Reanalysis of genotype distributions published in *Neurology* between 1999 and 2002

István Kocsis, MD, PhD; Barna Vásárhelyi, MD, PhD; András Györffy, VetD; and Balázs Györffy, MD

Abstract—The authors tested 123 genotypes described in 54 papers published in the journal Neurology between 1999 and 2002 to ascertain whether these genotype distributions deviated from Hardy–Weinberg equilibrium (HWE). Unreported deviations from HWE in 19 genotype distributions described in 11 of the papers were discovered. The authors also report additional information that could have been extracted after calculating HWE and conclude that HWE values should be mandatory in population genetic studies published in *Neurology*.

NEUROLOGY 2004;63:357–358

The gene frequencies at a locus in a randomly interbreeding diploid population and population genotype frequencies remain constant from generation to generation if mating is random and if mutation, selection, and migration do not occur. A mathematical formula created by Hardy and Weinberg[1-4] allows for this equilibrium to be tested in studied populations.

The Hardy–Weinberg law is suitable for testing the hypothesis of panmixia (random mating within a breeding population) and evolutionary stasis. It is a null hypothesis in genetic studies for testing the suitability of the enrolled reference population and for detecting even weak associations between genotype and disease. However, Hardy–Weinberg equilibrium (HWE) is not always calculated and published in papers investigating population genetics, as we found after reviewing a sample of papers published in recent volumes of *Neurology*.

Subjects and methods. We reanalyzed genotypes described in papers published in *Neurology* between 1999 and 2002. Enrollment criteria were investigation of biallelic genetic polymorphisms with Mendelian inheritance and case-control study design. We calculated HWE values for each published genotype in each subject group using the Arlequin program (http://anthropologie.unige.ch/arlequin/).[6,7] The level of significance was set at $p < 0.05$.

Results. After recalculating distributions of 123 genotypes described in 54 papers, we found 19 genotypes deviated from HWE (table). In none of the cases did the authors mention that the HWE p value was under 0.05. Significant deviation from Hardy–Weinberg expectation occurred in either control or investigated populations in three papers, in healthy reference genotypes in four papers, and in diseased genotypes in five papers.

Discussion. Genotype frequencies may deviate significantly from the expected values for several reasons. Deviations should appear in 1 of every 20 genotypes, if the level of significance is set at $p = 0.05$. However, regardless of the exact mechanism used to determine deviations, the fact that the distribution of the studied genotype deviates from HWE should be mentioned by authors. Another frequent, although often unreported, cause for deviations is methodologic bias. The applied method might, for example, not allow for adequate discrimination between hetero- and homozygotes, potentially leading to the alteration of genotype distribution. One or more of the model's assumptions may also be incorrect. Nonrandom mating (inbreeding or an allele effect on the mating) or gene flow may be occurring, selection may be operating, or something else evolutionarily interesting could be happening. Deviations could also be the result of a sampling error. The sample size could be too small, it might not be from a single population, or genotypes could have a different likelihood of being included in the sample.[5] Taking these possibilities into consideration, authors should always calculate HW values.

We also found articles in which the populations being investigated deviated from HWE, but this fact was unreported. This would appear to discount the associations, or the lack thereof, put forward in several of the papers we sampled in our study. These include the association between the *GSTM3 AA* genotype and disability in patients with multiple sclerosis and between *APOE* promoter (−219 G/T) and Alzheimer disease (AD) in Finnish populations as well as the lack of an association between *NOS3* Glu/Asp polymorphism and ischemic stroke, between sporadic inclusion body myositis and PrP codon 129 methionine homozygosity, between *CST3* +73 *G/A*

From the First Department of Pediatrics (Dr. Kocsis) and Second Department of Internal Medicine (Drs. A. Györffy and B. Györffy), Semmelweis University, and Research Laboratory for Pediatrics and Nephrology (Dr. Vásárhelyi), Hungarian Academy of Sciences and Semmelweis University, Budapest, Hungary; and Charité (Dr. B. Györffy), Institute of Pathology, Humboldt University, Berlin, Germany.

Received September 23, 2003. Accepted in final form March 8, 2004.

Address correspondence and reprint requests to Dr. István Kocsis, First Department of Pediatrics, Semmelweis University, Budapest, Hungary; e-mail: kopist@gyer1.sote.hu

Table *Papers in* Neurology *between 1999 and 2002 in which deviations from Hardy–Weinberg equilibrium (HWE) were not noted*

Vol: page nos	Genetic polymorphism	Affected population	Reported			n	Expected			HWE *p* value
			AA	*AB*	*BB*		*AA*	*AB*	*BB*	
53:418–420	*NOS3* gene (Glu/Asp)	Control	154	203	36	393	166	179	48	0.0079
54:552–557	*GSTM3*	Control	221	64	15	300	213	79	8	0.0015
55:1235	Prion protein gene (Met/Val)	Controls	156	44	198	398	79	197	122	0.0001
		Controls	129	25	146	300	67	149	84	0.0001
		Patients	21	1	19	41	11	21	9	0.0001
		Patients	14	1	7	22	10	10	2	0.0001
56:1593–1595	*ACE I/D*	Patients	120	239	74	433	132	214	87	0.020
57:337–339	*CST3*−157 *G/C*	Patients	140	33	6	179	137	39	3	0.0387
	CST3+73 *G/A*	Controls	180	40	8	228	175	49	4	0.0091
58:124–126	*D1.1*	Controls	16	32	42	90	12	41	37	0.0393
	D1.8	Controls	17	7	5	29	14	12	3	0.0343
58:881–884	*Nurr1* 7048G7049	Patients	162	48	15	225	154	64	7	0.0049
58:1566–1568	IL-1B (−511)	Controls	95	107	15	217	101	94	22	0.0380
59:59–66	*APOE* promoter (−491)	Patients	94	18	4	116	91	23	2	0.0312
	APOE promoter (−219)	Controls	67	102	19	188	74	88	26	0.0429
		Patients	57	118	100	275	49	134	92	0.0477
		Patients	120	361	194	675	134	333	208	0.0444
59:724–728	*OPRM1*	Patients	193	29	8	230	187	41	2	0.0002
59:756–758	*A2M-Ile1000Val*	Patients	61	48	23	132	55	60	17	0.0244

polymorphism and AD in Japanese patients, between dopamine receptor 1.1 or 1.8 and blepharospasm, and between interleukin-1B (−511) polymorphism and AD (see the table).

We found low HWE values in patient populations in studies investigating *ACE D* allele in dementia, *Nurr1* gene in subjects with Parkinson disease, *APOE* promoter −491 *AA* genotype in AD in Spanish subjects, *APOE* promoter −219 *TT* genotype in AD in English and French subjects, μ-opioid receptor subunit gene −172 *TT* genotype in subjects with idiopathic generalized epilepsy α_2-macroglobulin, and *Val1000* allele in subjects with AD. Although the authors found significant associations in these studies, the deviations of genotype distribution in the patient populations would have been further evidence for the observed associations. This was also the case when the authors of a paper from the sample found no association between the cystatin C promoter and increased risk of AD. In spite of the lack of significant difference between allele or genotype frequency of control and patient groups, there could be some association between this genetic variant and increased risk for the disease, according to the low HWE value.

These cases suggest that authors should always calculate HWE and consider the results when evaluating population genetic studies investigating the significance of a genetic polymorphism of Mendelian inheritance in human disease.

References

1. Hardy GH. Mendelian proportions in a mixed population. Science 1908; 28:49–50.
2. Weinberg W. Über den Nachweis der Vererbung beim Menchen. Jahresh Verein f Vaterl Naturk in Wüttemberg 1908;64:368–382.
3. Roughgarden J. Theory of population genetics and evolutionary ecology. Englewood Cliffs: Prentice Hall, 1996.
4. Stern C. The Hardy–Weinberg law. Science 1943;97:137–138.
5. Hedrick PW. Genetics of populations. New York: Van Nostrand Reinhold, 1983.
6. Schneider S, Roessli D, Excoffier L. Arlequin ver. 2.000: a software for population genetics data analysis. University of Geneva, Switzerland: Genetics and Biometry Laboratory, 2000.
7. Guo SW, Thompson EA. Performing the exact test of Hardy–Weinberg proportion for multiple alleles. Biometrics 1992;48:361–372.