Am. J. Hum. Genet. 63:947–954, 1998

STATISTICAL GENETICS '98
Bayesian Linkage Analysis, or: How I Learned to Stop Worrying and Love the Posterior Probability of Linkage

Veronica J. Vieland

Departments of Preventive Medicine and Environmental Health, Division of Biostatistics, and Psychiatry, University of Iowa College of Medicine, Iowa City

> *We should really like to know, at the end of study, the probability that we have found a linkage, as pointed out by Cedric Smith (1959) >35 years ago.* (Elston 1997)

> *Bayes' Theorem has the advantage that, unlike any other method of statistical inference, it gives the answer directly as a probability.…This probability has a direct meaning, and does not need to be hedged about with qualifications, unlike a significance level.* (Smith 1959)

Bayesianism permits us to measure directly the probability that we have found a linkage, or what is called the posterior probability of linkage (PPL). No other statistical paradigm does this. Smith advocated a Bayesian approach to linkage analysis as early as 1959, but the approach seems never to have gained general popularity (but see, e.g., Hauser and Boehnke 1993; Thomas et al. 1997). If what we really want to know is the PPL, and if there exists a statistical method for measuring the PPL, then why is the method not more widely applied? Is there some reason that we cannot get what we really want?

The Bayesian Paradigm and Bayes's Theorem

The Bayesian school of statistical inference is generally associated with two contentious notions: first, that probability is a measure of degree of belief (rather than, e.g., the more common notion that probability is a measure of long-run frequency); and, second, that the business of statistical inference is the orderly ranking of beliefs along a correspondingly subjective probability scale (rather than, say, the measurement of evidence based on data). (For the Bayesian perspective, see, e.g., Edwards et al. [1963]; for two other paradigms in statistical genetics, also see Vieland and Hodge [1998].) But, as Smith (1959) pointed out, the only truly Bayesian tool needed for calculation of the PPL is Bayes's theorem, which itself is not contentious at all.

In its most familiar form, Bayes's theorem states simply that, for a set of mutually exclusive and exhaustive outcomes $A_1,\ldots A_n$ and for any other outcome B,

$$P(A_i \mid B) = P(B \mid A_i)P(A_i)/\sum_{j=1..n} P(B \mid A_j)P(A_j)$$
$$= P(B \mid A_i)P(A_i)/P(B)\ . \qquad (1)$$

As a tool for calculation of conditional probabilities for events, the theorem, which follows from the first principles of probability theory, is unimpeachable. (I have given the theorem in the form most useful to my argument here, following, e.g., Feller [1968, p. 124]; but see also Hacking [1965] and Stuart and Ord [1987] for historical accounts of the postulate.)

If we define probability as a long-run frequency of occurrence, then only things that can occur (events) can have probability. On the other hand, if we define probabilities as degrees of belief, then propositions (hypotheses) become the sorts of things that have probability. Thus, in spite of the fact that not all hypotheses can be said to "occur" in anything but a metaphoric sense, a Bayesian has no problem interpreting equation (1) in those cases in which B represents some set of data, A_i is a hypothesis and the $P(A_j)$'s represent the prior probability distribution associated with the space of all possible hypotheses ("prior" in the sense of holding *before* looking at the data).

Received August 10, 1998; accepted for publication August 19, 1998; electronically published October 2, 1998.

Address for correspondence and reprints: Dr. Veronica J. Vieland, Departments of Preventive Medicine and Environmental Health, Division of Biostatistics and Psychiatry, University of Iowa College of Medicine, Iowa City, IA 52240. E-mail: veronica-vieland@uiowa.edu

0002-9297/98/6304-0005$02.00

A simple example will illustrate why this might be problematic. Let p be the parameter of a binomial distribution (the probability, say, that a coin lands heads), and let the hypothesis be H:$p = \frac{1}{8}$. A Bayesian has no particular difficulty in assigning a prior probability to H, because this can be done with reference to a state of mind. For instance, if I believe that most coins are "fair," then I may assign a relatively small prior probability to H. Generalizing to the probability distribution of the set of all possible values of p, I might argue that, even before the coin is tossed, a higher probability should be assigned to values near $\frac{1}{2}$ than to values farther away; but you might feel that, because you have no prior basis for preferring one value of p over another, the prior distribution should assign equal probability to all possible values.

For a Bayesian, this poses no problem, and we are each free to choose whichever prior-probability distribution better suits us. One consequence of adopting the Bayesian perspective, then, is that reasonable people can apply Bayes's theorem to the same data and, depending on their preferences regarding the priors, arrive at very different *posterior* probabilities for the hypothesis of interest, A_i ("posterior" in the sense of holding *after* looking at the data). This is often taken as invalidating Bayesianism in scientific applications (e.g., see Hacking 1965; Royall 1997). But this "subjective" property of the posterior probability is not a consequence of Bayes's theorem per se; it is a consequence of the nature of the hypothesis to which Bayes's theorem has been applied.

Suppose, by contrast, that the prior-probability distribution corresponds to a "chance set-up" (Edwards 1972, p. 50), in which each possible outcome occurs with some specific probability. For example, suppose that a lottery is conducted by first rolling a die, then picking one of six urns in accordance with what the die shows, and then drawing a ticket from the chosen urn. If the six urns give six different probabilities that our ticket matches the ticket drawn, then we would not hesitate to calculate the probability of winning as the weighted average of a winning ticket taken over the six urns, using as the weights the prior probabilities of each urn, in this example $\frac{1}{6}$ in each case; and we would have no difficulty in applying Bayes's theorem to obtain the posterior probability that the observed ticket had been drawn from any particular urn. Moreover, all of us, given the set-up (and barring computational errors), would arrive at the same number for this posterior probability.

Genetic segregation is a chance set-up, in the sense that there is a physical placement of genetic material along chromosomes, a set of stochastic laws governing meiotic processes, and, therefore, a true probability distribution for the set of all possible outcomes of any given meiotic event. The hypothesis of linkage represents a subset of such outcomes and is therefore readily interpretable as the kind of thing that has probability, quite without appealing to subjective considerations. We already make liberal use of other prior-probability distributions in linkage analysis, with similar justification. For instance, offspring probabilities are regularly calculated by averaging across all possible mating types, weighting the average by the prior probabilities (or population frequencies) of the mating types, where the prior-probability distribution of mating types is governed by a set of parameters—namely, the allele frequencies.

Of course, the posterior-probability distribution in the offspring depends on the prior gene frequencies, and there may be disagreements as to the best way to measure these, so that reasonable people could disagree as to which particular values should be used. But this is not to say that all choices for the gene frequencies are equally valid: the selection of the prior probabilities themselves is subject, in this case, to empirical verification. Indeed, I have never heard this used as an argument against stipulating the gene frequencies. In the case of genetic transmission, we may not know exactly what the chance set-up is—unlike the situation in the urn example—but we are on equal *conceptual* footing when we make use of the prior distribution of the recombination fraction θ in linkage analysis as we are when we appeal to the prior distribution of the urns in the lottery example.

Bayes's theorem becomes a tool for calculation of the posterior probability of *any* hypothesis only if we subscribe to the subjectivist view of things; but, as a tool for calculation of the posterior probability for hypotheses that refer to specific outcomes of chance set-ups, it does not require the subjectivist perspective—or entail the subjectivist consequences. The prior distribution of θ is a wholly empirical matter, just as is the prior distribution of mating types. Indeed, geneticists do talk uninhibitedly about "the probability of [the hypothesis of] linkage"—and not because we are all Bayesians.

Thus, there is no *principled* reason why we cannot calculate what we really want, the PPL. Of course, there is still the very real *practical* question of how the priors should be specified. But before I return to that question in detail, it will be helpful to first take a closer look at the PPL itself—to see just what it is; how it differs from some other, more familiar quantities; and to consider the implications of the PPL for establishing, and especially replicating, linkage for complex disorders.

What the PPL Is

The fundamental Bayesian operation applied to linkage analysis is really very simple in spirit. We start with a prior-probability distribution for θ, $f(\theta)$ and then, by considering the data D, we transform the prior distri-

bution into a posterior distribution $f(\theta|\mathrm{D})$ of θ, given the data, as in

$$f(\theta|\mathrm{D}) \underline{\Delta} \frac{f(\mathrm{D}|\theta)f(\theta)}{\int\limits_{\theta \leqslant \theta \leqslant 1/2} f(\mathrm{D}|\theta)f(\theta)\ d\theta} = \frac{f(\mathrm{D}|\theta)f(\theta)}{f(\mathrm{D})}\ , \qquad (2)$$

which is simply the analogue, in the continuous case, of the discrete form of Bayes's theorem, given above. (I will assume throughout that we are evaluating the evidence for linkage between a putative trait locus and one marker at a time.)

Equation (2) can, in turn, be used to calculate what we really want to know, the PPL. Let $\mathrm{H_L}$ represent the hypothesis that a trait locus and a given marker are linked, and let $\mathrm{H_{\bar{L}}}$ be the hypothesis that they are unlinked, so that $P(\mathrm{H_{\bar{L}}}) = [1 - P(\mathrm{H_L})]$. Then the PPL is simply the definite integral of equation (2):

$$\begin{aligned} \mathrm{PPL} \ \underline{\Delta}\ & P(\mathrm{H_L}|\mathrm{D}) \\ &= \int\limits_{0 \leqslant \theta < 1/2} f(\theta|\mathrm{D})\ d\theta \\ &= \int\limits_{0 \leqslant \theta < 1/2} \left[\frac{f(\mathrm{D}|\theta)f(\theta)}{\int\limits_{0 \leqslant \theta \leqslant 1/2} f(\mathrm{D}|\theta)f(\theta)\ d\theta} \right] d\theta\ . \qquad (3) \end{aligned}$$

Note that the prior $f(\theta)$ is assumed to be continuous on the interval $0 \leqslant \theta < \frac{1}{2}$, but with positive mass over the point $\theta = \frac{1}{2}$, so that $P(\mathrm{H_{\bar{L}}}) > 0$. An alternative form will be important later:

$$\begin{aligned} \mathrm{PPL} \ \underline{\Delta}\ & P(\mathrm{H_L}|\mathrm{D}) = [f(\mathrm{D}|\mathrm{H_L})P(\mathrm{H_L})]/ \\ & \quad [f(\mathrm{D}|\mathrm{H_L})P(\mathrm{H_L}) + f(\mathrm{D}|\mathrm{H_{\bar{L}}})P(\mathrm{H_{\bar{L}}})] \\ &= \left[\int f(\mathrm{D}|\theta,\mathrm{H_L})f(\theta|\mathrm{H_L})\ d\theta \right] / \\ & \quad \left[\int f(\mathrm{D}|\theta,\mathrm{H_L})f(\theta|\mathrm{H_L})\ d\theta \right. \\ & \quad \left. + \int f(\mathrm{D}|\theta,\mathrm{H_{\bar{L}}})f(\theta|\mathrm{H_{\bar{L}}})\ d\theta \right]\ . \qquad (4) \end{aligned}$$

The particular function of the data, $f(\mathrm{D}|\theta)$, can take any number of forms, but, in general, we can substitute the antilog of the usual LOD score. It happens that $\mathrm{LOD}(\theta) \underline{\Delta} \log_{10}[P(\mathrm{M,T}|\theta)/P(\mathrm{M,T}|\theta = \frac{1}{2})] \propto \log_{10} P(\mathrm{M}|\mathrm{T},\theta)$ for marker data M and trait data T (Clerget-Darpoux et al. 1986; Elston 1989). Moreover, although I have written equation (2) in terms of the probability $f(\mathrm{D}|\theta)$, the equation is really based on the underlying likelihood $\mathrm{L}(\theta) \ \underline{\Delta}\ k \times f(\mathrm{D}|\theta)$, where k is an arbitrary constant absorbing all functions of the data that do not involve θ (combinatorial terms, etc.). Because k is constant for all values of θ, it cancels out of the equation. Thus, provided that we are interested in conditioning on all the trait data, as is generally done (Ewens and Shute 1986; Hodge and Elston 1994; Vieland and Hodge 1996), any linkage-analysis program set up to calculate LOD scores can be used to calculate equation (2). Below, I return to the specification of the prior $f(\theta)$.

Equations (2) and (3) provide a simple and elegant method for updating the PPL as we accumulate new data. Suppose that we have two sets of data, $\mathrm{D_1}$ and $\mathrm{D_2}$, and that we have calculated the posterior distribution on the basis of the first set of data $\mathrm{D_1}$ alone. We can then substitute this posterior distribution $f(\theta|\mathrm{D_1})$ in place of the prior distribution $f(\theta)$ in equation (2), replace the data $\mathrm{D_1}$ by new data $\mathrm{D_2}$, and recalculate equation (2), which results in a new posterior distribution: $f(\theta|\mathrm{D_2})$. We can then reapply equation (3), integrating over $f(\theta|\mathrm{D_2})$, in order to calculate a new, updated PPL.

This procedure for updating the posterior distribution of θ and the PPL can be applied repeatedly, allowing us to accumulate linkage evidence across as many sets of data as we like. At each step of the way, we get what we want: a measure of the PPL itself, based on the total available evidence—that is, carrying over information regarding the distribution of θ, obtained on the basis of earlier data, toward the measurement of the evidence based on new data. This seems wholly in keeping with the spirit of scientific inquiry, for we would not usually require that investigators return to a complete state of ignorance between experiments; rather, we would like to be able to carry forward what we have learned from one experiment when we interpret the results of the next one.

The PPL and the Likelihood Ratio (LR)

In understanding what equations (3) and (4) *are,* it may be helpful to contrast them with a couple of familiar things that they are *not*. For instance, the antilog of the LOD score is an LR, which we can write, in a general form, as $P(\mathrm{D}|\theta)/P(\mathrm{D}|\theta = \frac{1}{2})$. As is well known, an LR is a measure of the relative probabilities of the data, given the two hypotheses, but is not a measure of the probability ratio of the hypotheses themselves (e.g., see Elston 1994). It can, however, be transformed into such a probability ratio, according to Bayes's theorem, via multiplication by the corresponding ratio of prior probabilities. This yields an alternative formulation of Bayes's theorem: "the posterior odds…is equal to the product of the prior odds and the likelihood ratio" (Stuart and Ord 1987 pp. 280–281; also see, e.g., Edwards 1972; Elston 1994).

We can also obtain posterior odds for linkage by dividing equation (4) by the corresponding equation for $P(H_{\bar{L}}|D)$, which gives

$$\frac{P(H_L|D)}{P(H_{\bar{L}}|D)} = \frac{f(D|H_L)}{f(D|H_{\bar{L}})} \times \frac{P(H_L)}{P(H_{\bar{L}})} \quad . \qquad (5)$$

This suggests an equivalence between $f(D|H_L)/f(D|H_{\bar{L}})$, as it appears in equation (5), and the LR, $P(D|\theta)/P(D|\theta = \frac{1}{2})$. But the two are, in fact, distinct.

An LR can only be used to compare "simple versus simple" hypotheses; in other words, it can only be evaluated for two specific values of θ at a time. This is precisely because the LR is calculated without reference to the prior distribution $f(\theta)$. We know—at least in principle—how to calculate the probability distribution of D for any *particular* value of θ, but what is $P(D|\theta < \frac{1}{2})$? To calculate $P(D|\theta < \frac{1}{2})$, we need to average the probabilities, for each particular value of θ, over the range $\theta < \frac{1}{2}$, weighting the average by the prior probabilities $f(\theta|\theta < \frac{1}{2})$. Without the prior probabilities, we cannot calculate this average and must therefore be satisfied to consider only one particular value of θ at a time.

The expression $f(D\,|\,H_L)$ in the numerator of equation (5) is shorthand for the full weighted average, $\int f(D\,|\,\theta,H_L) f(\theta\,|\,H_L)\ d\theta$, as in equation (4); but the expression $P(D\,|\,\theta)$ in the numerator of the LR represents the probability of D at one particular value of θ. (Because, under $H_{\bar{L}}$, θ can take only the value $\frac{1}{2}$, the corresponding denominators are the same.) This is a major strength of the LR in the many contexts in which we cannot or will not (for philosophical reasons) ascribe probability distributions to our parameters: precisely because the LR does not involve the prior distributions, in the absence of reasonable priors the LR can still be used as a measure of statistical evidence (Edwards 1972; Royall 1997). When we are able to specify the prior-probability distribution, however, the LR is inherently wasteful of information regarding its shape.

When it comes to picking the single value of θ to use in the numerator of the LR, a rich and elegant body of statistical theory, which owes its foundations in large measure to R. A. Fisher, instructs us to be particularly interested in that value $\theta = \hat{\theta}$ that maximizes the likelihood (i.e., the maximum-likelihood estimator), since it is this value that gives us the "best" evidence against the null hypothesis of no linkage. This is a perfectly satisfying approach so far as the LR itself goes (e.g., see Edwards 1972; Royall 1997), but we cannot then use the LR calculated at this one point, $\theta = \hat{\theta}$, to derive the posterior odds for linkage per se; rather, we can arrive only at the posterior odds for $\theta = \hat{\theta}$—namely, $[P(D|\theta = \hat{\theta})/P(D|\theta = \frac{1}{2})] \times [P(\theta = \hat{\theta})/P(\theta = \frac{1}{2})]$. If we wanted, we could, of course, multiply the LR evaluated at $\theta = \hat{\theta}$ by the prior odds for linkage, but note that the resulting quantity would always be greater than the true posterior odds for linkage.

The PPL and the Posterior Type I Error Rate

Another familiar quantity, which sometimes goes by the same name as the PPL, is the probability of linkage *given that we have rejected the null hypothesis* of no linkage. It was this latter quantity to which Morton appealed when he originally recommended a significance criterion of LOD $\geqslant$ 3 (Morton 1955). Using this criterion was meant to ensure (under certain conditions) that the probability that there really is linkage, given that we have observed a LOD $\geqslant$3, would be >95% (also see Morton 1998). There is nothing wrong with referring to this as a "posterior probability of linkage," but it is distinct from the PPL as defined above.

Morton was interested in a conventional statistical test of the hypothesis of linkage versus the hypothesis of no linkage, such that, whenever the LOD score exceeds some predetermined cutoff, we decide in favor of linkage. The test is characterized in terms of the probabilities of two types of errors that we might incur in making a decision: (*a*) the type I error probability α, or the probability that we reject the hypothesis of no linkage when it is true; and (*b*) the type II error probability β that we fail to reject the hypothesis of no linkage when it is false. Then, if we know the prior probability of linkage, $P(H_L)$, we can calculate the probability that there really is linkage associated with any given choice of the significance cutoff. For example, using a LOD cutoff of 3 for our significance test and applying Bayes's theorem, we calculate

$$\begin{aligned} &P(H_L|\,\text{LOD} \geqslant 3) \\ &\quad = [P(\text{LOD} \geqslant 3|H_L)P(H_L)]/ \\ &\qquad [P(\text{LOD} \geqslant 3|H_L)P(H_L) \\ &\qquad\quad + P(\text{LOD} \geqslant 3|H_{\bar{L}})P(H_{\bar{L}})] \\ &\quad = [(1-\beta)P(H_L)]/[(1-\beta)P(H_L) + \alpha P(H_{\bar{L}})] \quad . \end{aligned} \qquad (6)$$

Although, in some sense, equation (6) also represents a kind of posterior probability of linkage, it is not at all the same thing as the PPL. Equation (6) is a function of the predictive error rates of a particular decision-making procedure—namely, the rule telling us to reject the hypothesis of no linkage whenever the LOD score is $\geq$3. But the rule that we choose to employ in making such a decision surely has no bearing whatsoever on the actual probability that the chosen marker is linked to the disease. Accordingly, equation (6) is more aptly called "one

minus the posterior type I error rate of the test procedure" (as Morton did indeed call it). In statistics, this quantity is called the "*reliability*" of the test (Stuart and Ord 1987; Morton 1998); but, because that word has a different connotation in epidemiology, and to underscore that it is also a posterior measure, calculated after looking at the data, hereafter I will call it the "posterior reliability" (PR).

In equation (3) we calculate the probability of linkage conditional on the data D, whereas in equation (6) we calculate the probability of linkage conditional on having made a particular decision—namely, to reject the hypothesis of no linkage on the basis of the observed value of some statistic (not the PPL). This decision is, of course, related to the data, but only indirectly (e.g., see Hacking 1965; Edwards 1972; Royall 1997; Vieland and Hodge 1998). Why is it, after all, that Morton required a PR >95%? We are accustomed to accepting a type I error rate of 5% for many statistical tests, so it seems reasonable to adopt this level of significance with respect to the posterior type I error rate of the test as well. But, really, we should be interested in much smaller values of the PPL. Would you leave your umbrella at home because the probability of rain was only 45%?

Replication of Linkage and the PPL

I have mentioned above that equations (2) and (3) could be used to accumulate linkage evidence across data sets, but to use them in this way has important implications for how we decide when a linkage finding has been replicated. It is widely accepted that, to establish genetic linkage, any initial finding must be followed by replication in an independent set of data, and recent articles have reiterated the importance of maintaining stringent significance criteria in both initial and follow-up studies. For instance, Lander and Kruglyak (1995) recommend requiring an initial finding at a significance level of $\sim 4.9 \times 10^{-5}$ (for general pedigrees), followed by replication at a significance level of $\sim 1.7 \times 10^{-3}$ in an independent set of data. In addition, they and others (Elston 1994; Morton 1998) have stressed the need to maintain high power to detect linkage at both the initial and the follow-up stages.

The reason for both the stringency of the significance levels and the requirement of high power is that even a small *P* value can correspond to an unacceptably low PR, under certain circumstances (Smith and Sturt 1976; Génin et al. 1995; Morton 1998). For example, if the prior $P(H_L)$ is small, significance criteria need to be calibrated appropriately, in order to maintain an acceptable PR (Morton 1955). Complications such as multiple testing also may require that significance levels be sufficiently stringent to maintain the PR while compensating for genomewide type I error probabilities (Lander and Kruglyak 1995).

Similarly, as equation (6) shows, the PR is a function of both the type I and type II error rates of the testing procedure, so that, when the latter is large relative to the former, the PR can be low even for small *P* values (Stuart and Ord 1987; Elston 1994; Lander and Kruglyak 1995; Morton 1998). In this case, significance levels again need to be adjusted downward, in order to maintain a high PR. This problem is exacerbated by the fact that, although the power of a genomic screen to detect linkage to any one locus in a complex system may be only moderate to begin with, power to replicate at one *particular* locus in a replication study will almost certainly be worse (Suarez et al. 1994). Moreover, in linkage studies we cannot control or even accurately measure the power, because it depends on the underlying genetic model for the trait. Thus, it is really not possible to calculate the PR on the basis the *P* value at all.

These are the sorts of considerations to which Smith (1959, p. 301) was referring (in the passage that I have used as the second epigraph of the present article) when he said that significance levels needed to be "hedged about with qualifications" (also see Elston 1994). The *P* value is only reliable as a measure of posterior probability insofar as we control the type I and type II error rates of the test procedure on the basis of which it is calculated. That is to say that the need to control the error rates of our statistical procedures arises entirely from the simple fact that we are not directly measuring the PPL: controlling the error rates is our only line of defense against those unmeasured forces that cause a given *P* value to correspond to a range of possibilities for the PPL. If there were a one-to-one correspondence between *P* values and the PPL to begin with—if a *P* value of, say, .001 *meant* that the PPL was, for instance, 75%—then there would be no need to worry about significance levels; indeed, there would be no reason to perform significance tests to begin with!

Accepting a linkage finding based on only weakly significant *P* values in each of a few separate data sets carries an unacceptably low PR, but using the data sets in concert to measure the overall evidence for linkage in the form of the PPL carries no such danger. The point is not that a Bayesian analysis would yield more evidence for linkage across disparate data sets but, rather, that the only sure way to control the possibilities of misinterpreting the evidence with respect to the PPL, if indeed that is what we really wish to do, is to measure the PPL itself.

Bayesian analysis allows us to interpret the accumulated weight of the evidence for (or against) linkage, even when it is accumulated across multiple sets of data, each one only moderately informative on its own. The traditional requirement of *independent* replication would

appear to preclude the possibility that conclusive evidence can be obtained on the basis of evidence that accumulates in small increments, but perhaps, in studying complex disorders, we need to rethink this requirement. Of course there is another reason entirely that we require independent replication of any scientific finding—and that has to do with reproducibility of laboratory results. This is not a statistical issue, and extrastatistical methods (e.g., cross-typing of DNA at two independent laboratories or independent validation of diagnoses) are needed for this purpose.

But What about the Priors?

Of course, I have left the key difficulty for last: how do we specify the prior distributions? I think that the burden of assuming a particular form for $f(\theta)$ is less than might be supposed, but the argument here hinges on understanding exactly how the PPL is calculated, which is why I have left this topic for last. As equation (4) makes clear, there are really two separate decisions to be made: (*a*) what is the *prior-probability distribution of θ, given that there is linkage;* and (*b*) what is the *prior probability of linkage?* (The distribution of θ under $H_{\bar{L}}$ is simply 1 for $\theta = \frac{1}{2}$ and is 0 elsewhere).

The choice of a form for (*a*), $f(\theta|\theta < \frac{1}{2})$, will depend on experimental design. For example, when linkage to a particular candidate gene (and not just to what is sometimes called a "candidate locus," or a marker with prior evidence of linkage) is being tested, $f(\theta|\theta < \frac{1}{2})$ is given entirely by the genetic distance between the gene and the marker: if the marker is 1 cM away, then $f(\theta|\theta < \frac{1}{2}) = 1$ for $\theta = .01$ and $f(\theta|\theta < \frac{1}{2}) = 0$ elsewhere. On the other hand, if the study design calls for a genomic screen, then, under the assumption of linkage and at the marker closest to the trait locus, the appropriate distribution is the distribution of the distance between a random (trait) locus and its nearest marker, along a fixed marker map. (For calculation of a related distribution, see Elston and Lange 1975.) Specification of this distribution is straightforward, on the basis of the assumption of a uniform distribution for θ across the human genome, which appears reasonable on the basis of experimental results (Morton 1998). Existing approximations can be refined by taking into account, for instance, separate male and female recombination distances in those genomic regions where these vary substantially.

The critical point is that $f(\theta|\theta < \frac{1}{2})$ depends only on the density of the marker map, not on the genetic model for the trait: this distribution relates the location of one particular marker to one *particular* trait locus at a time (on the assumption that two or more trait loci do not share a closest marker) and is therefore independent of the total number of trait loci that exist. The specification of this portion of the prior distribution simply models the distance between pairs of loci, in the initial "chance set-up," in accordance with the underlying laws of genetic inheritance. Of course, we might still end up misspecifying this prior. As in any statistical analysis, it is important to evaluate the dependence of our results on the particular modeling assumptions used, and measurement of the PPL should be evaluated with respect to robustness to choice of the priors.

The choice of a form for (*b*), the prior probability of linkage between any given marker and a trait locus, might seem, at first glance, to be more problematic. This is the part of the distribution that *does* depend on the total number of underlying trait loci, since the more such loci there are, the greater the prior probability of linkage between the marker and at least one of the trait genes. Since the true number of trait genes cannot be known in advance, it is often said that we cannot specify the prior probability of linkage for a complex disorder. But notice that the smallest prior probability of linkage, under the assumption that there is at least one trait gene, occurs when there is exactly one trait gene. This means that assuming the existence of only one trait gene is equivalent to minimizing the prior probability of linkage—that is, maximizing the prior probability of no linkage $P(H_{\bar{L}})$. As equation (4) shows, the PPL depends on the relative magnitude of $P(H_{\bar{L}})$ with respect to the total posterior-probability distribution of θ; and larger values for $P(H_{\bar{L}})$ in the denominator reduce the magnitude of the total ratio. Thus, by assuming that there is exactly one trait gene, we arrive at a conservative way to calculate the PPL: in the presence of multiple genes, the true PPL will always be larger—all other things being equal—than the calculated value.

Again, the particular value of $P(H_{\bar{L}})$ will depend in part on study design. For instance, it will depend on the lengths of specific chromosomes or regions being tested (e.g., see Génin et al. 1995), differing, for example, when all chromosomes are considered or when only the autosomes are considered. In experimental organisms we can calculate this quantity directly from breeding data, whereas in humans $P(H_{\bar{L}})$ may involve some approximation. Various calculations for the probability that two loci are syntenic suggest a number in the narrow range .051–.054 (Morton 1998); under the assumption that linkage is detectable at $\leqslant 40$ cM, Elston and Lange (1975) calculated that the prior probability of linkage for two random loci is ~2%. It is possible that this number can be improved on in certain contexts, but, in most applications, small adjustments are unlikely to have large effects on the PPL. Again, this assumption can be tested.

Thus, once we look in detail at how equation (4) is put together, we see that to be Bayesian in this context

is not such an onerous undertaking: the necessary prior distributions can be established on sound empirical grounds, without the need to appeal to "subjective" considerations; and, what is more, they can be appropriately specified in the absence of information regarding the genetic model for the trait.

There is, of course, another very serious objection to the use of equation (2) in the study of complex disorders—namely, that "to arrive at a posterior probability of linkage, we must…assume prior probabilities for all the various possible modes of inheritance" (Elston 1994, p. 10). This criticism goes to the very heart of any likelihood-based analysis—not just a Bayesian one: if we do not know the true, underlying genetic model, then we cannot write the likelihood in its correct form. Whether this objection will pose a serious obstacle to Bayesian applications of the sort that I have described remains to be put to rigorous testing. I suspect—and preliminary data confirm (author's unpublished data)—that misspecification of the likelihood will be no more deleterious in Bayesian applications than in the calculation of LOD scores themselves. Greenberg et al. (1998*a*) and others have shown that, for LOD scores, accurate specification of the trait model is not essential (also see Greenberg et al. 1996, 1998*b*).

Of course, we will never be certain, in any given application, that our calculation of the PPL has *not* been influenced by misspecification of the model, although this possibility can be explored by changing the model and seeing what happens to the PPL. But what is our alternative? Many would argue that the only acceptable alternative, when the genetic model is unknown, is to restrict ourselves to the so-called model-free methods (see Schaid 1998 [in this issue]), whose validity is independent of the true mode of inheritance. This means that we must restrict our outcome measures to *P* values.

I find the underlying logic of this argument peculiar. What we *really* want to measure is the PPL, and the *P* value does not measure it. In this case, isn't restricting ourselves to the calculation of P values, simply because they are amenable to exact calculation, a bit like looking for our lost keys under a lamppost, just because that is where the light happens to be? Do you prefer to get—exactly—what you *don't* want, or would you rather have approximately *what you really do want?*

Acknowledgments

This essay represents, in part, work done in collaboration with Jian Huang and Kai Wang; I am indebted to Sue Hodge and John Ashkenas for critical readings of earlier versions. This work was supported by National Institute of Mental Health grants R01 52841 and K02 01432.

References

Clerget-Darpoux F, Bonaïti-Pellié C, Hochez J (1986) Effects of misspecifying genetic parameters in lod score analysis. Biometrics 42:393–399

Edwards AWF (1972) Likelihood. Cambridge University Press, London

Edwards W, Lindman H, Savage LJ (1963) Bayesian statistical inference for psychological research. In: Neyman J (ed) Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability. University of California Press, Berkeley

Elston RC (1989) Man bites dog? The validity of maximizing lod scores to determine mode of inheritance. Am J Med Genet 34:487–488

——— (1994) P values, power, and pitfalls in the linkage analysis of psychiatric disorders. In: Gershon ES, Coninger CR (eds) Genetic approaches to mental disorders. American Psychiatric Press, Washington, DC

——— (1997) Algorithms and inferences: the challenge of multifactorial diseases. Am J Hum Genet 60:255–262

Elston RC, Lange K (1975) The prior probability of autosomal linkage. Ann Hum Genet 38:341–350

Ewens WJ, Shute NC (1986) A resolution of the ascertainment problem. I. Theory. Theor Popul Biol 30:388–412

Feller W (1968) An introduction to probability theory and its applications. Vol 1, 3d ed. John Wiley & Sons, New York

Génin E, Martinez M, Clerget-Darpoux F (1995) Posterior probability of linkage and maximal lod score. Ann Hum Genet 59:123–132

Greenberg DA, Abreu P, Hodge SE (1998*a*) The power to detect linkage in complex disease by means of simple LOD-score analyses. Am J Hum Genet 63:870–879

——— (1996) Affecteds-only linkage methods are not a panacea. Am J Hum Genet 58:892–895

——— (1998*b*) Power, mode of inheritance, and type I error in lod scores and affecteds-only methods: reply to Kruglyak. Am J Hum Genet 62:202–204

Hacking I (1965) Logic of statistical inference. Cambridge University Press, London

Hauser ER, Boehnke M (1993) The posterior probability of linkage Am J Hum Genet Suppl 53:1012

Hodge SE, Elston RC (1994) Lods, wrods, and mods: the interpretation of lod scores calculated under different models. Genet Epidemiol 11:329–342

Lander E, Kruglyak L (1995) Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results. Nat Genet 11:241–247

Morton NE (1955) Sequential tests for the detection of linkage. Am J Hum Genet 7:277–318

——— (1998) Significance levels in complex inheritance. Am J Hum Genet 62:690–697

Royall R (1997) Statistical evidence: a likelihood paradigm. Chapman & Hall, New York

Schaid DJ (1998) Transmission disequilibrium, family controls, and great expectations. Am J Hum Genet 63:935–941 (in this issue)

Smith CAB (1959) Some comments on the statistical methods used in linkage investigations. Am J Hum Genet 11:289–304

Smith CAB, Sturt E (1976) The peak of the likelihood curve in linkage testing. Ann Hum Genet 39:423–426

Stuart A, Ord JK (1987) Distribution theory. Vol 1, 5th ed. In: Kendall's advanced theory of statistics. Oxford University Press, New York

Suarez BK, Hampe CL, Van Eerdewegh P (1994) Problems of replicating linkage claims in psychiatry. In: Gershon ES, Coninger CR (eds) Genetic approaches to mental disorders. Am Psychiatric Press, Washington, DC

Thomas DC, Richardson S, Gauderman J, Pitkäniemi J (1997) A Bayesian approach to multipoint mapping in nuclear families. Genet Epidemiol 14:903–908

Vieland VJ, Hodge SE (1996) The problem of ascertainment for linkage analysis. Am J Hum Genet 58:1072–1084

——— (1998) *Statistical Evidence: A Likelihood Paradigm,* by Richard Royall (review). Am J Hum Genet 63:283–289