas de acuerdo con el riesgo de procrear DacDCV. En este modelo, ninguna familia quedó en el grupo de riesgo moderado (Anexo).

Regresión de Poisson

Como en el modelo anterior, la unidad estadística fue la familia, y la variable que se analizó fue el número de casos con DacDCV. Se incluyó la nueva variable TMFA, bajo el supuesto de que los casos afectados dependen del tamaño de la familia, lo que garantiza no obtener incompatibilidades (por ejemplo, mayor número de casos afectados que casos susceptibles). El modelo se ajustó con la variante de regresión robusta del paquete estadístico informático comercial STATA, que garantiza estimaciones libres de los efectos de apalancamiento por observaciones atípicas. Las variables relevantes fueron TRAS5 (Gc5), TRAS2 (Gc2) y BRA1 (pp) (Tabla 5). Como el número de eventos (variable a predecir) tendría que depender de la variable TMFA, se incluyó como covariante en el modelo.

Todo se mantuvo igual, y aparentemente, la variable a predecir depende mucho más de los factores genéticos que del TMFA (valor P de la variable TMFA en la Tabla 6).

Entre las predicciones que ofrece el modelo de regresión de Poisson, aparece nuevamente el riesgo moderado (ver Anexo).

Discusión

Los tres modelos ofrecen resultados semejantes en cuanto a las variables más relevantes y a la adecuada clasificación en grupos de riesgo, lo mismo cuando la unidad estadística es el sujeto que cuando es un estudio familiar.

En los tres modelos hay variables que indican la presencia del carácter, y aumentan el riesgo considerablemente. Estas son la presencia de puntos de ruptura en los brazos pp y los cromosomas involucrados 21 y 22. Otra variable que en un caso no alcanzó la significación estadística fue la presencia del cromosoma 9, incluida en los modelos de regresión logística y de

Tabla 2. Predictores en los tres modelos según la regresión logística

Modelos	Variables incluidas	Coefficientes	Significación
Modelo 1	TRAS5	1.355	0.011
Modelo 2	TRAS2	2.079	0.004
	TRAS5	2.303	0.001
Modelo 3	TRAS2	2.184	0.003
	TRAS5	2.916	0
	BRA1	1.790	0.013

Poisson. Ello demuestra la importancia de la presencia de estas variables en la comprensión del fenómeno genético que se estudia. Los resultados son similares a los de Cans y cols. en 1993 [6].

En ese mismo trabajo se apunta, además, que en estos cromosomas, el riesgo es aún mayor si los puntos de ruptura se localizan en ambos brazos cortos. A pesar de utilizar datos provenientes de diferentes poblaciones y modelos, ambas investigaciones llegaron a resultados similares. Ello prueba la validez del enfoque y la capacidad predictiva de las variables elegidas. Numerosas investigaciones reportan al cromosoma 9 como muy frecuente en la DacDCV, a partir de la segregación de translocaciones que lo incluyan [16, 17].

16. Sankoff D, Deneault M, Turbis P, Allen C. Chromosomal distributions of break-points in cancer, infertility, and evolution. *Theor Popul Biol.* 2002;61(4):497-501.

17. Harper PS. *Practical genetic counselling.* 6th ed. Verlag: Hodder Arnold; 2004.

Tabla 3. Ajuste de los modelos de regresión múltiple

Modelos	R ²	R ² ajustada	Error estándar de la estimación
1 ^a	0.137	0.115	0.21473
2 ^b	0.307	0.270	0.19494

^aPredictores en el modelo: BRA1.

^bPredictores en el modelo: BRA1, TRAS5.

Tabla 4. Coeficientes de los modelos de regresión múltiple

Modelos	Coeficientes no estandarizados		Coeficientes estandarizados Beta	Significación
	B	Error estándar		
1 BRA1	0.236	0.095	0.370	0.017
2 BRA1	0.286	0.088	0.449	0.002
TRAS5	0.251	0.082	0.420	0.004

Tabla 5. Salida de la regresión de Poisson sin el predictor ‘tamaño de la familia’

Regresión de Poisson		Número de observaciones = 41 LR chi ² (3) = 115.58 Prob > chi ² = 0			
Log verosimilitud = -79.619332		Pseudo R ² = 0.4206			
Número de eventos	Coefficiente	Error estándar	z	P> z	[95% Conf., interval]
TRAS5	10.75	1.06	10.16	0	8.68 12.83
TRAS2	6.62	0.89	7.43	0	4.87 8.36
BRA1	6.23	0.82	7.63	0	4.63 7.83
_cons	-18.15	0.99	-18.36	0	-20.09 -16.21
TMFA	No se incluyó en el modelo				

Tabla 6. Salida de la regresión de Poisson con el predictor “tamaño de la familia”

Regresión de Poisson		Número de observaciones = 41 LR chi ² (4) = 20.50 Prob > chi ² = 0.0004			
Log verosimilitud = -25.59		Pseudo R ² = 0.29			
Número de eventos	Coefficiente	Error estándar	z	P> z	[95% Conf., interval]
TRAS5	2.53	0.74	3.41	0.001	1.08 3.99
TRAS2	1.96	0.70	2.82	0.005	0.60 3.33
BRA1	1.28	0.66	1.94	0.053	-0.01 2.57
_cons	-2.83	0.80	-3.55	0.000	-4.39 -1.27
TMFA	0.06	0.07	0.81	0.418	-0.08 0.19

El cromosoma 22 también está en el Gc5, por lo que también aparece en todos los modelos como cromosoma que aumenta el valor de riesgo. En la base de datos empleada, no se encontró ninguna DacDCV que presentara este cromosoma; sin embargo, en el trabajo de Cans y cols. 1993 [6] sí, por la alta prevalencia de la t(11;22) muy relacionada con la presencia de este tipo de descendencia.

El 80% de las familias se clasificaron en grupos de riesgo de la misma manera por los tres modelos (Anexo). Teniendo en cuenta este resultado, y la coincidencia en la selección de las variables, fue evidente la relevancia de las variables seleccionadas como predictoras. De forma general, la regresión de Poisson permitió mayor precisión; es decir, un valor teórico de riesgo más cercano al 'observado'. Solo en la familia 34, este modelo mostró un valor de riesgo muy distante de la realidad (Anexo): en esa familia hay tres personas, y dos de ellas son DacDCV. En este modelo nuevamente se apreció la presencia del Gc2 (cromosoma 9), lo cual tal vez redujo el valor teórico de riesgo de DacDCV, debido a que de los tres predictores seleccionados en la regresión de Poisson (Gc5, Gc2 y pp), solo presenta puntos de ruptura en los brazos cortos de ambos cromosomas de la translocación (pp).

Del análisis de la alta precisión de la regresión de Poisson, es preciso destacar que 1) en la familia 40, es el único modelo que permitió clasificarla de la mejor manera, según los datos obtenidos. Esta familia solo presentó Gc2 (variable con valor = 1) de los tres predictores seleccionados en este modelo, lo cual destaca la importancia de que en una translocación se encuentre el cromosoma 9 y ello se relacione con la probabilidad de tener DacDCV [18]; y 2) los estudios de familias diferentes con la misma translocación, presentaron diferentes valores de riesgo (esto no había sucedido hasta ahora), aunque quedan clasificadas dentro del mismo grupo de riesgo. Ello se explica cuando se tiene en cuenta la diferencia, desde el punto de vista de las expresiones matemáticas de los modelos, que en el caso de la regresión de Poisson implica utilizar información propia de cada familia, lo que provoca diferencias entre familias diferentes pero con la misma translocación.

La clasificación de las familias por grupos de riesgo en la regresión de Poisson y en la regresión logística coincide en 95%; mientras que en la regresión múltiple coincide en 80% con la regresión de Poisson y la regresión logística. El alto porcentaje de coincidencia entre la regresión de Poisson y la regresión logística puede deberse a que, independientemente de las características de cada modelo, ambos presentan las mismas variables (Gc5, Gc2 y pp) como predictores. Con respecto a la regresión múltiple se observó que toda translocación que presente puntos de ruptura en los brazos cortos (como la familia 29) es de alto riesgo, independientemente de si presenta DacDCV o no.

En este modelo, los predictores seleccionados son la presencia de puntos de ruptura en ambos brazos cortos (pp, BRA1 = 1) y Gc5 (cromosomas 21 y 22, TRAS5 = 1). Además, en la base de datos no hay ninguna translocación con ambos predictores a la vez. En

este modelo no se presentan muchas combinaciones para que una familia sea clasificada de alto riesgo, de manera que si cumple con al menos uno de ellos es suficiente para que se clasifique de alto riesgo, independientemente de que en lo 'observado' haya historia reproductiva con DacDCV. En este modelo, el parámetro que acompaña a la variable pp es mayor que el que acompaña a la variable Gc5. Por ello, cuando una familia presenta una translocación con ambos puntos de ruptura en brazos cortos, tiene un mayor riesgo esperado de acuerdo con este modelo.

Las familias 29 y 40 no se clasificaron de igual manera por los tres modelos (Anexo). La familia 40 debería de estar clasificada como de alto riesgo; sin embargo, solo el modelo de la regresión de Poisson lo logra, por las razones explicadas. Según lo 'observado', la familia 29 debería de estar clasificada como de riesgo moderado; sin embargo, solo la regresión logística así lo establece. En esta última familia, la regresión de Poisson tuvo un valor de riesgo cercano al rango de riesgo moderado. La familia 7 estuvo mal clasificada por los tres modelos (alto riesgo). La razón fue que la translocación que caracteriza a esta familia contiene el cromosoma 22, el cual forma parte del Gc5 (variable seleccionada con mucho peso por los tres modelos), pero esa familia no posee historia reproductiva con DacDCV. Además, en toda la base de datos es la única familia que posee este cromosoma. Por tanto, al quedar incluido el cromosoma 22 en el mismo grupo cromosómico del cromosoma 21, presente en varias translocaciones y con historia de DacDCV, la familia se clasifica con alto riesgo.

Los tres modelos propuestos clasifican adecuadamente a las familias, según la experiencia de cada una. Las pocas discrepancias son normales cuando las estimaciones son empíricas y dependen de datos 'observados', por lo que el riesgo 'real de recurrencia' difiere [19].

Es muy importante resaltar la manera en que se procedió. Se propuso una estimación del riesgo de DacDCV, mediante una metodología que ofreció los resultados a partir de modelos. Específicamente, se obtuvieron estimaciones del riesgo de DacDCV a partir de la regresión logística en una base de 200 casos, la regresión múltiple en una base de datos de 41 familias (las cuales conforman los 200 casos anteriores) y la regresión de Poisson (41 familias). Cada uno de estos métodos poseen elementos positivos y negativos que se discutieron oportunamente.

La estimación del riesgo de DacDCV dista mucho de ser la mera obtención de una cifra: es el análisis cuidadoso de los resultados con todos los métodos propuestos y las conclusiones, además de la experiencia del especialista. Esta característica añade un componente de racionalidad importante en la metodología propuesta.

Agradecimientos

Se agradece el trabajo conjunto de los asesores genéticos de la red provincial de Genética Médica de Pinar del Río y en especial a la Lic. Olga Luisa Quiñones Masa y al Dr. Reinaldo Menéndez García, por sus colaboraciones para esta investigación.

18. Panasiuk B, Danik J, Lurie IW, Stasiwicz-Jarocka B, Lesniewicz R, Sawicka A, et al. Reciprocal chromosome translocations involving short arm of chromosome 9 as a risk factor of unfavorable pregnancy outcomes after meiotic malsegregation 2:2. *Adv Med Sci.* 2009;54(2):203-10.

19. Nussbaum RL, McInnes RR, Willard HF, editors. *Thompson and Thompson. Consejo genético y evaluación del riesgo.* 7th ed. Barcelona: Elsevier Masson; 2008.

Anexo

Clasificación de las familias en grupos de riesgo: alto, moderado y bajo*

Familias	Regresión logística	Regresión múltiple	Regresión de Poisson
1	Riesgo moderado	Bajo riesgo	Riesgo moderado
2	Alto riesgo	Alto riesgo	Alto riesgo
3	Riesgo moderado	Alto riesgo	Riesgo moderado
4	Alto riesgo	Alto riesgo	Alto riesgo
5	Bajo riesgo	Bajo riesgo	Bajo riesgo
6	Bajo riesgo	Bajo riesgo	Bajo riesgo
7	Alto riesgo	Alto riesgo	Alto riesgo
8	Alto riesgo	Alto riesgo	Alto riesgo
9	Alto riesgo	Alto riesgo	Alto riesgo
10	Alto riesgo	Alto riesgo	Alto riesgo
11	Alto riesgo	Alto riesgo	Alto riesgo
12	Bajo riesgo	Bajo riesgo	Bajo riesgo
13	Bajo riesgo	Bajo riesgo	Bajo riesgo
14	Bajo riesgo	Bajo riesgo	Bajo riesgo
15	Bajo riesgo	Bajo riesgo	Bajo riesgo
16	Bajo riesgo	Bajo riesgo	Bajo riesgo
17	Bajo riesgo	Bajo riesgo	Bajo riesgo
18	Bajo riesgo	Bajo riesgo	Bajo riesgo
19	Bajo riesgo	Bajo riesgo	Bajo riesgo
20	Bajo riesgo	Bajo riesgo	Bajo riesgo
21	Bajo riesgo	Bajo riesgo	Bajo riesgo
22	Bajo riesgo	Bajo riesgo	Bajo riesgo
23	Bajo riesgo	Bajo riesgo	Bajo riesgo
24	Bajo riesgo	Bajo riesgo	Bajo riesgo
25	Bajo riesgo	Bajo riesgo	Bajo riesgo
26	Bajo riesgo	Bajo riesgo	Bajo riesgo
27	Bajo riesgo	Bajo riesgo	Bajo riesgo
28	Bajo riesgo	Bajo riesgo	Bajo riesgo
29	Riesgo moderado	Alto riesgo	Bajo riesgo
30	Riesgo moderado	Bajo riesgo	Riesgo moderado
31	Riesgo moderado	Alto riesgo	Riesgo moderado
32	Bajo riesgo	Bajo riesgo	Bajo riesgo
33	Alto riesgo	Alto riesgo	Alto riesgo
34	Riesgo moderado	Alto riesgo	Riesgo moderado
35	Bajo riesgo	Bajo riesgo	Bajo riesgo
36	Bajo riesgo	Bajo riesgo	Bajo riesgo
37	Alto riesgo	Alto riesgo	Alto riesgo
38	Riesgo moderado	Bajo riesgo	Riesgo moderado
39	Bajo riesgo	Bajo riesgo	Bajo riesgo
40	Riesgo moderado	Bajo riesgo	Alto riesgo
41	Alto riesgo	Alto riesgo	Alto riesgo

* En la mayoría de las familias coinciden los tres modelos. En las familias 1, 3, 30, 31, 34 y 38 solo coinciden los modelos de regresión logística y de Poisson. Solo en las familias 29 y 40 no hay coincidencia entre los modelos descritos.