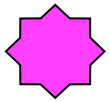
Mutations 35delG and M34T of *GJB2* gene among children with early onset hearing loss in Estonia

Variations on statistical topics of a paper submitted to *Laryngoscope*



Presented by T. Möls on BI JC, 2th February 2010

Mutations 35delG and M34T of *GJB2* gene among children with early onset hearing loss in Estonia

Rita Teek [1,2]

Katrin Kruustük [3]
Riina Žordania [4]
Kairit Joost [4]
Tiia Reimand [1,2]

Tõnu Mõls [6]

Eneli Oitmaa [5]
Maarja Kõiv [1]
Tiina Kahre[1,2]
Neeme Tõnisson [1]

Mart Kull (†) [3]

Katrin Õunap [1,2]

Referred patients + wrote manuscript

Referred patients

Statistics

Molecular studies

Supervised study in Dept. of Ear Clinic

Supervised study in Dept. of Genetics, corresponding author

[1] Dept. of Genetics, United Laboratories, UT Hospital

[2] Dept. of Pediatrics, UT

[3] Ear Clinic, UT Hospital

[4] Tallinn Children's Hospital

[5] Asper Biotech, Tartu

[6] Dept. of Bioinformatics, Inst. of Cell and Molecular Biology, UT

(Submitted to **Laryngoscope** 2009)

Background

Hearing loss is the most common sensory disorder worldwide.

Approximately one to two children in 1000 (0.1 – 0.2 %) are born with hearing loss.

Etiology of deafness is heterogeneous

> 60% preasumable genetic, in most cases monogenic
(100 mapped loci and 46 causally implicated genes)

> Mostly autosomal-recessive genes

> Of all pre-lingual, non-syndromic, recessive deafness cases about 50% can be related to *GJB2* gene mutations
(*GJB2* has single coding exon, protein belongs to connexins implicated in cap-junctional intercellular communication).

Purpose of study

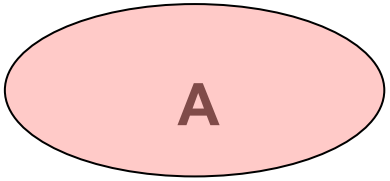
To find out the distribution and expressivity of 35delG and M34T mutations in *GJB2* gene among newborns and children with early onset hearing loss in Estonia.

Purpose of the presentation

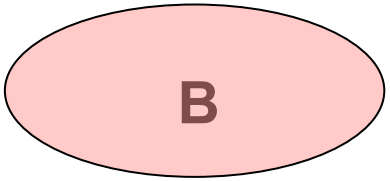
To show some statistical methods used in analysis of distribution and expressivity of mutant alleles.

The Data

Two datasets – **A and B –** were collected for study:



A screened cohort of 998 neonates born in one month: anonymous samples of dried blood spots on Guthrie cards separately for Northern, Western and Southeast Estonia.



A group of 233 probands who were referred to genetic counseling from January 2000 to March 2009 from the whole Estonia, with early onset hearing loss as a main complaint.

Plan of statistical analysis (1)

Task 1. Test if the probability of mutant allele (35delG or M34T) in the three parts of Estonia is equal (use data **A**).

> ***If the H_0 (equality of probabilities) can not be rejected, join the three datasets in a single one to characterize the whole Estonia and estimate on this the mutant alleles probabilities p_G and p_M.***

Task 2. Test if population of newborns (data **A**) is in Hardy-Weinberg equilibrium (HWE).

> ***If HWE can not be rejected, suppose that alleles are distributed randomly and independently. In particular, suppose that the lethality of embryos carrying mutant allele(s) equals to the population mean prenatal lethality.***

Plan of statistical analysis (continued)

Task 3. Test if the probability of allele 35delG (or allele M34T) among children with early onset hearing loss (data **B**) is equal to the allele probability, estimated from data **A** (test if the H_0 is true).

If H_0 is rejected and $p_B > p_A$, consider the mutant allele as a risk factor.

Task 4. Compare the observed frequencies of genotypes in patients (data **B**) with the expected (when using allele probability from data **A)** frequencies, assuming also HWE.

If the frequency of heterozygous genotype is significantly higher in B than expected from A, the mutant allele is not recessive.

Task 1

Can we analyse all three Estonian Regions together?

Dataset A – newborns analysed (1)

Regions	Newborns screened	35delG alleles		M34T alleles	
Northern	608	31		36	
Western	111	5		10	
South-Eastern	279	15		12	
Whole Estonia	998	51		58	

Dataset A – testing H_0: “probabilities of mutants in Regions are equal“

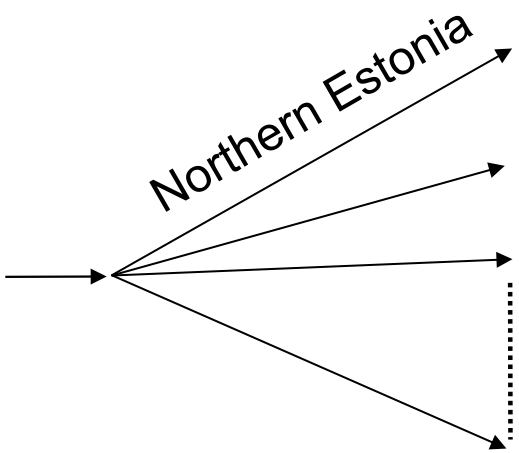
Regions	Newborns screened	35delG alleles	35delG probability	M34T alleles	M34T probability
Northern	608	31	0.0255	36	0.0296
Western	111	5	0.0225	10	0.0450
South-Eastern	279	15	0.0269	12	0.0215
Fisher’s test*, H_0=”probabilities equal“		P = 0.9773		P = 0.1981	

*Modified

Conclusion: No evidence, that probabilities differ between Regions. Data of three Regions can be joined.

Statistical details

There are several methods for testing, if a multinomially distributed vector of counts has the probability distribution stated by H_0. Most popular is the Pearson's χ^2.



Expected if H_0	Observed
31.07	31
5.67	5
14.26	15
	etc.

$$X^2 = \sum \frac{(\text{expected} - \text{observed})^2}{\text{expected}}$$

We have $\chi^2 = 0.1185$.
If H_0 is true, then this or greater value appears with probability of P=0.9425, so it is quite normal.

The expected frequencies were calculated as, for example: of 2×998=1996 screened alleles 2×608=1216 (=60.92%) were taken from Northern Estonia. Assuming uniform distribution of 51 G-allele over Estonia, 0.6092×51=31.07 M-allele are expected found in Northern Estonia.

Question: why in Table P = 0.9773, by Pearson's χ^2. P = 0.9425.

Dataset A – newborns analysed (2)
Estimating probabilities of 35delG and M34T:
$p_G = 51 / (2 \times 998) = 0.0256$; $p_M = 58 / (2 \times 998) = 0.0291$

Regions	Newborns screened	35delG alleles	35delG probability	M34T alleles	M34T probability
Northern	608	31	0.0255	36	0.0296
Western	111	5	0.0225	10	0.0450
South-Eastern	279	15	0.0269	12	0.0215
Whole Estonia	**998**	**51**	**0.0256**	**58**	**0.0291**

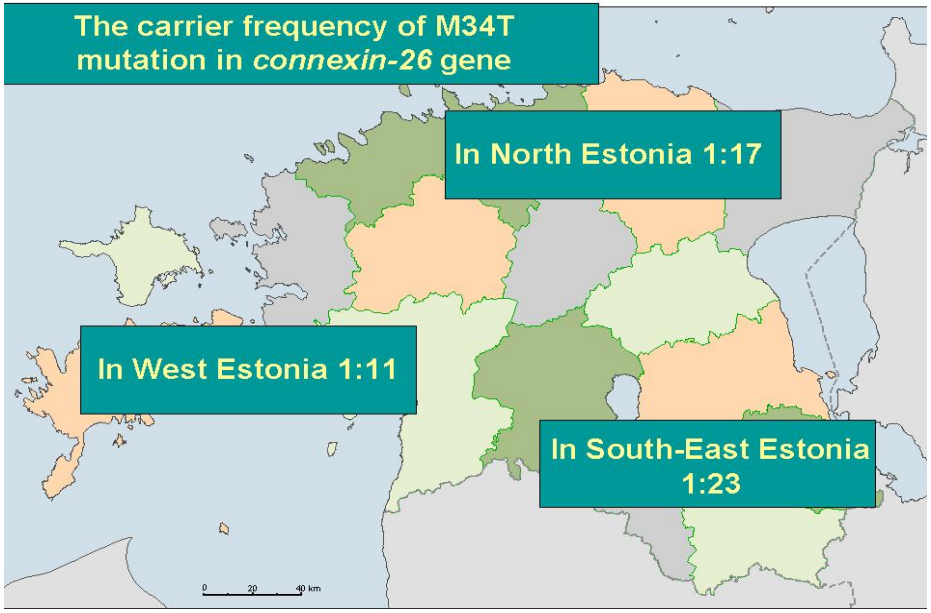
Conclusion: ML estimate for 35delG probability is $p_G = 0.0256$
ML estimate for M34T probability is $p_M = 0.0291$

Comments

Carrier frequency of mutations 35delG and M34T in *GJB2* gene in Europe

Country	35delG	M34T
Austria [30] [44]	1:60-1:110	-
Belgium [38]	1:190	-
Belgium		-
Bulgaria [38])	1:157	-
Czech Republic [38]	1:48.7	-
Denmark [38]	1:47.5	-
Finland [32] [40]	1:43-1:63	1:26
France (Brittany) [38]	1:96	-
France [38]	1:200	-
France [24]	1:66	1:43
Germany [38]	1:50	-
Greek [38]	1:33	-
Greek [9]	1:28	-
Holland [38]	1:44.5	-
Italy [38]	1:32	-
Malta [38]	1:36	-
Norway [38]	1:190	-
Portugal [38]	1:45	-
Slovenia [38]	1:182	-
Spain ([38]	1:40	-
Turkey [38]	1:37.5	-
United Kingdom [38]	1:119	-
British/Irish [35]		1:25.2
Estonia (this study)	**1:22**	**1:17**

The alleele probability p is in medical texts usually presented by carrier frequency, which has from statistical point of view a little bit fuzzy meaning, because homosygotes and heterozygotes are both carriers but have quite different meaning.



Task 2

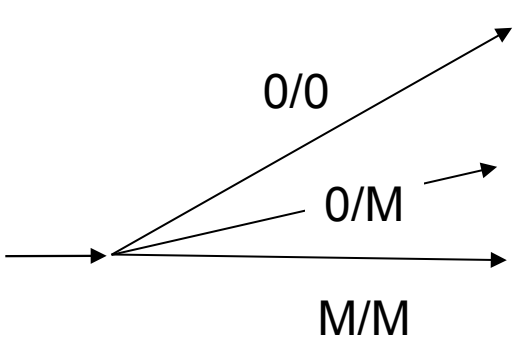
Is the population of newborns in HWE?

Testing HWE in newborns does not give evidence that the population is out of HWE with respect alleles 35delG and M34T:

Genotype	A = 35delG		A = M34T	
	Observed	Expected	Observed	Expected
0/0	949	947.7	940	940.8
0/A	47	49.7	58	56.3
A/A	2	0.65	0	0.84
P-value	P > 0.25		P > 0.40	

Conclusion. The childrens data show, that alleles 0 and M are probably transmitted independently and the sample is in HWE.

Statistical details



Expected if H_0	Observed	Possible, if H_0, combinations				
940.8	940	940	939	938	941	:
5.67	58	57	59	60	56	:
0.845	0	1	0	0	1	:

For testing the HWE, the Pearson's χ^2 can not be used, because it is an asymptotic method, which is relevant for large numbers of observations, in each table cell at least 5 or more expected or observed values. In the present example we have in one cell 0 observations. Therefore we use 'modified' Fisher's method.

We consider all possible components for multivariate vector of frequecies, having sum 998. Some of combinations are unrealistic, the other have higher probabiliti. Now we sum probabilities of all the possible combinations, which are equal or less than the observed combination. If the resulting sum is small, the H_0 must be rejected.

Tasks 3 and 4

How are the M- and G-alleles related with hearing loss?

Observed and expected, under HWE, **genotype frequencies** among newborns and patients. P-values are calculated using multinomial distribution with estimates for allele probabilities.

$\mathbf{P_1}$ is the probability that observed frequencies deviate from the expected ones by chance,

$\mathbf{P_2}$ is the sum of probabilities of the genotype distributions possible at random sampling and having probability P_1 or less (this is analogous to **Fisher's P**).

Expected2 frequencies that follow from HWE with M34T probability estimates from **A** data.

Expected3 and the related P_1 is calculated assuming that one M/M-embryo died before birth.

Analysis of M34T (= M) genotypes

Genotype	Frequency (patients)			Frequency (children)		
	Observed	Expected	Expected2	Observed	Expected	Expected3
0/0	213	206.8	219.6	940	940.8	939.9
0/M	13*	25.4	13.2	58	56.3	58.2
M/M	7	0.78	0.2	0	0.84	0.90
All	233	233	233	998	998	999
P_1	0.0000000417		$2 \cdot 10^{-10}$	0.0226		0.0197
P_2	0.0000000512			0.0226		

* Five of 13 0/M patients were G/M.

Observed and expected, under HWE, frequencies of genotypes among patients. Considering genotypes with 35delG allele, genotypes having M43T allele were omitted and, vice verse, M43T genotypes were analyzed without genotypes having 35delG. The expected frequencies were calculated using allele probabilities estimated from newborns sample.

Genotype	A = 35delG		A = M34T	
	Observed	Expected	Observed	Expected
0/0	118	202.3	118	125.4
0/A	22	10.6	8	7.5
A/A	73	0.14	7	0.11
All	213	213	133	133
P-value	≈0		≈0	

Genotype	A = 35delG		A = M34T	
	Observed	Expected	Observed	Expected
0/0	118	202.3	118	125.4
0/A	22	10.6	8	7.5
A/A	73	0.14	7	0.11

Conclusions. Among patients, the homozygote A/A is strongly more frequent than one can expect in case of HWE. This is even more evident, if to calculate the expected frequencies from children's data (see previous slide, Column Expected2).

It follows also, that

(1) M/M-individuals are subjected to phenotypic anomalies.

(2) M34T is a recessive allele (in heterozygous genotypes does not significantly increase the probability of phenotypic anomalies).

(3) G/G- and 0/G- individuals are subjected to phenotypic anomalies while 0/G- genotype is less damaged if compared to G/G.

Tanks for Your Attention !