Am. J. Hum. Genet. 66:2001–2004, 2000

Report

The Impact of Transmission-Ratio Distortion on Allele Sharing in Affected Sibling Pairs

Celia M. T. Greenwood[1] and Kenneth Morgan[1,2,3]

Departments of [1]Human Genetics and [2]Medicine, McGill University, and [3]The Montreal General Hospital Research Institute, Montreal

There have been recent reports of transmission-ratio distortion (TRD) or segregation distortion in families not selected for genetic disease. If TRD exists but is ignored, linkage studies searching for disease genes in affected relatives may be misinterpreted. We show that the identical-by-descent sharing patterns for affected sib pairs are strongly affected by TRD and, further, that the estimated statistical significance of a sib-pair linkage study may be extremely biased. However, we also show that, if TRD is suspected during the planning stage of a study, the planned sample size of the study needs to be increased by only a small amount to maintain the desired power.

Naumova et al. (1998) have suggested the existence, in Centre d'Étude du Polymorphisme Humain (CEPH) families that are not selected for genetic disease, of an X-linked locus at which there is increased probability of transmission of the grandpaternal allele, through the mother to a male child. In 62% of transmissions, the grandpaternal X-chromosome allele, rather than the grandmaternal allele, was transmitted to the males. Transmission-ratio distortion (TRD) has also been observed at several autosomal loci in mice and humans (e.g., see Herrmann et al. 1999; Paterson and Petronis 1999; Paterson et al. 1999; Pardo-Manuel de Villena et al. 2000). Autosomal TRD in humans has been observed, by Eaves et al. (1999), at the INS-IGF2 VNTR, although in this case a specific allele is preferentially transmitted, and these authors have pointed out that TRD will alter the statistical significance of allele sharing in linkage studies of affected sib pairs. Excess allele sharing in affected sib pairs may indicate not a shared disease susceptibility but, rather, sharing due to deviation from Mendelian inheritance patterns.

We can evaluate the impact of X-linked TRD, such as that observed by Naumova et al. (1998), on the expected allele sharing between pairs of male sibs affected with the same disease. It is a function of the biased probability of allele transmission and the disease risk associated with the locus under study. Let τ be the probability that the grandpaternal (the mother's father's) allele on the X chromosome is transmitted to a male child. Under Mendelian inheritance, $\tau = .5$. In the presence of TRD, the probability that two male sibs share a gene identical by descent (IBD) on the X chromosome is $P(\text{IBD} = 1) = \tau^2 + (1 - \tau)^2$. The expected allele sharing between affected male sibs, for a locus on the X chromosome, can be expressed as a function of the increased risk of disease to a brother of an affected proband (Risch 1990*b;* Hallmeyer et al. 1996). Assume that a disease gene is linked, with a recombination fraction of 0, to the marker under study and that there is linkage equilibrium. Let z_1 be the proportion of brother pairs who are expected to share an X-linked allele IBD, given that both are affected with the same disease. Then, $z_1 = 1 - \{[2\tau(1 - \tau)]/\lambda_{XS}\}$, where λ_{XS} is the locus-specific relative risk of disease for a male sib. Figure 1 shows values of the expected allele sharing for affected brothers (z_1), as a function of τ and λ_{XS}. Even when the locus has no effect on disease risk ($\lambda_{XS} = 1$), the expected allele sharing will, in the presence of TRD, be biased away from .5. If $\lambda_{XS} = 1$ but $\tau = .62$, the expected sharing increases to .53. Similarly, an observed sharing of $z_1 = .60$ could correspond to a locus conferring an excess risk of $\lambda_{XS} = 1.28$ or to a locus that undergoes distortion ($\tau = .72$) but that has no effect on disease risk. Hence, unknown TRD can lead to incorrect inferences in sib-pair linkage studies.

Received January 27, 2000; accepted for publication March 14, 2000; electronically published April 20, 2000.

Address for correspondence and reprints: Dr. Celia M. T. Greenwood, Division of Medical Genetics, L10-109, Montreal General Hospital, 1650 Cedar Avenue, Montreal QC H3G 1A4, Canada. E-mail: celia@bagel.epi.mcgill.ca

0002-9297/2000/6606-0032$02.00

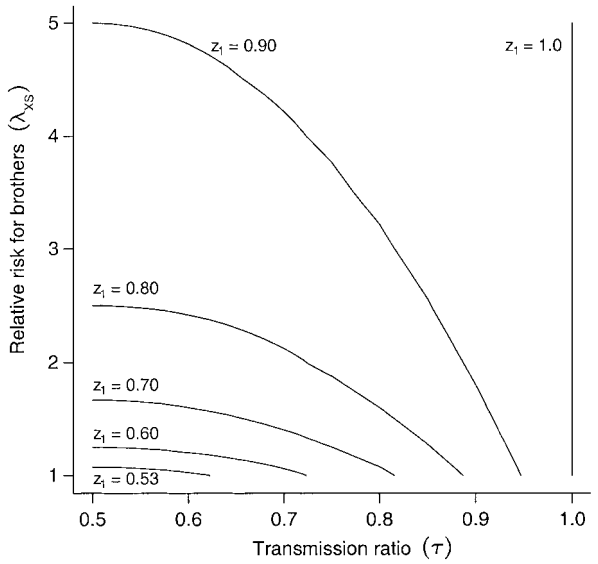

Figure 1 Contours of z_1 at an X-linked locus for affected brother pairs, as a function of the transmission ratio parameter, τ, and λ_{XS}.

Tests for linkage can be modified to take TRD into account. A simple test for X linkage in an affected-sib-pair study compares z_1 with the value expected when there is no linkage. Under the assumption of the normal approximation to the binomial, figure 2 shows that the estimated significance can be drastically biased if TRD is ignored. (There will usually be sufficient sib pairs in a linkage study that the normal approximation is adequate.) The ratio of the *P* value under the correct model to the *P* value that ignores TRD is shown as a function of τ, sample size *N*, and λ_{XS}, and the ratio increases rapidly with τ and *N*. For 30 brother pairs, if $\tau = .62$ and $\lambda = 2$, the true *P* value will be 2.6 times larger than a reported *P* value that ignores TRD. When $N = 100$, the ratio rises to 19.1.

For sib pairs with type 1 diabetes and a *DR3/DRX* genotype (where *DRX* is not *DR4*) at the MHC locus, Cucca et al. (1998) observed excess allele sharing at *DXS1068*, which is the same marker observed, by Naumova et al. (1998), as showing TRD in CEPH families. The diabetic brother pairs showed the most sharing (observed allele sharing $\hat{z}_1 = .74$; Cucca et al. 1998). In the absence of TRD, such allele sharing would be consistent with an excess risk of diabetes, $\lambda_{XS} =\sim 1.9$. On the other hand, if $\tau = .62$, then this amount of sharing is compatible with a locus conferring a risk of 1.8. Although the change in λ_{XS} is small, the statistical significance of such a pattern of sharing would have been overstated. Cucca et al. (1998) do not give the number of brother pairs, but, if it is assumed that 25% of the 97 sib pairs are of this type (a conservative estimate, given the excess of male diabetics in the sample), then the estimated *P* value would be .0142 for $\tau = .62$, versus a *P* value of .00677 in the absence of TRD ($\tau = .5$).

For a chosen value of λ_{XS}, it can be seen, in figure 1, that the allele sharing changes slowly with TRD up to moderate values of τ (i.e. , <.7). For example, for a locus conferring a twofold risk, z_1 will only increase from 75%, in the absence of TRD, to 76.4%, when $\tau = .62$. In fact, TRD has only a moderate effect on the power of a linkage study designed to detect a locus with a fixed value of λ_{XS}.

The power to detect linkage, by use of large sample approximations for the binomial distribution, can be expressed as

$$1-\beta = P\left[Z > \frac{c_\alpha\sqrt{\mathrm{Var}_0(z_1)} + E_0(z_1) - E_A(z_1)}{\sqrt{\mathrm{Var}_A(z_1)}}\right]$$

$$= P\left\{Z > \frac{[\lambda_{XS}c_\alpha\sqrt{2\tau(1-\tau)(1-2\tau+2\tau^2)} + 2\sqrt{n}(1-\lambda_{XS})\tau(1-\tau)]}{\sqrt{2\tau(1-\tau)[\lambda_{XS}-2\tau(1-\tau)]}}\right\}, \quad (1)$$

where *Z* is a standard normal random variable and where the subscripts 0 and A refer to expectations taken in the presence of TRD, under the null hypothesis of no linkage and the alternative hypothesis of linkage, respectively. The quantity c_α is the upper quantile of the

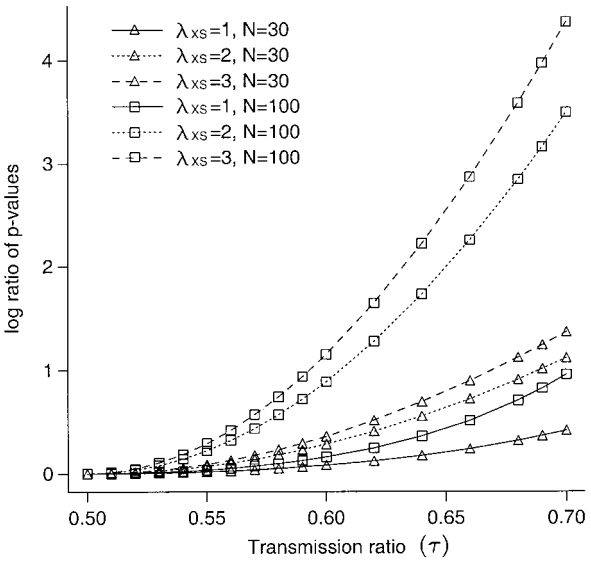

Figure 2 Inflation of significance levels when TRD is ignored, as a function of the λ_{XS}, the number of affected brother pairs, *N*, and the transmission ratio parameter τ. Vertical axis is $\log_{10}$ of the ratio of the true *P* value (which takes TRD into account) to the *P* value that assumes no TRD ($\tau = .5$). τ is shown only in the range .5–.7, to magnify the scale of the vertical axis.

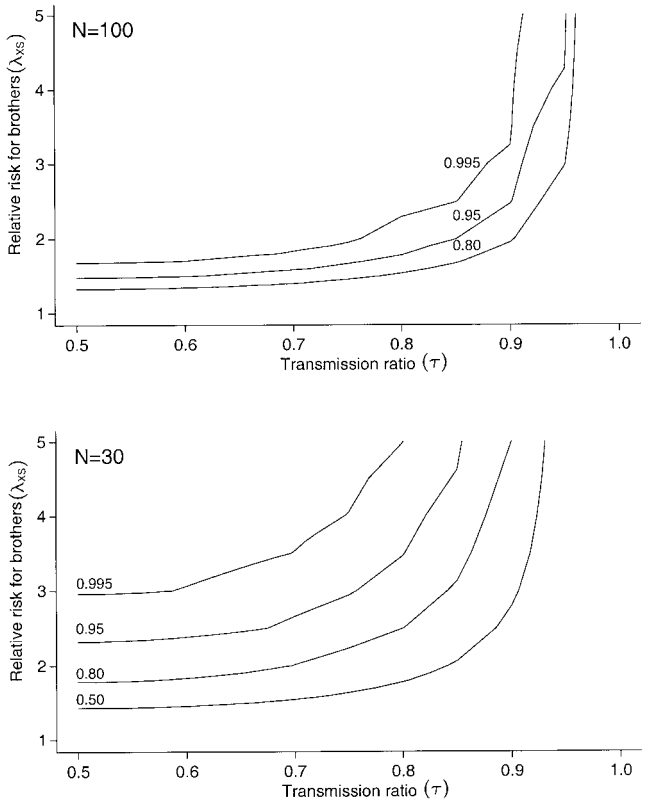

Figure 3 Contours of expected power, $1 - \beta$, for tests of X-chromosome linkage that use N affected brother pairs (see eq. [1]). Type I error $\alpha = .05$.

normal distribution associated with a type I error of α for a one-tailed test, and β is the type II–error rate.

Figure 3 shows power contours for linkage tests based on affected brother pairs, and it can be seen that power decreases as TRD increases, for a fixed value of λ_{XS}. However, the power loss is small except for large values of TRD. For a sample of 30 affected brother pairs, for example, the power to detect a locus conferring a risk of 2.0 when $\tau \leq .62$ is >87%; however, if $\tau \geq .75$, then the power drops rapidly. Equation (1) can be easily inverted, to estimate the required sample size for a sib-pair linkage study, while taking into account the possibility of TRD. If $\tau = .62$ then adding 11% more sib-pairs will guarantee enough power. If τ is as high as .7, then the sample size will need to be increased by one-third, and the sample-size inflation increases rapidly for values of $\tau > .7$. These inflation factors do not depend on either the value of λ_{XS} or on the chosen type I error or power.

Similarly, the effect of autosomal TRD on affected-sib-pair linkage studies can be examined. Let τ_{pat} represent the probability that the grandpaternal allele at an autosomal locus is transmitted from a father to a child, and let τ_{mat} represent the probability that the grandpaternal allele is transmitted from a mother to a child. Let z_j, $j = 0, 1, 2$, be the expected proportion of affected sib pairs who share j alleles IBD. At an autosomal locus, if $\tau_{pat} = \tau_{mat}$, then the average sharing, $z_2 + z_1/2$, is the same as that for the X chromosome when the putative disease locus has no effect on risk. Let λ_O be the relative risk of disease to a parent or offspring of the proband, and let λ_M be the relative risk for an MZ twin. Then the allele sharing proportions are expected to be

$$z_0 = \frac{P(\text{IBD} = 0)}{P(\text{IBD} = 0) + \lambda_O P(\text{IBD} = 1) + \lambda_M P(\text{IBD} = 2)} ,$$

$$z_1 = \frac{\lambda_O P(\text{IBD} = 1)}{P(\text{IBD} = 0) + \lambda_O P(\text{IBD} = 1) + \lambda_M P(\text{IBD} = 2)} ,$$

$$z_2 = 1 - z_0 - z_1 .$$

Under an additive model (Risch 1990*a*), $\lambda_M - 1 = 2(\lambda_O - 1)$; figure 4 shows the average sharing $z_2 + z_1/2$, expected with TRD under this model. Allele sharing is more sensitive to TRD than in figure 1; however, as in the X-chromosome case, the power of a study design is reduced by only a small amount when $\tau < .7$ (results not shown).

What are plausible levels of TRD in humans? Although almost complete distortion ($\tau \geq .95$) has been observed (e.g., Silver 1993) in mice and *Drosophila,* in humans such extreme distortion has not been observed.

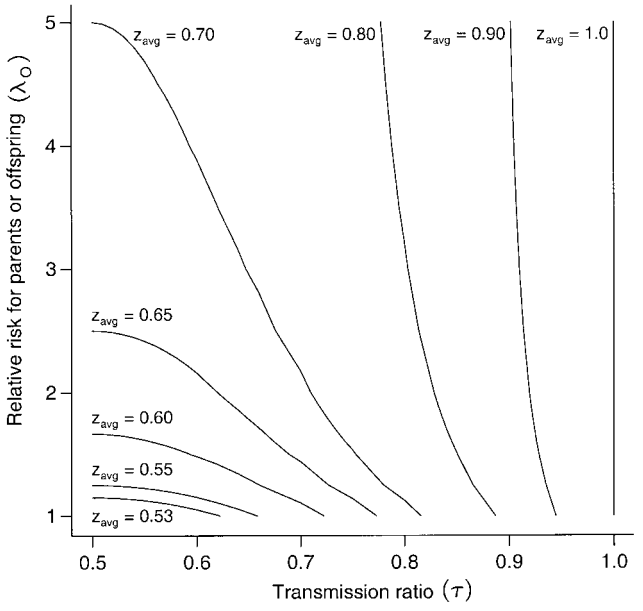

Figure 4 Contours of expected allele sharing, $z_{avg} = z_2 + z_1/2$, for affected sib pairs, for an autosomal gene without sex-specific effects, under a model with no dominance variance.

Naumova et al. (1998) observed $\tau = .62$ in CEPH families; Paterson and Petronis (1999) reported excess sharing of 63% IBD in diabetic sister pairs at a locus that shows excess female sharing in CEPH families (also see Zavattari et al. 2000). Such IBD sharing would correspond to $\tau \approx .75$ if there were no linked disease locus. Modest levels of distortion can result from selection due to meiotic drive (Pardo-Manuel de Villena et al. 2000) rather than to embryonic lethality. If maternal meiotic drive is the mechanism that leads to TRD, then TRD will rarely be $\geqslant$75% (Pardo-Manuel de Villena et al. 2000).

In summary, if TRD exists but is ignored, the statistical significance of a linkage study can be substantially overstated (Eaves et al. 1999; Paterson and Petronis 1999). However, TRD can be taken into consideration for sample-size calculations for sib-pair studies, and assuming a value of $\tau < .7$ will lead to only a small increase in the required sample size. These two conclusions may appear to be contradictory. However, when a study is designed, calculations will be performed for a fixed value of the relative risk, λ, whereas ignoring TRD during an analysis is equivalent to estimating an λ value that is much too large.

If the marker locus is in linkage disequilibrium with the susceptibility locus, or if a candidate gene is being studied, then TRD may increase or decrease the statistical significance of allele sharing. For example, Eaves et al. (1999) showed that, in CEPH families, class III alleles of INS-IGF2 VNTR are less often transmitted from heterozygous class I/III parents to offspring (44%–49% are expected to receive the class III allele). However, offspring with polycystic ovary syndrome (PCOS) are more likely to receive a class III allele from a heterozygous father. The PCOS risk associated with the INS-IGF2 VNTR was underestimated when the segregation distortion was ignored.

It has been shown, in the absence of either TRD or segregation distortion, that to initially genotype only the affected sibs of nuclear families is the most cost-effective strategy for sib-pair studies (Hauser et al. 1996), and this design has become popular. Nevertheless, the ascertainment of unaffected sibs makes it possible to test for segregation distortion (Spielman et al. 1993). To assess whether TRD may be a confounding factor in disease-mapping studies, transmission patterns from grandparents to unaffected grandchildren need to be examined. For example, if grandparents are available, then the sib-pair sharing could be examined in subgroups defined by the grandparental origin of the alleles. It would be worthwhile to collect three-generation families unselected for disease, so that TRD could be evaluated across the human genome.

Acknowledgments

We thank Anna Naumova for helpful discussions. This research was partially supported by grants from the Canadian Networks of Centres of Excellence program (the Canadian Genetic Diseases Network and the Mathematics of Information Technology and Complex Systems). C.M.T.G. is supported by a Medical Research Council of Canada Postdoctoral Fellowship.

References

Cucca F, Goy JV, Kawaguchi Y, Esposito L, Merriman ME, Wilson AJ, Cordell HJ, et al (1998) A male-female bias in type 1 diabetes and linkage to chromosome Xp in MHC HLA-DR3-positive patients. Nat Genet 19:301–302

Eaves EA, Bennett ST, Forster P, Ferber KM, Ehrmann D, Wilson AJ, Bhattacharyya S, et al (1999) Transmission ratio distortion at the INS-IGF2 VNTR. Nat Genet 22:324–325

Hallmeyer J, Hebert JM, Spiker D, Lotspeich L, McMahon WM, Petersen PB, Nicholas P, et al (1996) Autism and the X chromosome: multipoint sib-pair analysis. Arch Gen Psychiatry 53:985–989

Hauser ER, Boehnke M, Guo S-W, Risch N (1996) Affected-sib-pair interval mapping and exclusion for complex genetic traits: sampling considerations. Genet Epidemiol 13: 117–137

Herrmann BG, Koschorz B, Wertz K, McLaughlin KJ, Kispert A (1999) A protein kinase encoded by the *t complex responder* gene causes non-mendelian inheritance. Nature 402:141–146

Naumova AK, Leppert M, Barker DF, Morgan K, Sapienza C (1998) Parental origin-dependent, male offspring-specific transmission-ratio distortion at loci on the human X chromosome. Am J Hum Genet 62:1493–1499

Pardo-Manuel de Villena F, de la Casa-Esperon E, Briscoe TL, Sapienza C (2000) A genetic test to determine the origin of maternal transmission distortion: meiotic drive at the mouse *Om* locus. Genetics 154:333–342

Paterson AD, Naimark DM, Petronis A (1999) The analysis of parental origin of alleles may detect susceptibility loci for complex disorders. Hum Hered 49:197–204

Paterson AD, Petronis A (1999) Transmission ratio distortion in females on chromosome 10p11-p15. Am J Med Genet 88:657–661

Risch N (1990*a*) Linkage strategies for genetically complex traits. I. Multilocus models. Am J Hum Genet 46:222–228

——— (1990*b*) Linkage strategies for genetically complex traits. II. The power of affected relative pairs. Am J Hum Genet 46:229–241

Silver LM (1993) The peculiar journey of a selfish chromosome: mouse *t* haplotypes and meotic drive. Trends Genet 9:250–254

Spielman RS, McGinnis RE, Ewens WJ (1993) Transmission test for linkage disequilibrium: the insulin gene region and insulin-dependent diabetes mellitus (IDDM). Am J Hum Genet 52:506–516

Zavattari P, Esposito L, Nutland S, Todd JA, Cucca C (2000) Transmission-ratio distorion at Xp11.4-p21.1 in type 1 diabetes. Am J Hum Genet 66:330–332