Journal of Medical Genetics 1989, **26**, 305–308

X linked or autosomal recessive? A new approach to an old problem

ANDREW P READ
From the Department of Medical Genetics, St Mary's Hospital, Manchester M13 0JH.

SUMMARY Families in which a single male is affected with a disease which might be either X linked recessive or autosomal recessive present problems in counselling. Before female relatives can be counselled, the probabilities of each mode of inheritance must be assessed, taking into account the prior probabilities, the pedigree structure, any DNA probe data, and any carrier testing data. The widely used linkage analysis package LINKAGE can be used to do the calculation, which is much simpler than the conventional Bayesian method.

When a man has a disease which might be either X linked or autosomal recessive, and there is no family history to indicate the mode of inheritance, counselling his female relatives is difficult. The risks to his sister or daughter differ greatly according to the mode of inheritance. Wolff *et al*[1] discuss a typical example, two families where an isolated male has a disease which may be either X linked Becker muscular dystrophy (BMD) or autosomal recessive limb girdle muscular dystrophy (LGMD). In each case a sister of the affected man requested counselling about recurrence risks.

In these cases it is necessary to calculate the relative likelihood of X linked or autosomal inheritance in the light of all the available information: the population gene frequencies, the pedigree structure, any DNA probe results, and any carrier tests such as creatine kinase (CK) estimations. Traditionally[1–4] a Bayesian calculation has been used. This approach is indisputably valid in principle. It also fits well with the predilection of geneticists for Bayesian methods. The problem is that the calculation can be very formidable. Few of us could confidently derive an expression such as this[1]:

$$\frac{9/128\ \mu\ \delta\ ha\ hb\ hc\ \theta\ (1-\theta)^2\ (\theta^3+(1-\theta)^3)+1/64\ \mu\ \delta\ hb\ hc\ \theta(1-\theta)}{9/128\ \mu\ \delta\ ha\ hb\ \theta(1-\theta)\ [hc\ \theta^3(1-\theta)+\theta^4+hc\ (1-\theta)^4+\theta(1-\theta)^3]+1/64\ \mu\ \delta\ hb\ \theta(1-\theta)\ (hc+1)+1/64\ \mu\ \delta+3/128\ kb\ \mu\ (kc+1/2+ka\ kc+1/2\ ka)}$$

and even evaluating it arithmetically without making mistakes is tricky. A particular problem with the Bayesian approach is that the form of calculation is different in each family, so that it is not readily computerised.

I wish to suggest an alternative approach, based on linkage analysis programs (LIPED[5] or LINKAGE[6]). The way these programs calculate lod scores is to calculate the total likelihood of a pedigree, including any relevant data, for example, CK, on two alternative assumptions, namely that the loci in question are linked or are not linked. The ratio between these two likelihoods, expressed as a logarithm, is the lod score. It is well known that these programs can be used to calculate genetic risks. In this case the likelihood of the pedigree is calculated on the alternative assumptions that the proband is or is not a carrier of the disease. The ratio of these likelihoods is the odds that the proband is a carrier. I suggest that in exactly the same way the programs can calculate the likelihood of pedigrees on the two alternative assumptions that the disease is X linked or autosomal recessive. This may be seen as an alternative, easier way of calculating Bayesian joint probabilities (that is prior × conditionals), while the final odds ratio is an expression of the Bayesian posterior probability (that is, joint probability/sum of joint probabilities).

Received for publication 17 October 1988.
Revised version accepted for publication 2 December 1988.

Method

THE LINKAGE PROGRAM PACKAGE

The methods and examples given here should be read in conjunction with the documentation for the LINKAGE package,[7] which gives instructions for setting up the data and pedigree files and explains their use. The analyses shown used the programs PEDPOINT, PREPLINK, and MLINK in version 3·5 of LINKAGE.

GENE FREQUENCIES AND MUTATION RATES

The data file requires gene frequencies and muta-

tion rates. For the autosomal recessive option the gene frequency is estimated by Hardy-Weinberg from the population incidence, and the mutation rate can be set to zero. Including reasonable mutation rates does not significantly alter the outcome. For the X linked case, the gene frequency q must be set so that the probability an arbitrarily chosen woman is a carrier is $2pq \approx 2q$. Grimm,[8] following Haldane, has shown that this probability is

$$\frac{4\mu+2\mu f}{1-f}$$

(μ is the mutation rate and f the fitness of affected males)

and that the incidence of the disease in males is

$$\frac{3\mu}{1-f}.$$

For BMD and LGMD suitable values[8] are:
BMD $q=5{\cdot}0\times10^{-5}$ $\mu=5{\cdot}5\times10^{-6}$
LGMD $q=6{\cdot}6\times10^{-3}$ $\mu=0$.

RELATIVE PROBABILITIES ON PEDIGREE DATA ALONE

The LINKAGE programs require phenotypes at a second locus in addition to the disease locus. Where there is no relevant marker data, the trick[9] is to include a dummy locus with two alleles having gene frequencies 1 and 0, with everybody in the pedigree homozygous for the common allele. The program then calculates the probabilities correctly.

RELATIVE PROBABILITIES INCLUDING DNA MARKER TYPES

If the marker is X linked, the data file for X linked analysis will specify the true recombination frequency, and the data file for the autosomal analysis will have the recombination fraction 0·5. In general, the same pedigree is used for both analyses; however, with an X linked marker it may be necessary to change marker types of some sons for the autosomal analysis. If a father is type 1, the autosomal analysis will not allow a son to be type 2 because it assumes that the marker is autosomal. The son's type should be changed to 2–1. Since the marker is not linked to the disease, this has no effect on the likelihood or carrier risk, provided that it is done with males whose parents are included in the pedigree. The reverse situation (an autosomal marker in an X linked analysis) will rarely occur but may necessitate recoding of heterozygous males.

A correction is necessary to the likelihood in the autosomal case before it can be compared to the X linked likelihood. For each male without parents in the pedigree, the calculated likelihood must be divided by the gene frequency of the marker allele he carries. The gene frequency will be less than 1, so the effect is to increase the likelihood of the autosomal pedigree. This unfortunate complication is necessary because the autosomal calculation assumes he is homozygous, with likelihood q^2, whereas the X linked calculation assumes he is hemizygous, with likelihood q. There is no problem with females or with males whose parents are included in the pedigree. In practice the correction is very simply applied; for instance, in the example below, we have two males with no parents, and we simply divide by 1/4.

RELATIVE PROBABILITIES INCORPORATING CK DATA

Again a dummy locus must be included if there is no genuine marker data. Liability classes are defined for each CK level found, plus one for women with no CK value. Women whose CK is to be considered are coded as affected. The significance of the CK is contained in the penetrances assigned to that liability class in the PREPLINK file. The penetrances for genotypes 1–1, 2–1, and 2–2 are set to the CK likelihood that she is not a carrier (P_N), the CK likelihood that she is a carrier (P_C), and zero, respectively. Thus if the h value (P_N:P_C) is 2, the penetrances should be 0·67, 0·33, and 0 for genotypes 1–1, 2–1, and 2–2. The effect of this manoeuvre is to assign to her the genotypes 1–1, 2–1, and 2–2 with probabilities 0·67, 0·33, and 0, respectively.

The same liability classes must also be used for the autosomal analysis, but here the penetrances for genotypes 1–1, 2–1, and 2–2 must be set to P_N, P_N, and 0. Within each liability class the same figure (P_N) must be used for the penetrance of the normal genotype in the X linked and autosomal analyses.

Any other carrier test, whether for the X linked or the autosomal disease, can be incorporated in a similar way.

Results

The pedigree (figure) is adapted from Wolff *et al.*[1] Translated into the input format for LINKAGE it appears as in table 1. For the autosomal analysis, subjects 9 and 10 are recoded as heterozygous (see Methods). When the pedigree is put through the PEDPOINT or MAKEPED programs, preparatory to running MLINK, subject 8 is flagged as the proband for risk calculation.

The two data files (set up using PREPLINK) are shown in table 2. Results of the analysis (using MLINK) were: X linked analysis: log likelihood −8·822886, carrier risk 0·47763. Autosomal analysis: log likelihood −8·935090, carrier risk 0·66667.

The autosomal likelihood must be corrected for

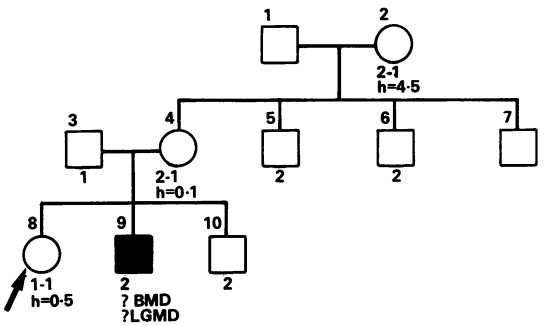

FIGURE *Example pedigree. Marker types are for a probe linked to BMD with 5% recombination. Creatine kinase h values are odds (h:1) of normal:BMD carrier.*

TABLE 1 *Pedigree in fig 1 coded for input to the linkage program. Sex: 1=M, 2=F. Affection status: 1=unaffected, 2=affected. Marker type: 1 0=allele 1 only, 0 1=allele 2 only, 1 1=both alleles, 0 0=no data.*

Subject	*Father*	*Mother*	*Sex*	*Affection status*	*Liability class*	*Marker type*
01	00	00	1	1	1	0 0
02	00	00	2	2	2	1 1
03	00	00	1	1	1	1 0
04	01	02	2	2	3	1 1
05	01	02	1	1	1	0 1
06	01	02	1	1	1	0 1
07	01	02	1	1	1	0 0
08	03	04	2	2	4	1 0
09	03	04	1	2	1	0 1
10	03	04	1	1	1	0 1

TABLE 2 *Data files for finding the likelihood of the pedigree and carrier risk assuming the disease is (a) BMD and (b) LGMD.*

	BMD			*LGMD*		
Gene frequencies						
Locus 1 (disease)	0·9995, 0·0005			0·9933, 0·0067		
Locus 2 (marker)	0·5, 0·5			0·5, 0·5		
Mutation rate at disease locus	0·0000055			0·0		
Penetrances in females:						
Genotype	1–1	2–1	2–2	1–1	2–1	2–2
Liability class 1	0·0	0·0	1·0	0·0	0·0	1·0
Liability class 2	0·82	0·18	0·0	0·82	0·82	0·0
Liability class 3	0·09	0·91	0·0	0·09	0·09	0·0
Liability class 4	0·33	0·67	0·0	0·33	0·33	0·0
Recombination fraction		0·05			0·50	
Calculate risks?		yes			yes	

males with no parents in the pedigree as explained in Methods. There are two (Nos 1 and 3); the marker gene frequencies are 0·5, therefore the likelihood must be divided by $(0{\cdot}5)^2$, that is, multiplied by 4. The corrected log likelihood is therefore $-8{\cdot}935090+\log(4)=-8{\cdot}333030$.

The log likelihood ratio is $8{\cdot}822886-8{\cdot}833303=0{\cdot}489856$. The odds ratio is the antilog of $0{\cdot}489856=3{\cdot}09{:}1$ in favour of LGMD.

For subject 8, the risk that she carries BMD= (chance that the disease is BMD)×(chance that she is a carrier if the disease is BMD)=$1/4{\cdot}09\times0{\cdot}47763=0{\cdot}117$. Similarly, the risk that she carries LGMD is $3{\cdot}09/4{\cdot}09\times0{\cdot}66667=0{\cdot}504$.

Discussion

In the particular case of distinguishing BMD and LGMD, molecular genetics may come to the rescue. We[10 11] and others[12] have shown that dystrophin cDNA probes detect gene deletions in about 60% of men with BMD, and we have been able to use these probes to make definitive diagnoses of BMD in several men with ill defined neuromuscular disease. Immunological study of the dystrophin in muscle biopsy specimens[13] will probably allow unambiguous diagnosis of all BMD patients. However, the problem of X linked versus autosomal recessive inheritance will remain for many other diseases, and even for BMD where the affected male is dead or unavailable for testing.

The necessity of making two separate runs of the linkage program for the two genetic hypotheses increases the risk of error. A correct result depends on keeping all probabilities strictly comparable between the two runs. This can involve problems of three sorts. First, logical problems: for example, the affected person is bound to be male if he has BMD, but could have been male or female if he has LGMD. Should the resulting factor of 2 difference appear in the quoted likelihoods, given that we would not have asked the question in the first place had it been an affected female? (The answer is yes.) Second, genetic problems: for example, might CK be raised in carriers of LGMD, and what is one to make of the earlier claims that 40% of LGMD is sporadic? Third, computational problems: the need to recode some males as heterozygous and correct the likelihood for the autosomal calculation, as described above. Only the computational problems are specific to the linkage method; the other problems will arise whatever method of calculation is used.

There is little real value in being able to calculate odds to three decimal places, but it is often useful to know what the range of probabilities is in a particular pedigree, and how much it might be changed by, for example, DNA typing some unaffected brothers. For these purposes one needs a computer program which can be set up and then fed with a range of possibilities. Thus I suggest that the advantages of the linkage approach are:

(1) Generality: once the input conditions (the gene frequencies and mutation rates) are set up for a pair of diseases, the program can tackle any pedigree without readjustment. This is particularly useful for exploring the range of possible values in a situation where the true gene frequencies are uncertain.
(2) Reliability: there is much less risk of arithmetical slips.
(3) Accessibility: anyone who can use the LINKAGE package for lod score calculations or risk analysis can follow the methods given here and perform the analysis. Most active genetics departments contain at least one such person, whereas few contain anybody able and willing to undertake the conventional Bayesian calculation.

References

[1] Wolff G, Mueller CR, Grimm T. Benign muscular dystrophy: risk calculation in families with consanguinity. *J Med Genet* 1989;**26**:299–304.
[2] Grimm T. Genetic counselling in Becker type X-linked muscular dystrophy. II. Practical considerations. *Am J Med Genet* 1984;**18**:719–23.
[3] Young ID, Nugent Z, Grimm T. Autosomal recessive or sex linked recessive: a counselling dilemma. *J Med Genet* 1986;**23**: 32–4.
[4] Blank CE. Modern genetics and neuromuscular disorder. *Lancet* 1987;**i**:626–7.
[5] Ott J. Estimation of the recombination fraction in human pedigrees: efficient computation of the likelihood for human linkage studies. *Am J Hum Genet* 1974;**26**:588–94.
[6] Lathrop GM, Lalouel JM. Easy calculation of lod scores and genetic risks on small computers. *Am J Hum Genet* 1984;**36**: 460–5.
[7] Lathrop M. *Linkage analysis programs: user's guide version 0·6.* Salt Lake City: Howard Hughes Medical Institute, 1987.
[8] Grimm T. Genetic counselling in Becker type X linked muscular dystrophy. I. Theoretical considerations. *Am J Med Genet* 1984;**18**:713–8.
[9] Ott J. Documentation for Liped version 3, October 1976.
[10] Forrest SM, Smith TJ, Cross GC, *et al.* Effective strategy for prenatal prediction of Duchenne and Becker muscular dystrophy. *Lancet* 1987;**ii**:1294–7.
[11] Read AP, Mountford RC, Forrest SM, Kenwrick SJ, Davies KE, Harris R. Patterns of exon deletions in Duchenne and Becker muscular dystrophy. *Hum Genet* 1988;**80**:152–6.
[12] Wapenaar MC, Kievits T, Hart KA, *et al.* A deletion hot spot in the Duchenne muscular dystrophy gene. *Genomics* 1988;**2**: 101–8.
[13] Hoffman EP, Fischbeck KH, Brown RH, *et al.* Characterisation of dystrophin in muscular biopsy specimens from patients with Duchenne's or Becker's muscular dystrophy. *N Engl J Med* 1988;**318**:1363–8.

Correspondence to Dr A P Read, Department of Medical Genetics, St Mary's Hospital, Hathersage Road, Manchester M13 0JH.