

Erratum: Broman KW. 2001. Estimation of Allele Frequencies With Data on Sibships. *Genet Epidemiol* 20:307–15.

Karl W. Broman*

Department of Biostatistics, Johns Hopkins University, Baltimore Maryland

In the April 2001 issue of *Genetic Epidemiology*, in the article “Estimation of Allele Frequencies With Data on Sibships,” by Broman (20:307–15), there is an error on page 310, in the second paragraph under “Method 3: Accounting for Relationships.” The stated probabilities that an allele in the second sibling is not identical by descent (IBD) with one of the first sibling’s alleles, written as $p_l/(1 + p_l)$, are incorrect; we had missed two important cases. Let (g_{11}, g_{12}) denote the two alleles of the genotype of the first sibling, (g_{21}, g_{22}) denote the two alleles of the genotype of the second sibling, and $\mathbf{g} = (g_{11}, g_{12}, g_{21}, g_{22})$. Further, let \mathbf{pg} denote the genotypes for the two parents, and A denote the event “ g_{21} is not IBD with g_{11} or g_{12} .” We seek $\Pr(A|\mathbf{g})$, which we calculate by conditioning on the parents’ genotypes, as follows:

$$\Pr(A|\mathbf{g}) = \frac{\sum_{\mathbf{pg}} \Pr(\mathbf{pg}) \Pr(\mathbf{g}|\mathbf{pg}) \Pr(A|\mathbf{g}, \mathbf{pg})}{\sum_{\mathbf{pg}} \Pr(\mathbf{pg}) \Pr(\mathbf{g}|\mathbf{pg})}$$

The correct probabilities appear in Table I.

As a result of this error, the numbers in Table I of Broman (2001) were slightly wrong. We have rerun our computer simulations, after modifying our algorithm using the corrected probabilities, to obtain the true maximum likelihood estimates (MLEs). The corrected version of the table appears in Table II.

*Correspondence to: Karl W. Broman, Department of Biostatistics, Johns Hopkins University, 615 N. Wolfe St., Baltimore MD 21205. E-mail: kbroman@jhsph.edu

Received for publication 10 July 2002; Revision accepted 11 July 2002

Published online in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/gepi.10194

TABLE I. Probability That an Allele in the Second Sibling Is Not IBD With Either Allele in the First Sibling

Sibs' genotypes			
g_{11}	g_{12}	g_{21}	g_{22}
		Pr(g_{21} not IBD with g_{11} or $g_{21} g$)	
		Pr(g_{22} not IBD with g_{11} or $g_{21} g$)	
11	11	$p_1/(1 + p_1)$	$p_1/(1 + p_1)$
11	22	1	1
11	12	$p_1/(1 + p_1)$	1
11	23	1	1
12	12	$(p_1 + 2p_1p_2)/(1 + p_1 + p_2 + 2p_1p_2)$	$(p_2 + 2p_1p_2)/(1 + p_1 + p_2 + 2p_1p_2)$
12	13	$2p_1/(1 + 2p_1)$	1
12	34	1	1

TABLE II. Corrected Version of Table I in Broman (2001)

Het	p			
	0.05	0.10	0.15	0.20
0.70	1.46	1.44	1.42	1.39
0.75	1.47	1.45	1.42	1.39
0.80	1.47	1.43	1.42	1.40
0.85	1.48	1.43	1.42	1.40
0.90	1.48	1.46	1.43	1.42

It is interesting to note that the numbers in the corrected table are somewhat smaller than those in the original version. It appears that the true MLEs have somewhat greater standard deviations (SDs) than our flawed algorithm, especially for the larger values of the allele frequency, p . While our original estimates exhibit a slight negative bias (-0.001 in the case of $p = 0.2$ with heterozygosity = 0.9), they have smaller SDs and thus smaller mean square errors than the true MLEs. Note that the true MLEs appear to be unbiased.

We thank Mary Sara McPeck for identifying this error.