Environmental Health Perspectives
Vol. 42, pp. 127-136, 1981

Detection of Genetic Effects of Environmental Agents

by Edmond A. Murphy*

The fundamental problems in population monitoring for genetic effects are twofold: the binomialized nature of the data and the lower power due to small risk of finding positive results. The binomial character is artificial, even forced, and can with advantage be replaced by more refined analysis, and by a focus on all mutations, not merely harmful ones. Moreover, a binomial treatment ignores accessory information (birth order, clustering, etc.). But this objective requires that an explicit model be used instead of nonparametric methods; a cancer may represent multiple independent hits that should be separately scored; sequencing of a codon or its product may show multiple distinct changes.

A sage . . . made the discovery that the flesh of swine . . . might be cooked . . . without the necessity of consuming the whole house to dress it.

Charles Lamb: *Dissertation on Roast Pig*

There is a pompous humility that I deplore in others, which I am about to represent as a virtue in myself. One who is not well up in a field is apt to harp on the very elementary aspects; and such is our conceit, that we hope that such topics are not merely elementary, but also elemental. Just so, the demi monde of mathematics uses florid formulas, but the pure mathematician wonders about obvious things: what "greater than" may mean, or whether a sheet of paper has two surfaces.

The Problem

The fundamental problem in the whole system of monitoring for genetic effects in populations seems to me to reside in two features: first, the data are more or less binomial in type—that is, they deal with individuals that either have, or have not, some particular characteristic; and secondly, the probability of a successful search, that is, of finding an aberrant case, is very small, perhaps one chance in a million or some small multiple thereof. If we find that, in the face of some environmental exposure, there are authentic increases in these rates that are due neither to defective sampling nor to the vagaries of chance, well and good: we are equipped to do battle over preventive measures. But failing to detect such changes, we can reassure, or be reassured, only if our tests have had high enough power. Now in this context, power means large samples, and large samples mean trouble. For either we must be dealing with an abnormality so conspicuous that the evidence is straightforward and recognition of all cases assured—stillbirths, sex ratios, multiple births—measures which will certainly have been exploited; or we must have an expensive system of vigilation and endless agonizing over the excellence and uniformity of diagnostic standards. Such an enterprise is cumbersome, and there must be at least systematic doubts as to whether it will work at all. Unfortunately, a cogent system of validation will, like belling the cat, surely prove at least as cumbersome as the main enquiry.

My misgivings will no doubt be obvious: that even with an enormous budget, methods demanding massive samples will not work. If, then, we are not to go this way, we are driven back on attacking our two fundamental problems: binomiality and low spontaneous rates.

Binomiality

Who uses the binomial distribution and why?

The binomial is the simplest of all, nontrivial,

*Division of Medical Genetics, Johns Hopkins University, Baltimore, Maryland 21205.

nondegenerate, distributions. If an event had fewer than two possible, distinct, outcomes, it could not be random. Now dichotomies of the binomial kind are imputed in two contexts: penetrating fundamentality and crass, even fatuous, naivity. Examples of fundamental binomiality would be existence and nonexistence, or matter and antimatter, or positive and negative electrical charges. Examples of crass binomiality would be "Americans" and "non-Americans"; or, in the tradition of Sellar and Yeatman (*1*), "good kings" and "bad kings." The essential difference is this: fundamental binomiality deals with something so near to the limit of our conceptual resolving power, that we are at a loss to find a constituent that will furnish any more basic term of discourse. There are not, for instance, degrees of existence. If something does not, and cannot, exist it is not an object of science. The gap between existence and its contrary is so enormous, that all other features assume trivial proportions. In such cases, we binomialize because we cannot escape doing so. It is arguable that the ultimate realities are dichotomous. This topic is the field of the fundamental scientist's fundamental scientist, the pure mathematician, the philosopher.

Crass binomiality is a less awesome phenomenon, in which the gap between the two states is much less wide, and may even be purely constructive. The difference between male and female, so the activists tell us, is largely fictive; and while sober scientists may not altogether corroborate this view, nevertheless we have no criterion of sex to which we can appeal for distinguishing unambiguously in disputed cases. The difference between "black" and "white" persons may conceivably have sociological meaning, but—in the United States at least—scarcely any biological merit whatsoever. For, insofar as race is a colorimetric term it is obvious, without any further formal documentation (*2*), that color behaves as virtually an uninterrupted variable* and not a dichotomous one; and if race is a matter of genes, as is usually implied, studies on genetic admixture have repeatedly shown the shallowness of the dichotomy (*3-5*). Yet race is still widely used as a crass binomial character. Crass binomiality is, like as not, a product of ambition, distraction, uglification and derision; it pays scant heed to detail, to logical structure; it is, as we have seen, a kind of dead end of statistical power. It is the tool of the lazy; of the superficial; and of the logically timorous, who do not wish to venture any surmise that may prove false.

*Such a variable is often inaccurately termed "continuous."

You will notice that I have crassly binomialized binomiality itself. In fact, I am quite sure that there is an uninterrupted ontological scale from the fundamental to the crass (Table 1). Rating on the scale of fundamentality-crassness depends on three features: natural grouping, reproducibility of the grouping, and the extent to which there exist more worthwhile and efficient means of assessment. So, to return at last to my topic, we might wonder where on this scale case-monitoring for new genetic mutations lies. My own belief is that it is towards the crass end of the continuum. Let me try to marshall my arguments.

In the first place, we must have doubts whether, even at a crass level, binomialization rather than polytomization, is an adequate description. The more we look, the more evidence we find of genetic heterogeneity. Consider, for instance, the alarming rate at which McKusick's (*6*), Borgaonkar's (*7*), and other catalogs are growing; the evidence of Harris (*8*), and Boyer (*9*) on the ubiquity of chemical polymorphisms, mostly (so far) of structure and enzymic function; the horrendous diversity of tissue types (*10*). More and more it grows on us that not two, but multiple, alleles per locus may be the rule. Figure how much mutation there must be, and watch the anxious frowns on the faces of the population geneticists trying to reconcile this profusion with classical theories of genetic load. Then recall that we are crassly dividing our observations into those with, and those without, harmful mutations (which, I suppose, wear white hats and black hats, respectively).

Now, to be sure, if we are to make our inferences from coarse studies on populations—coarse not from choice but for logistical reasons—why, we can probably do little better. Or again, if our sole interest is the assessment of the scale of damage, the binomial method may do well enough. But this approach is to ignore mechanisms and how we may use them to refine our means of inference. At least one should seriously consider

Table 1. Binomial representation: crassness and fundamentality.

Fundamentality	
Existence	Nonexistence
Positive electrical charge	Negative electrical charge
Hemoglobin type A	Hemoglobin type S
Male	Female
Righthanded	Lefthanded
Clever	Stupid
Mad	Sane
Negro	Caucasian
Employable	Unemployable
Friendly	Unfriendly
Crassness	

that we may have hold of the wrong end of the stick. The pragmatic objective of our enquiry is to prevent harmful mutations; but we should not confuse the pragmatic objective with the means of enquiry. Etiologically, the point of the point may not, perhaps, be the harm, but the mutation. One concern of society is to prevent fires. But that does not mean that the only relevant channel of research to prevent them is by counting fires and looking at when and where they occur, and whatever crude etiological factors that one might routinely think up without having any idea of what the nature of a fire is. The most profitable means might deal with the comparatively neutral field of the chemical properties of building materials. I would not approach prevention of harmful effects by monitoring harmful effects only. For if mutation is chemically what we believe it to be, is it necessarily true, is it even plausible, that those factors that cause deleterious mutations differ from those that produce neutral, or even beneficial ones?

Power

This line of thought I shall develop further. But to keep the argument orderly, let us revert to the second issue, sample size. It is, of course, a commonplace of epidemiology, that effects are studied, not simply in proportions but in aggregations. We know from probability theory, for instance, that the average proportions affected are not influenced by whether or not cases are mutually independent. But of course mutual nonindependence is itself valuable evidence of etiology; indeed, one definition of epidemiology is the study of the nonuniform distribution of cases in a population. Classical genetics has a similar definition: the study of biological variation. Unfortunately, implementing this definition of quantitative genetics has been seriously, and perhaps disastrously, blunted by equating variation with variance.

Now the simple binomialization of data and the interpretation of them as binomial by the methods usually employed by the unreflecting, take no account of nonindependence. To the contrary: the usual binomial procedures applied glibly to proportions are valid only if (among other conditions) the outcomes for all units are indeed independent. Where there is reason to suppose that this assumption is violated, the practice is to retreat into a subpopulation, for which independence does seem a warranted assumption, e.g., within a social class, or an occupation, or a family; and so successively. But there are logical weaknesses in this kind of successively conditioned inference.

Crass binomialization may be seen as a nonparametric analysis; but even in this "safe" and conservative field one can do better. Techniques such as the theory of runs have been devised to test for serial independence. One of the basic concerns where outcomes are not independent, is what the nature of the contagion may be. In genetics, we would think, for instance, of birth order. Thus, later members of a sibship are more likely to be affected by achondroplasia (because of paternal age); by Down's syndrome (because of maternal age); by Rh incompatibility (because of cumulative risk of sensitization); and by birth injury (because of multiparity and the risk of precipitate birth). But in congenital syphilis, clustering tends to occur in early pregnancies. Clustering in middle pregnancies may be due to domicile (during an intermediate stage of parental prosperity) in an area with a high level of some teratogen. But none of these highly informative patterns would show up in a crass binomial analysis, in which the number of affected progeny was regarded as sufficient (i.e., an adequately summarizing) statistic.

A natural extension of birth order (which is an ordering in a discrete, univariate, space) is ordering in multiple dimensions. But the foremost of these, clustering of residence or place of birth, is a much more complicated topic for two reasons. First, it is surprisingly difficult to give an adequate definition of what we mean by a cluster. The politically unscrupulous, like all wicked people, have a practical cunning for drawing boundaries in such a way as to manipulate the outcomes of elections. And precisely the difficulty of proving the gerrymander, is the lack of a sufficiently general theory of clustering against which the charge can be judged. Granted that we can draw closed loops, what are reasonable constraints on the shapes for these loops? Under the null hypothesis of uniform risk, provided that the loops are drawn beforehand, it does not matter what shapes they are. But for an unspecified alternate hypothesis where the loops are decided from the outcomes, interpretation is more difficult. There is a respectable theory of clustering (*11*); but I suppose that we have at least as much interest in interpreting the pattern as we have in merely rejecting the null hypothesis. Besides, the constitution of the alternate hypothesis (as is well known) determines where the rejection region lies. The null hypothesis might be rejected either because the occurrence of cases is too regular or too clumped. If, regardless of all etiological factors, precisely every hundredth child had a particular congenital defect, the hypothesis of uniform risk would be implausible. Yet the hypothesis of clustering would be

even more implausible. We must recognize the diversity of etiologies: for the unifying thread might be a toxin distributed down a river valley; or pollutant in the lee of a factory; or radiation to houses built on a particular hill; or the flow of traffic.

The other problem of clustering is the variation in the density of the population. Under the null hypothesis of uniform risk, it is clear that (other things being equal) the expectation of the number of cases will be proportional to the size of the population in the area. Even if extreme cold or dryness are factors, one can hardly expect much of a cluster at the north pole or the more desolate parts of the Sahara Desert. If an acceptible map of Canada by a projection that conserves area (Fig. 1a) were used to plot cases, clusters of almost anything whatsoever, (even social isolation!) will tend to occur in the big towns and, in the Western Provinces, at the most southerly parts, which is where the population is concentrated. But a rescaling to represent a uniform population density would allow genuine clusters to show up much more clearly. The map (Fig. 1*b*) inspired by that of Skoda and Robertson (*12*) but much simplified, could be made ever more refined in detail, so that all districts of all cities could be individually identified. Of course maps could be devised, all of which give uniform population densities, which are topologically equivalent, but yield different shapes. We could devise several forms, each designed to detect certain patterns. For instance, in Figure 2A is represented an imaginary terrain in which contour lines are shown with a superimposed grid. A river is flowing northwards, its descent being steepest, and therefore its flow fastest, in the sourthern, more mountainous part of its course. In Figure 2B, the repatterned map takes account of westerly wind, which is strongest in the northern plains and weakest in the protection of the hills. The scale is thus contracted from west to east in the northern part. (The wind has the effect of scattering and hence "diluting" the concentration of cases, which must be regathered by crowding the scale.) In Figure 2C we suppose that the effects are waterborne, and hence there is greatest scattering where flow is fastest, just as a dye in a stream will be dispersed faster and further down, than across, the stream. Thus the southern part of the river must be foreshortened; and where density of population for the grid is to be conserved, the lateral scale must be stretched. These two distorsions are merely illustrations. Remembering that my topic is genetic effects, we might capitalize on the fact that human populations tend to migrate along river valleys, as Cavalli-Sforza and colleagues (*13*) demonstrated some years ago.†

Further Increasing the Power

Granted that we have extorted from the data all the refinement and the ancillary information that the nonrandomness of the cases may furnish, and granted that there is a practical limit to the size of the sample we may hope for, what more can we do to increase the power? I can think of three devices.

Increase the Sensitivity

We have already considered how we may do this by switching the focus from gross and disabling phenotypes to equivalent random changes of any kind. The argument here is that the sensitive index is rate of mutation, not rate of disease. In this approach, one would be substituting genetic skill of the few specialists for the unpolarized watchfulness of the many generalists. The burgeoning field of analysis of nucleic acids by restriction enzymes offers enormous promise. But we note three practical problems that it generates. The investigative tool is at present technically complex and may always be so. It remains to be seen whether it will ever be accessible to, say, the field epidemiologist. Secondly, to capitalize on power, we will need to study parents and children, since mutation is cogently established by demonstrating genetic content —in genes or chromosomes—of the offspring that cannot have been inherited. I need hardly point out the problems arising from doubtful parentage, especially paternity. We usually check for this contamination by genetic compatibility of marker genes. But there is a logical trap; for of course we have to decide how much of any incompatibility may be ascribed to new mutation, and when we should abandon this explanation in favor of nonparentage (*15*). An obvious device is to exploit the difference between the degrees of uncertainty about nonpaternity and nonmaternity. (Of course one cannot ignore the inescapable confounding of any such effect with biological differences due to the sex of the disputed parent or to the special impact of the intrauterine milieu.) As we have noted, new mutation appears to increase with paternal age. (However, does nonpaternity also increase with the husband's age?). One may console oneself with the

†Recently the proceedings of a workshop on cartography and epidemiology held in 1976, have been published (*14*). None of it seems to be directed to rare events, or (more specifically) to detecting evidence of clustering.

FIGURE 1. Maps of Canada: (*a*) conventional (zenithal equidistant projection); (*b*) isodemographic. (This mapping, one of many possible mappings that are topologically and demographically equivalent, makes the density of the population uniform.) Hence, under the null hypothesis, that cases occur entirely at random, independently, and at constant risk, their distribution should follow Roach's probability algebra (*11*). Systematic departures from this algebra—due to heterogeneity of risk ("hot spots"), to nonindependence ("contagion"), or to channels of spread (wind, water, food distribution, etc.) will produce typical patterns of distortion.

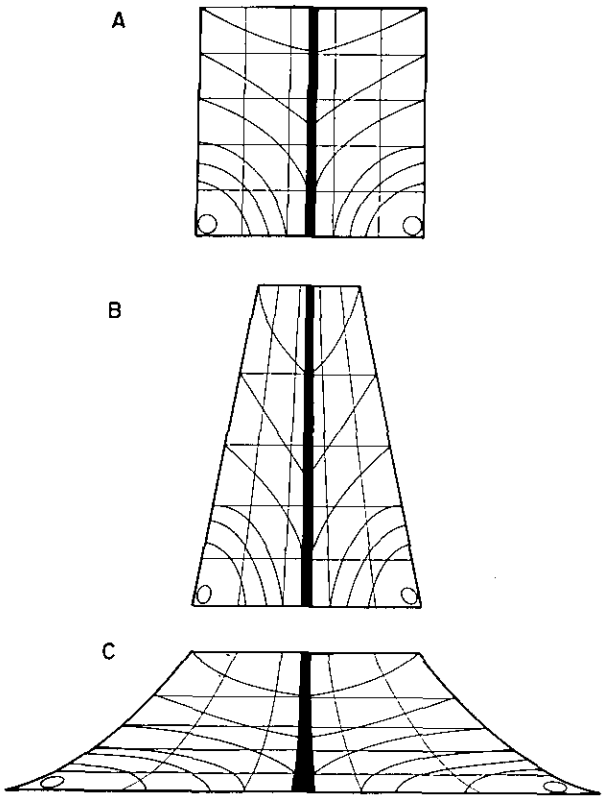

FIGURE 2. Etiological mapping. (A). The physically undistorted characteristics of an (imaginary) region. As in all three diagrams, a grid is shown in fine lines, contours of height in somewhat heavier lines, a river in solid black. The river rises in the hills in the south (bottom of the diagram), and flows northwards into a broadening plain, where there is progressively less physical protection from the prevailing westerly winds, and airborne factors are more rapidly and widely disseminated. In the early part of its courase, the river is falling most rapidly, as represented by the somewhat greater closeness of the contours; hence, linear flow is also fastest and (for instance) a degradable waterborne toxin will be more widely spread. (B) Compensatory distortion to offset changing velocity of wind. Areas are conserved, so that the foreshortening east-west must be compensated for by extension north-south. (C) Compensatory distortion to offset the changing rate of linear flow in the river. The latter two mappings should restore homogeneity of risk under specific hypotheses. For epidemiological purposes, it would in general be necessary also to make the maps isodemographic.

thought that the same distortions are likely to occur in different groups, and that the strength of our evidence will lie, not in absolute values, but in comparisons of otherwise similar groups.

The third snag is, in principle, easily remediable, but clings, like original sin, to everything we do. It is the defect that, by some inexorable law of weakest links, the more refined the studies that the scientist does in the laboratory, the more carelessly he conducts the sampling procedure. I suspect that the best strategy would be to have epidemiologists and biochemists handcuffed together in symbiotic pairs throughout the entire study. This step should ensure that as much attention is paid to the denominator as to the numerator. But, even so, we would need special supervision to make sure that they do not corrupt each other. There is an obvious danger that the epidemiologist may get interested in the biochemistry, and then all is lost.

Artificially Increasing the Rate of Mutation

For any given size of sample, the variance of the estimate increases as the proportion affected increases from 0% to 50%. Thus the implication of (say) a 10% difference in the proportions of affected subjects in two samples is least impressive where the trait is common (Table 2). However, paradoxically the maximum possible power of a test for effect may be enhanced by increasing the larger proportion of positives (Table 2). Could we then use this strategy? There are three approaches, each with its own particular problem.

First, experimental studies in man. There seem to be insurmountable ethical constraints here. I am not sure that I can defend them formally; but it is supposed that most mutations are harmful and the risk of inducing them unwarranted. It is doubtful that even a volunteer could ethically give consent *ex parte* for a procedure that may jeopardize the phenotype of any future progeny.

The second procedure, then, is experimentation

Table 2. Power in comparing proportions in two samples each of size 1000 from binomial distributions.

Percentage affected		Number of standard errors separating the means
Group 1	Group 2	
a. Fixed difference between proportions		
1	11	9.6
2	12	8.9
5	15	7.6
10	20	6.3
30	40	4.7
45	55	4.5
b. Fixed proportion in the first group		
1	2	1.8
1	3	3.2
1	5	5.3
1	10	9.0
1	20	14.6
10	15	3.4
10	20	6.3
10	25	9.0
10	30	11.5
10	35	14.0
10	40	21.7

either in animals or on human tissues *in vitro*. I am not sure that I want to get into the ramifications of such arguments by analogy. But, at least we have to be concerned about the bland assumption of proportionate effects. It is an old, and unresolved problem as to whether there is indeed a threshold below which mutation does not occur (*16*). By its nature, this problem is difficult to solve without experiments of immense size. There must also be some misgiving about the significance of the response of an isolated tissue. For we know enough about repair processes in the intact organism to have considerable doubts about the relevancy of rates of being affected in one tissue alone. Skin, for instance, is much more often called upon than other tissues to repair damage caused by ultraviolet rays.

The third procedure is to explore the whole problem at a finely divided level. Radiation of various types is much more rapidly and efficiently assessed than the incidence of a disease. If we understood precisely the mechanism of action of such radiation in mutation, or teratogenesis, we would be able to resolve our major problem. Of course this proposal has an almost surrealistic quality, which will doubtless not be soon attained. However I raise the issue purely on principle. Of course, even this solution depends on intrinsically discrete events and is hence subject to all the logical shortcomings of binomialized monitoring of individuals. But the discreteness is now obtruding at a vastly more refined level. My moral is that even in a discrete world we must not assume that at any particular level of individuation we have hit bedrock in resolving power. As I shall suggest presently, there may be, for instance, something intermediate between counting phenotypes and counting gamma rays.

Strategic Withdrawal from Crass Binomiality

So I am led to suggest that we try our best to break free from binomialization of our data, at least to the point at which the level of resolution is much greater than that with which we are too often content. There are several methods of doing so.

First, and by far the most important, whatever we can do to impart structure to our crass binomial model, we should do with enthusiasm. To say how this task is to be done would be to propose a blueprint for most of modern biology. Even the most Olympian ambitions of a later speaker would not aspire to such a goal. But at the least perhaps I can give a few illustrations of the kind of thing I mean.

In our monitoring system we register each person affected by cancer as a single event; and that is probably the best we can do on the evidence. But if we are in a position to impart structure to this binomialized outcome, we can do better. Suppose, for instance, that we accept the Nordling-Armitage (*7*, *8*) model that cancer is a multiple-hit process. Then, rather tritely, it is produced by multiple events which, ideally, could be individually counted. If we have, and seize, the opportunity to identify intermediate stages, we may do much better than by simply counting cancers. For instance, it is widely held (*19*) that bowel cancer occurs only in an antecedent polyp. If it takes one hit to convert normal bowel into a polyp, and a second hit to convert the polyp into a cancer, then we have two events to be scored, not one. The casual observations of my colleagues in gastroenterology are that polyps are perhaps an order of magnitude more prevalent than cancers. Hence we have, on this hypothesis, at least two events we can score, and at present, most epidemiologists are scoring cancers only, the less common of the two. But even that is not all. For estimates (*20*, *21*) have suggested that common cancers may be something like a seven-hit process; and one supposes that, in principle, the several intermediate stages differ and that each transition may be scored as an event. Of course, I am not putting all this up as a finished product, or even as one that I believe literally. The point is that if we succeed in getting "inside" the process, we may break through the binomial barrier into something that is statistically much more efficient.

But even within the individual patient there may quite possibly be much more information than we have tapped. A burning issue at present is whether multiple polyposis of the colon (a disease in which we ourselves have a special interest) is monoclonal or polyclonal. If, as is plausible, it proves to be polyclonal, then each polyp is a separate event. Accordingly, we should not merely record whether or not the patient is affected, but how many polyps are present. The demonstration of polyclonal origin depends on having a marker that will Lyonize (*22*). The only one widely useful at present is glucose-6-phosphate dehydrogenase (G6PD) for which there are many readily distinguishable alleles. No report on the findings for a suitable carrier has yet been published.

An admirable example of the use of the internal structure of the model and one on much more secure footing than bowel cancer (being much more readily accessible to study) is retinoblastoma. This tumor used to be called an autosomal dominant condition with incomplete penetrance (*23*). This

more likely than others; that there is degeneracy in the code; that where amino acids are concerned, a substitution of serine for proline requires that at least one event has occurred, of valine for argenine, at least two, and lysine for phenylalanine, at least three. They should be scored accordingly. As a concrete example, in Table 3 are given several descriptor that owes little allegiance to formal analysis, and none whatsoever to concrete biology. The way the term has been used by the insouciant is almost a parody of meaningless pseudo-parameterization. Needless to say, any effort to estimate it has, for the most part, been crassly binomial.

Knudson (*26*, *27*) ingeniously converted the phenomenon of retinoblastoma into a two-hit process which led to explicit predictions about recurrence risk in relatives of various types, in twins, in unilateral and bilateral cases, etc., which could be verified, and were. Now this model had two tremendous advantages. It identified a population at high risk; and that, as we have seen, may enhance power vastly. Even if we went no further, it must be evident that such people with heritable cancer are a natural monitoring device who should receive intensive study for hazard, an attention which to my knowledge they have not received. But there is more. Knudson worked out a quantitative relationship to the number of retinoblasts (the vulnerable cells) which agreed with the recurrence ratio. We might plausibly argue that two physically separated retinoblastomas in the same eye or, *a fortiori*, one in each eye, constitute two separate events and may be scored as such.

Pathology abounds in graded systems, devised for prognosis or to assess therapy, but with which the etiologist does nothing and, in the absence of any pathogenetic theory, can do nothing. The dermatologists, who of all physicians have the most accessible of systems to study, must surely see the skin as an admirable means for monitoring mutagenic insults. They talk about skin cancers with a malignancy of grade 1/2; but surely there must be some information waiting to be quarried out of the difference between such lesions and more malignant ones? Yet once again, there must presumably be more information to be extracted from the number of lesions there are and where exactly they occur. There must also be added insight from the (statistical) regression of the number of lesions on the extent of exposure (see below). Of course, one gets nothing for nothing; and to extract the information efficiently one needs to have some grasp of the structure of the process, and must be prepared to take some judicious, intellectual, risks by making surmises that one is spared in using the crass binomial. That is what I meant by including the timorous among the binomializers.

Second, one might do more, perhaps much more, by capitalizing on relationship of degrees of heterogeneity. It is easy, for instance, to overlook the difference between systematic and unsystematic heterogeneity. Regression analysis is more powerful than analysis of variance, precisely because, over and beyond exploring heterogeneity, it makes use of the information on relationship. The whole notion of bioassay is based on the idea that the probability of failure can be stated as an explicit function of dosage. But clearly this may be improved on. The usual form ascribed to this relationship has no explicit biological meaning. Why should the minimum lethal dose of a drug follow a Gaussian, logistic, or lognormal distribution, and what pharmacological insight would it furnish if it did? But let us look at the same problem from the reverse end. Suppose that we are studying the effect of a drug, the pharmacodynamics of which reflects an X-linked polymorphism, such as G6PD deficiency. If the genetic dose effect is linear, Lyonization will make the bioassay function a binomial of high order, which is close to Gaussian. In Negroes, in whom there is a high rate of deficiency, we would expect to see a bipartite distribution, and comparison between the shapes of the two component parts should allow us to separate genetic, from other, sources of variation. I imply nothing revolutionary about this particular example; it is more important as an illustration of the kind of elucidation that can emerge when the statistics are adapted to the data rather than the data to the statistics. Of course one could devise examples where the dose-response curve is not even monotonic.

Let us take another tack. A mutation, in the halcyon days of plant and Drosophila genetics, and still in the minds of the environmentalists, is a phenotypic phenomenon. For the modern molecular geneticist it is genotypic. And since evolution leads to, even consists in, progressive divorce between the genotype and the phenotype, the term "mutation" becomes more and more ambiguous, especially in higher organisms. I would venture the opinion that the almost total estrangement between the population geneticist and the molecular biologist, and of both from the internist, is due to failure to recognize this divorce. Genetic selection operates on phenotype. But what is inherited is genotype.

To the environmentalist, a mutation is a mutation is a mutation. But, for instance, we now know a great deal about point mutation—that is, muta-

Table 3. Evidence of mutational events in variant hemoglobins.[a]

Type of hemoglobin	Chain	Position	Wildtype		Mutant type		Number of mutations
			Amino acid	Code	Amino acid	Code	
A	Beta	6	Glutamic acid	GAA	Valine	GUA	1
C Harlem	Beta	6	Glutamic acid	GAA	Valine	GUA	2
	Beta	73	Glutamic acid	GAA	Asparagine	AAU	
Bristol	Beta	67	Valine	GU.	Aspartic acid	GAC	2
Not known	–	–	Phenylalanine	UUU UUL	Lysine	AAA AAG	3

[a]The information for this table was kindly furnished by Dr. K. D. Smith.

tion due to substitution of one nucleotide for another. We know that some substitutions are much term was given no precise meaning. "Penetrance" has been widely invoked without any attempt to define its logical properties (*24*) or without even any recognition of the fact that whether or not it has a heritable component makes a fundamental difference to its pattern of manifestation (*25*). Penetrance is an excellent example of a purely external hemoglobin variants from which variable numbers of mutations from the wild type may be inferred. They include an instance of two mutants within one triplet, and one where there are two mutants at different amino acids. My colleague, Kirby D. Smith, who kindly furnished this information, knows of no hemoglobin recorded so far in which there is an amino acid substitution that can be explained only by three or more mutations. We know that some amino acid substitutions may produce phenotypic changes, inconspicuous even by chemical criteria; that the immediate effect of a mutation may be masked because the mutant is recessive, or hypostatic to some mutant at another locus. It must be evident that in bewailing the minute risks of new mutations and the low power of our monitoring, we are starving in the midst of plenty, not to say opulence.

The facts are a good deal less sure in the detailed analysis of chromosomes, especially with the results now available by prophase banding. The field is at an exploratory stage and, while it is showing great promise, its formulations are at present too unstable for me to make any useful compact statement about it. But when the details are completed, I am confident that we shall have a tool at least tenfold more sensitive than what has been in use in the decade since Casperson introduced banding.

Conclusion

There is still insufficient recognition, even among professional epidemiologists, that to binomialize is the refuge of the destitute, a statistical device to be avoided at all costs. The basic problem is that it makes the data two-valued and that in consequence an absolute upper limit on what we can extract from the data, is set by the sample size. From a binomial test, one can never reject a null hypothesis of $p = 1/2$ at the 5% level from a sample of size 4. no matter how grossly false the hypothesis. This limitation does not apply to the z, t, or F distributions, for instance. There is never any excuse in serious epidemiology for splitting continuous, unimodal, distributions into categories; but more, there should be an unceasing drive to convert what appear to be ineluctably binomial data into data with biological meaning, logical structure, epidemiological import, and statistical power. Until this policy is tried, or shown to be not merely troublesome but impossible, I have no tears to shed for our misfortunes.

From the Division of Medical Genetics, Johns Hopkins University School of Medicine. The research for this paper was supported under grant number GM24736 of the National Institutes of Health and a grant from the Julia Baker Fund.

REFERENCES

1. Sellar, W. C., and Yeatman, R. J. 1066 and All That. Methuen, London, 1930.
2. Harrison, G. A., and Owen, J. J. T. Studies on the inheritance of human skin color. Ann. Hum. Genet. 28: 27-37 (1964).
3. Glass, B., and Li, C. C. The dynamics of racial intermixture—an analysis based on the American Negro. Am. J. Hum. Genet. 5: 1-20 (1953).
4. Steinberg, A. G., Stauffer, R., and Boyer, S. H. Evidence for a GM allele in the Gm system of American Negroes. Nature 188: 169-170 (1960).
5. Workman, P. L., Blumberg, B. S., and Cooper, A. J. Selection, gene migration and polymorphic stability in a U.S. white and Negro population. Am. J. Hum. Genet. 15: 429-37 (1963).
6. McKusick, V. A. Mendelian Inheritance in Man. Johns Hopkins University Press, Baltimore, 1966, 1968, 1971, 1975, 1978.
7. Borgaonkar, D. S. Chromosome Variation in Man: A Catalog of Chromosomal Variants and Anomalies. Johns Hop-

kins University Press, Baltimore, 1975; AR Liss, New York, 1977, 1980.
8. Harris, H., and Hopkinson, D. A. Average heterozygosity per locus in man: an estimate based on the evidence of enzyme polymorphisms. Ann. Hum. Genet. 36: 9-20 (1972).
9. Boyer, S. H., Kunkel, L. M., and Smith, K. D. The proportion of all point mutations which are unacceptable: an estimate based on hemoglobin, amino acid and nucleotide sequences. Can. J. Genet. Cytol. 20: 111-137 (1978).
10. Albert, E. D., Mickey, M. R., and Terasaki, P. I. Genetics of four new HL-A specificities in the Caucasian and Negro populations. Transplant Proc. 3: 95-100 (1971).
11. Roach, S. A. The Theory of Random Clumping. Methuen, London, 1968.
12. Skoda, L., and Robertson, J. C. Isodemographic Map of Canada. Geographical Paper No. 50. Information Canada, Ottawa, 1972.
13. Cavalli-Sforza, L. L., Barrai, I., and Edwards, A. W. F. Analysis of human evolution under random genetic drift. Cold Spring Harbor Symp. Quant. Biol. 29: 9-20 (1964).
14. DHEW. Proceedings, 1976 Workshop on Automated Cartography and Epidemiology, Arlington, Va., March 18, 1976, DHEW Publ. PHS 79-1254, 1979.
15. Murphy, E. A. Application of genetics to epidemiology. In: Trends in Epidemiology, G. T. Stewart, Ed., CC Thomas, Springfield, Ill., 1972, pp. 102-138.
16. Cornfield, J. Carcinogenic risk measurement. Science 198: 693-99 (1977).
17. Nordling, C. O. A new theory on the cancer-inducing mechanism. Brit. J. Cancer 7: 68-72 (1972)
18. Armitage, P., and Doll, R. The age distribution of cancer and a multistage theory of carcinogenesis. Brit. J. Cancer 8: 1-12 (1954).
19. Muto, T., Bussey, H. J. R., and Morson, B. C. The evolution of cancer of the colon and rectum. Cancer 36: 2251-70 (1975).
20. Armitage, P., and Doll, R. A two-stage theory of carcinogenesis in relation to the age distribution of human cancer. Brit. J. Cancer 9: 161-169 (1957).
21. Ashley, D. J. B. Colonic cancer arising in polyposis coli. J. Med. Genet. 6: 376-368 (1969).
22. Fialkow, P. J. Clonal origin and stem cell evolution of human tumors. In: Genetics of Human Cancer J. J. Mulvihill, R. W. Miller, and J. F. Fraumeni, Eds., Raven, New York, 1977.
23. Macklin, M. J. Inheritance of retinoblastoma in Ohio. Arch. Ophthal. 62: 842-851 (1959).
24. Trankell, A. Penetrance calculus in population genetics. Am. J. Hum. Genet. 8: 44-48 (1956).
25. Murphy, E. A. Incomplete penetrance and the ascertainment of recessive characters in man. 1. Introduction and distribution theory. J. Chron. Dis. 23: 9-28 (1970).
26. Knudson, A. G., Hethcote, H. W., and Brown, B. W. Mutation and childhood cancer: A probabilistic model for the incidence of retinoblastoma. Proc. Natl. Acad. Sci. (U.S.) 72: 5116-20 (1975).
27. Knudson, A. G. Genetic and environmental interaction in the origin of human cancer. In: Genetics of Human Cancer, J. J. Mulvihill, R. W. Miller, and J. Fraumeni, Eds., Raven, New York, 1977.