Proc. Nat. Acad. Sci. USA
Vol. 68, No. 4, pp. 820–823, April 1971

Mutation and Cancer: Statistical Study of Retinoblastoma

ALFRED G. KNUDSON, JR.

Graduate School of Biomedical Sciences and M. D. Anderson Hospital and Tumor Institute,
The University of Texas at Houston, Houston, Texas 77025

Communicated by James V. Neel, February 8, 1971

ABSTRACT Based upon observations on 48 cases of retinoblastoma and published reports, the hypothesis is developed that retinoblastoma is a cancer caused by two mutational events. In the dominantly inherited form, one mutation is inherited via the germinal cells and the second occurs in somatic cells. In the nonhereditary form, both mutations occur in somatic cells.

The second mutation produces an average of three retinoblastomas per individual inheriting the first mutation. Using Poisson statistics, one can calculate that this number (three) can explain the occasional gene carrier who gets no tumor, those who develop only unilateral tumors, and those who develop bilateral tumors, as well as explaining instances of multiple tumors in one eye.

This value for the mean number of tumors occurring in genetic carriers may be used to estimate the mutation rate for each mutation. The germinal and somatic rates for the first, and the somatic rate for the second, mutation, are approximately equal. The germinal mutation may arise in some instances from a delayed mutation.

The origin of cancer by a process that involves more than one discreet stage is supported by experimental, clinical, and epidemiological observations. These stages are, in turn, attributed by many investigators to somatic mutations. Ashley (1) has calculated that the common cancers are produced by a number of mutations that varies from 3–7, according to the specific cancer. It was pointed out by Ashley that cancer could still be a process occurring in two physiologically different stages, as suggested by studies in chemical carcinogenesis, the number of mutations involved in each stage varying from cancer to cancer. What is lacking, however, is direct evidence that cancer can ever arise in as few as two steps and that each step can occur at a rate that is compatible with accepted values for mutation rates. Data are presented herein in support of the hypothesis that at least one cancer (the retinoblastoma observed in children) is caused by two mutational events.

PATIENT DATA

The records of all retinoblastoma patients admitted to the M. D. Anderson Hospital, some 48 cases during the period 1944–1969, were reviewed. These cases are tabulated (Table 1) with respect to unilaterality or bilaterality, sex, age at diagnosis, and family history. Whenever possible, the number of tumors in each eye was estimated.

DISCUSSION

Several authors have concluded that retinoblastoma may be caused by either a germinal or a somatic mutation (2–5). The fraction of cases inheriting the mutation may be estimated indirectly. The percentage of all cases that are bilateral is approximately 25–30 (5), the present series being higher (48%) than that because of referral bias. All bilateral cases should be counted as hereditary because the proportion of affected offspring closely approximates the 50% expected with dominant inheritance (5). On the other hand, of the 70–75% of all cases that are unilateral, only 15–20% are thought to be hereditary (3, 5); thus, 10–15% of all cases are unilateral and hereditary. The percentage of all retinoblastoma patients with the hereditary form is, therefore, in the range 35–45; among these, 25–40% are unilateral and 60–75% are bilateral. In contrast, 55–65% of all retinoblastoma cases are of the nonhereditary form and all are unilateral. These distributions are summarized in Table 2.

Some patients that inherit the gene for retinoblastoma are never affected, although they transmit the trait to offspring who may become affected. The size of this group is difficult to estimate. Previous estimates range generally from 1–20%. Some authors have probably overcorrected for ascertainment bias. The estimate by Falls and Neel (2) of a range 1.5–10% represents an attempt to avoid this bias.

If the above estimates are correct, then carriers of the germinal mutation are distributed as follows:

Unaffected	1–10%
Unilateral	25–40%
Bilateral	60–75%

These data strongly suggest the possibility that tumors are distributed in accord with a Poisson distribution. If the Poisson distribution is followed, the mean number of tumors, m, formed per gene carrier may be estimated.

The necessary calculation is complicated by the fact that more than one tumor can occur per eye and that multiple tumors may appear unilaterally. For mean numbers of tumors of one–five, the fractions that should by chance be unilateral and bilateral may be calculated as shown in Table 3. It is seen that $m = 3$ is the value most compatible with the distributions noted above for unaffected, unilaterally affected, and bilaterally affected carriers of the mutant gene.

This distribution of tumors additionally predicts a multiplicity distribution within affected hereditary cases. It is difficult to test this prediction because so many tumors are very large when seen and, in fact, this situation was true for all of our unilateral cases, for both eyes in nine of the bilateral cases, and for one eye in the remaining bilateral cases. However, a tumor count was possible in one eye of each of 14 bilateral cases (Table 1). The expected numbers for one eye among bilateral cases may be calculated from the binomial distribution, when terms of the unilateral (o,r and r,o) classes are eliminated, and are shown in Table 4. In the last columns of Table 4 are shown the numbers of tumors observed in single eyes in our 14 cases and in 52 cases of retinoblastoma re-

TABLE 1. *Cases of retinoblastoma*

(a) Bilateral cases							(b) Unilateral cases				
Case	Hospital number	Sex	Age at diagnosis (months)	No. of tumors: right	No. of tumors: left	Family history	Case	Hospital number	Sex	Age at diagnosis (months)	No. of tumors	Family history
1	03712	F	8	*	*	no	24	00198	F	48	*	no
2	11571	M	3	*	*	no	25	03705	M	22	*	no
3	12163	F	11	*	*	no	26	06600	M	33	*	no
4	17076	M	2	*	1	no	27	08847	M	38	*	no
5	18025	M	60	1	*	affected sib.	28	08997	F	47	*	no
6	18237	M	22	*	*	no	29	19118	M	50	*	no
7	20291	F	4	3	*	no	30	24986	M	32	*	no
8	22699	F	18	2	*	no	31	26961	M	28	*	no
9	24729	F	30	*	*	no	32	33131	F	31	*	no
10	30464	M	3	2	*	no	33	36322	F	29	*	no
11	30470	F	6	*	1	no	34	40061	F	21	*	no
12	37837	M	7	*	2	affected sib.	35	40306	M	46	*	no
13	41134	M	9	3	*	no	36	41628	F	36	*	no
14	43391	F	4	5	*	no	37	45583	F	73	*	no
15	44176	F	13	*	*	no	38	46371	M	29	*	no
16	44649	M	18	*	4	no	39	47070	F	15	*	no
17	46860	F	24	*	*	no	40	52892	M	52	*	no
18	59704	F	44	1	*	no	41	54321	M	24	*	no
19	67024	F	5	*	*	no	42	61533	M	8	*	no
20	68422	M	12	*	1	no	43	64465	F	19	*	no
21	69224	M	3	*	*	no	44	64622	M	36	*	no
22	72656	M	12	*	1	no	45	68543	F	34	*	no
23	74616	M	15	1	*	father and paternal uncle	46	69502	F	27	*	no
							47	74498	M	10	*	no
							48	76092	F	8	*	no

* Number of tumors not determined.

ported by Stallard (6). The data clearly fall between $m = 2$ and $m = 3$.

This calculation permits another; namely, that of the probability that a mutant cell will develop into a tumor. If n is the total number of cells in the two retinae that have the potential for tumor formation, then m/n expresses the probability that a cell with the inherited mutation will develop into a tumor cell. The retinoblastoma is derived from a cell that generates both the inner neuroblastic layer (which gives rise to the ganglion and amacrine cells) and the outer neuroblastic layer (which gives rise to the bipolar, horizontal, and visual receptor cells). The order of magnitude of the number of retinoblasts is probably reflected by the order of magnitude of the number of ganglion cells, cells derived from the early-differentiated inner molecular layer. The number of ganglion cells has been estimated at 2×10^6 per retina, or 4×10^6 for two eyes (7). Using this value as an approximation for n, we find that the probability (m/n) that a cell with the inherited mutation will develop into a tumor cell is 0.75×10^{-6}. Since a majority of hereditary cases occurs within the first 2 years of prenatal and postnatal life, the probability, expressed as a mutation rate per year at either member of an autosomal gene pair, would be one-fourth of this value, or approximately 2×10^{-7} per year. This estimates the rate of a second mutation in cells already abnormal as a result of an inherited mutation at another gene site.

TABLE 2. *Distribution of retinoblastoma cases by type and laterality (3, 5)*

	Bilateral	Unilateral	Total
Hereditary	25–30%	10–15%	35–45%
Nonhereditary	0	55–65%	55–65%
Total	25–30%	70–75%	100%

Unilateral cases that do not result from germinal mutations constitute 55–65% of the total cases. If these cases arise by mutations at the same sites as in the hereditary form, then what is the relation between the somatic and germinal mutation rates for the first mutation? Let the two rates for this first mutation be μ_g (germinal) and μ_s (somatic), and the probability of the second event be m/n for both cases. Furthermore, let i be the total incidence of retinoblastoma, f_h the fraction of the hereditary cases, f_n the fraction of nonhereditary cases, q the population frequency of the germinal mutant gene, and s the coefficient of selection for the germinal mutant gene. Then the incidence of hereditary cases is:

$$f_h \cdot i = 2q(1 - e^{-m}).$$

Since $\mu_g = sq$, then

$$f_h \cdot i = 2\mu_g(1 - e^{-m})/s,$$

and

$$\mu_g = s \cdot f_h \cdot i/2(1 - e^{-m}).$$

The incidence of nonhereditary cases, $f_n \cdot i$, is equal to the product of the probability of the first mutation in somatic cells in the first year of fetal and postnatal life, $2\mu_s \cdot n$, and

TABLE 3. *Expected distributions of tumors for various mean numbers*

Total tumors both eyes (r)	Probability (Poisson)	Fraction unilateral $2(1/2)^r$	Frequencies for various values of m				
			1	2	3	4	5
0	e^{-m}	...	0.368	0.135	0.050	0.018	0.007
1	me^{-m}	1	0.368/0*	0.271/0	0.150/0	0.073/0	0.034/0
2	$\frac{m^2e^{-m}}{2!}$	1/2	0.092/0.092	0.135/0.135	0.112/0.112	0.073/0.073	0.042/0.042
3	$\frac{m^3e^{-m}}{3!}$	1/4	0.015/0.046	0.045/0.135	0.056/0.168	0.049/0.147	0.034/0.131
4	$\frac{m^4e^{-m}}{4!}$	1/8	0.002/0.013	0.011/0.079	0.021/0.147	0.024/0.172	0.021/0.148
5	$\frac{m^5e^{-m}}{5!}$	1/16	0.000/0.003	0.002/0.034	0.006/0.095	0.010/0.147	0.011/0.158
6	$\frac{m^6e^{-m}}{6!}$	1/32	0.000/0.001	0.000/0.012	0.002/0.049	0.003/0.101	0.004/0.136
7	$\frac{m^7e^{-m}}{7!}$	1/64	...	0.000/0.003	0.000/0.022	0.001/0.059	0.002/0.098
≧8	$\sum_{r=8}^{r=\infty} \frac{m^re^{-m}}{r!}$	$2(1/2)^r$	...	0.000/0.001	0.000/0.010	0.000/0.050	0.000/0.132
Totals: none			0.368	0.135	0.050	0.018	0.007
unilateral			0.477	0.464	0.347	0.233	0.148
bilateral			0.155	0.399	0.603	0.749	0.845

* Unilateral/bilateral.

TABLE 4. *Frequencies of tumors in one eye of bilateral cases for various values of mean number (m) for both eyes**

Tumors, one eye	Expected frequencies (%)				Observed numbers		Observed frequency
	m = 1	m = 2	m = 3	m = 4	14 cases, present series	52 cases, Stallard (6)	66 cases (%)
1	77	59	43	32	7	28	53
2	20	29	33	31	3	14	26
3	3	10	17	21	2	7	14
4		2.5	6	11	1	3	6
5			1.8	4.2	1		1.5

* The frequencies of bilateral tumors for various values of r and m are obtained from Table 3. The relative distribution of tumors between right (d) and left (s) eyes is obtained from the binomial expansion, $(d/2 + s/2)^r$. On multiplication, one obtains the expected distribution of tumors among bilateral cases, as illustrated for $m = 1$:

r	Frequency, bilateral (Table 2)	Frequency of tumors one eye			
		1	2	3	4
2	0.092	0.092			
3	0.046	0.023	0.023		
4	0.013	0.004	0.006	0.004	
5	0.003	0.000	0.001	0.001	0.000
	Totals	0.119	0.030	0.005	
	Per cent	77	20	3	

the probability of the second event, m/n, from which

$$f_n \cdot i = 2\,\mu_s \cdot n \cdot m/n = 2\,\mu_s \cdot m,$$

and

$$\mu_s = f_n \cdot i/2m.$$

The relationship between mutation rates is therefore:

$$\mu_s/\mu_g = f_n/f_h \cdot 1/ms \cdot (1 - e^{-m}).$$

An estimate of s is difficult, as it has become lower with therapeutic success. For purposes of this calculation, we assume a value of $s = 0.5$. Using this and values noted above for the other parameters, we calculate

$$\frac{\mu_s}{\mu_g} = \frac{0.60}{0.40} \cdot \frac{1}{(3)\,(0.5)} \cdot (0.95) \approx 1.$$

It is apparent that the two mutation rates are of a similar order of magnitude.

The incidence of retinoblastoma is approximately 5×10^{-5}, from which the germinal mutation rate, μ_g, may be calculated:

$$\mu_g = \frac{f_h \cdot i \cdot s}{2(1 - e^{-m})} = \frac{(0.4)\,(5 \times 10^{-5})\,(0.5)}{(2)(0.95)} =$$

5×10^{-6} per generation.

This rate is close to that of $6\text{–}7 \times 10^{-6}$ calculated by Vogel (3). In the nonhereditary form, the mutation rate is expressed per year, which, assuming a generation time of 25–30 years, yields an estimate of approximately 2×10^{-7} per year.

Although the above data are incompatible with two *independent* mutational "second events", they do not constitute

direct evidence that a single independent "event" of any kind is involved. If a second, single event is involved, the distribution of bilateral cases with time should be an exponential function, i.e., the fraction of the total cases that develops in a given period of time should be constant, as expressed in the relationship $dS/dt = -kS$, and $\ln S = -kt$, where S is the fraction of survivors not yet diagnosed at time t, and dS is the change in this fraction in the interval dt. As shown in Fig. 1, this is indeed the case. By contrast, the fractional decrease in unilateral cases per unit time does not show this relationship (Fig. 1). Although 15–20% of the unilateral cases should be of the hereditary type and so contaminate the data, the observations more nearly fit the anticipated two-mutation expression, $\ln S = -kt^2$, derived by Burch (8). That a difference in mean age at diagnosis exists between unilateral and bilateral cases has been noted previously. The respective mean ages at diagnosis for bilateral and unilateral cases have been reported in other series as 15 and 24 months (9) and 15 and 29 months (10), and in the present series are 15 and 32 months.

The exponential decline in new hereditary cases with time reflects the occurrence of a second event at a constant rate in a declining population of embryonal cells. For the nonhereditary cases, this declining population of cells must experience two independently occurring events. New cases of both types occur only in childhood because the embryonal cells vanish.

The data presented here and in the literature are consistent with the hypothesis that at least one cancer, retinoblastoma, can be caused by two mutations, each of which occurs at a rate of the order of 2×10^{-7} per year. One of these mutations may be inherited as a result of a previous germinal mutation that occurs at about the same rate. Those patients that inherit one mutation develop tumors earlier than do those who develop the nonhereditary form of the disease; in a majority of cases those who inherit a mutation develop more than one tumor. On the other hand, the probability that an individual not inheriting a mutation would develop more than one tumor is vanishingly small, so that nonhereditary cases are invariably unifocal.

The two-mutation hypothesis is consistent with current thought on the mutational origin of cancer. Ashley (1) has reviewed two-hit and multiple-hit theories of carcinogenesis and concluded that the common cancers are produced by about 3–7 mutations. Interestingly, one of the lowest estimates was for brain tumors, which are, like retinoblastoma, derived from neural elements. Ashley suggests that the two-stage hypothesis of initiation and promotion may still be correct, in that each stage may result from more than one mutation.

In the present series of 48 cases, three, all bilateral, are familial. In one case an affected father and his affected brother had unaffected parents. The same situation of affected sibs with unaffected parents exists for our other two cases. Such instances have been noted repeatedly in the past, the most dramatic report being that of Macklin (11), who found not only four families resembling those in the present series, but also ten families with affected cases in more widely separated relationships. Occasionally an unaffected parent may actually have had a retinoblastoma that has undergone spontaneous regression, but this must happen very rarely (2). On the other hand, if one attempts to attribute these occurrences to decreased penetrance, an inconsistency develops. Prior to the appearance of cases in a branch of a family, penetrance is very low; after its occurrence, penetrance is very high. As pointed out by Neel (12), this circumstance is precisely that discussed by Auerbach (13) with respect to the dominantly inherited split hand, or lobster claw, deformity. Auerbach compared this pattern with that found in *Drosophila*, in response to chemical mutagenesis by nitrogen mustard, and refers to it as an example of delayed mutation, or premutation.

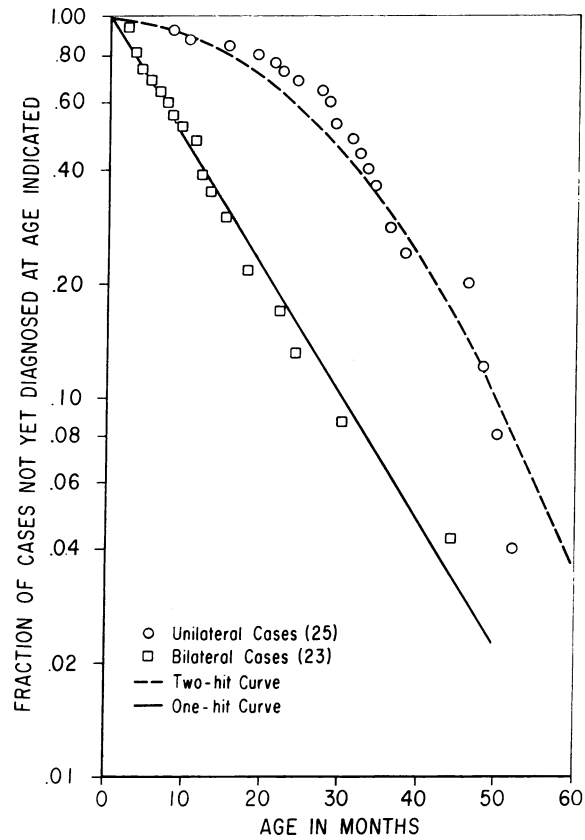

FIG. 1. Semilogarithmic plot of fraction of cases of retinoblastoma not yet diagnosed (S) vs. age in months (t). The one-hit curve was calculated from $\log S = -t/30$, the two-hit curve from $\log S = -4 \times 10^{-5} t^2$.

1. Ashley, D. J. B., *Brit. J. Cancer*, **23**, 313 (1969).
2. Falls, H. F., and J. V. Neel, *Arch. Ophthalmol.*, **46**, 367 (1951).
3. Vogel, F., *Z. Menschl. Vererb. Konstit-Lehre*, **34**, 205 (1957).
4. Smith, S. M., and A. Sorsby, *Ann. Hum. Genet.*, **23**, 50 (1958).
5. Schappert-Kimmijser, J., G. D. Hemmes, and R. Nijland, *Ophthalmologica*, **151**, 197 (1966).
6. Stallard, H. B., *Tr. Ophth. Soc. U.K.*, **82**, 473 (1962).
7. Van Buren, J. M., in *The Retinal Ganglion Cell Layer* (Charles C Thomas, Springfield, Ill., 1963), p. 64.
8. Burch, P. R. J., *Proc. Roy. Soc.*, **162 B**, 223 (1965).
9. Leelawongs, N., and C. D. J. Regan, *Amer. J. Ophthalmol.*, **66**, 1050 (1968).
10. Jensen, O. A., *Acta Ophthalmol.*, **43**, 821 (1965).
11. Macklin, M. T., *Amer. J. Hum. Genet.*, **12**, 1 (1960).
12. Neel, J. V., in *Methodology in Human Genetics*, ed. W. J. Burdette (Holden-Day, San Francisco, 1962), p. 203.
13. Auerbach, C., *Ann. Hum. Genet.*, **20**, 266 (1956).