Help File for My Variant Analysis

Contents

1 Background

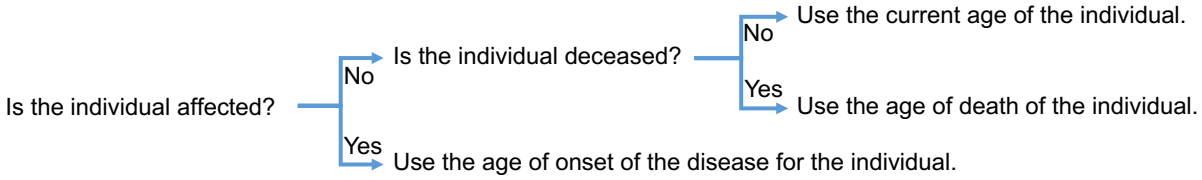
This website (`analyze.myvariant.org`) implements various methods for cosegregation analysis. Each method analyzes a pedigree and gives a value that can be used to determine whether a variant is pathogenic or benign with higher values indicating more pathogenic results. For more information please refer to the following author's papers. Note that the FLB part of the website is built upon the LINKAGE or FASTLINK program and the CSLR and meioses counting uses the CoSeg R package so references for those are included as well.

- CSLR: Mohammadi L, Vreeswijk MP, Oldenburg R, et al. A simple method for co-segregation analysis to evaluate the pathogenicity of unclassified variants; BRCA1 and BRCA2 as an example. *BMC Cancer*. 2009;9:211. `https://www.msbi.nl/cosegregation/default.aspx`
- FLB: Thompson D, Easton DF, Goldgar DE. A full-likelihood method for the evaluation of causality of sequence variants from family data. *Am J Hum Genet*. 2003;73(3):652-655.
- Meioses Counting: Jarvik GP, Browning BL. Consideration of Cosegregation in the Pathogenicity Classification of Genomic Variants. *Am J Hum Genet*. 2016;98(6):1077-1081.
- FLB(Fastlink) : Schäffer AA, Gupta SK, Shriram K, Cottingham RW. Avoiding recomputation in linkage analysis. *Hum Hered*. 1994;44(4):225-237.
- FLB(LINKAGE)Lathrop GM, Lalouel JM. Easy calculations of lod scores and genetic risks on small computers. *Am J Hum Genet*. 1984;36(2):460-465.
- CSLR/Meioses Counting and CoSeg R package: Ranola JM, Liu Q, Rosenthal E, and Shirts B. A comparison of cosegregation analysis methods for the clinical setting. *Familial Cancer*. 2017; In press.

2 File Formatting

The file can be uploaded as a plain text file or an excel file using the example ped file as a template. For the plain text file, the format of the pedigree file consists of 9 columns separated by a single space along with a header. Note that the header, or first line, can be anything as the webtool ignores it completely. For both formats the 9 columns in order are:

1. Individual ID - This column gives the unique number assigned to the current individual. Numbers are not assigned in any particular order.
2. Father - This column gives the Individual ID of the current individual's biological father or 0 if the individual's father is not in the pedigree. Note that both parents should be known or both should be unknown. Also, all individuals in the pedigree should be connected (i.e. If a pedigree has the proband's grandfather then the proband's father and mother need to be listed in the pedigree even if no other information about them is known. Likewise, siblings need to have their parents included in the file as that is how they are identified as siblings.).
3. Mother - This column gives the Individual ID of the current individual's biological mother or 0 if the individual's mother is not in the pedigree. Note that both parents should be known or both should be unknown. Also, all individuals in the pedigree should be connected (i.e. If a pedigree has the proband's grandfather then the proband's father and mother need to be listed in the pedigree even if no other information about them is known).
4. Sex - This column represents the individual's sex where 1 stands for male and 2 stands for female.
5. Affection Status - This column shows the individual's affection status where 0 stands for unknown, 1 stands for unaffected, and 2 stands for affected. Individuals should be marked as affected for a particular gene if they have any of the phenotypes listed below.
   (a) BRCA1 or BRCA2 - Breast cancer or ovarian cancer
   (b) MLH1 - Colon cancer, colon polyps, endometrial cancer, grouped small bowel cancer, stomach cancer, ovarian cancer, or urinary tract cancers
6. Age - This column represents one of 3 possible ages rounded to the nearest whole number. If the individual is affected it should be the age of onset of the disease. If the individual is unaffected, it should be the current age of the individual if alive or age at death. For more help, refer to the flow chart below. If the age is unknown, it is okay to estimate it based on a sibling or parent's age. A good rule of thumb is the individual is the same age as their siblings and 25 years younger than their parents.

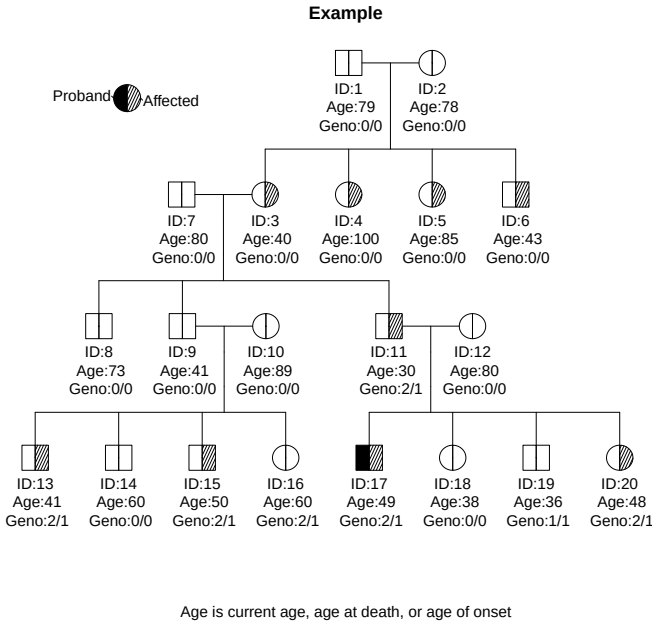


7. Allele1 - This column gives the status of the first allele of the genotype in question where 0 is unknown, 1 is normal (negative for the VUS), and 2 is the variant(positive for the VUS). Note that Alelle1 and Allele2 are interchangeable. Also both alleles should be known or unknown. (Normally we assume the individual has at only one copy of the VUS as having two would be extremely rare.)
8. Allele2 - This column gives the status of the second allele of the genotype in question where 0 is unknown, 1 is normal (negative for the VUS), and 2 is the variant(positive for the VUS). Note that Alelle1 and Allele2 are interchangeable. Also both alleles should be known or unknown. (Normally we assume the individual has at only one copy of the VUS as having two would be extremely rare.)
9. Proband - This column should be 0 for all members except the proband. Note that there should only be a single proband and that they should be a carrier of the variant and affected.

3 Example Pedigree

An example pedigree file is plotted below for reference.

Individual	Father	Mother	Sex	Affection	Age	Alelle1	Allele2	Proband
1	0	0	1	0	79	0	0	0
2	0	0	2	0	78	0	0	0
3	1	2	2	2	40	0	0	0
4	1	2	2	2	100	0	0	0
5	1	2	2	2	85	0	0	0
6	1	2	1	2	43	0	0	0
7	0	0	1	1	80	0	0	0
8	7	3	1	1	73	0	0	0
9	7	3	1	1	41	0	0	0
10	0	0	2	0	89	0	0	0
11	7	3	1	2	30	2	1	0
12	0	0	2	1	80	0	0	0
13	9	10	1	2	41	2	1	0
14	9	10	1	1	60	0	0	0
15	9	10	1	2	50	2	1	0
16	9	10	2	1	60	2	1	0
17	11	12	1	2	49	2	1	1
18	11	12	2	1	38	0	0	0
19	11	12	1	1	36	1	1	0
20	11	12	2	2	48	2	1	0



4 Penetrance Classes

Penetrance parameters for disease and SEER models were adapted from the published works listed below.

1. Howlader N, Noone AM, Krapcho M, Miller D, Bishop K, Altekruse SF, Kosary CL, Yu M, Ruhl J, Tatalovich Z and others. 2013. SEER Cancer Statistics Review, 1975-2013. SEER Cancer Statistics Review. Bethesda, MD: National Cancer Institute.

2. Quehenberger F, Vasen HF, van Houwelingen HC. 2005. Risk of colorectal and endometrial cancer for carriers of mutations of the hMLH1 and hMSH2 gene: correction for ascertainment. J Med Genet 42(6):491-6.

3. Wang F, Fang Q, Ge Z, Yu N, Xu S, Fan X. 2012. Common BRCA1 and BRCA2 mutations in breast cancer families: a meta-analysis from systematic review. Mol Biol Rep 39(3):2109-18.

4.1 ATM

	Female		Male	
Age Range	Homozygous Normal	Heterozygous	Homozygous Normal	Heterozygous
[0,20)	0	0	0	0
[20,30)	0.004	0.01	1.2e-06	7.2e-06
[30,40)	0.012	0.03	1.9e-05	0.000114
[40,50)	0.024	0.0602	8.5e-05	0.00051
[50,60)	0.052	0.13	0.00027	0.00162
[60,70)	0.086	0.215	0.00067	0.00402
[70,∞)	0.131	0.3283	0.0012	0.0072

4.2 BRCA1

	Female		Male	
Age Range	Homozygous Normal	Heterozygous	Homozygous Normal	Heterozygous
[0,20)	0	0	0	0
[20,30)	4.0997e-05	0.047524	1.2e-06	0.00017
[30,40)	0.00189916	0.18042	1.9e-05	0.0012
[40,50)	0.00878848	0.3736	8.5e-05	0.003
[50,60)	0.0275136	0.5752	0.00027	0.0062
[60,70)	0.05646	0.6889	0.00067	0.012
[70,∞)	0.0793	0.785	0.0012	0.018

4.3 BRCA2

	Female		Male	
Age Range	Homozygous Normal	Heterozygous	Homozygous Normal	Heterozygous
[0,20)	0	0	0	0
[20,30)	0.00040997	0.015985	1.2e-06	0.0018
[30,40)	0.00189916	0.12691	1.9e-05	0.012
[40,50)	0.00878848	0.2608	8.5e-05	0.027
[50,60)	0.0275136	0.4571	0.00027	0.047
[60,70)	0.05646	0.5554	0.00067	0.068
[70,∞)	0.0793	0.6449	0.0012	0.083

4.4 CHEK2

	Female		Male	
Age Range	Homozygous Normal	Heterozygous	Homozygous Normal	Heterozygous
[0,20)	0	0	0	0
[20,30)	5.2e-05	0.00093	0.0004457	0.0007991
[30,40)	0.00389	0.00281	0.0019968	0.0035803
[40,50)	0.01305	0.02062	0.005068	0.009086
[50,60)	0.03822	0.05965	0.0145085	0.0260137
[60,70)	0.07997	0.1245	0.03386	0.06071
[70,∞)	0.11808	0.18402	0.05771	0.10348

4.5 MEN1

	Female		Male	
Age Range	Homozygous Normal	Heterozygous	Homozygous Normal	Heterozygous
[0,20)	0.00054	0.06851	0.00049	0.06851
[20,30)	0.00197	0.30529	0.002247	0.30529
[30,40)	0.00388	0.71125	0.00501	0.71125
[40,50)	0.0069	0.90842	0.00802	0.90842
[50,60)	0.01564	0.96661	0.01276	0.96661
[60,70)	0.02971	0.99672	0.02031	0.99672
[70,∞)	0.10412	0.99891	0.03135	0.99891

4.6 MLH1

	Female		Male	
Age Range	Homozygous Normal	Heterozygous	Homozygous Normal	Heterozygous
[0,20)	0	0	0	0
[20,30)	0.00117	0.0103	0.000525	0.00554
[30,40)	0.00375	0.0394	0.00183	0.0266
[40,50)	0.0106	0.148	0.0064	0.104
[50,60)	0.0273	0.356	0.01816	0.246
[60,70)	0.0544	0.519	0.0426	0.384
[70,∞)	0.0915	0.628	0.0809	0.521

4.7 MSH6

	Female		Male	
Age Range	Homozygous Normal	Heterozygous	Homozygous Normal	Heterozygous
[0,20)	0	0	0	0
[20,30)	0.00134	0.037	0.00077	0.01016
[30,40)	0.00555	0.07731	0.00263	0.02139
[40,50)	0.01473	0.127234	0.00829	0.04015
[50,60)	0.03496	0.24224	0.02536	0.10982
[60,70)	0.06735	0.44613	0.05639	0.26541
[70,∞)	0.10421	0.65798	0.09296	0.46772

4.8 PMS2

	Female		Male	
Age Range	Homozygous Normal	Heterozygous	Homozygous Normal	Heterozygous
[0,20)	0	0	0	0
[20,30)	0.00134	0.02366	0.00077	0.0118
[30,40)	0.00555	0.04954	0.00263	0.02483
[40,50)	0.01473	0.08676	0.00829	0.04693
[50,60)	0.03496	0.19066	0.02536	0.12596
[60,70)	0.06753	0.34858	0.05639	0.26013
[70,∞)	0.10421	0.49644	0.09296	0.39772