

PIAGE FAQs

Frequently asked questions

What is the theory behind the power/sample size estimation implemented in PIAGE?

Power and sample size estimations are based on Gauderman's (2002) asymptotic approach for power and sample size estimations in direct studies of GxE. Hardy-Weinberg equilibrium and independence of genotypes and environmental exposures in the population are assumed. The estimates are based on genotypic codes ($G=1$ ($G=0$) for individuals who carry a (non-) risk genotype), which depend on the mode of inheritance (dominant, recessive, or multiplicative). A conditional logistic regression approach is used, which employs a likelihood-ratio test with respect to a biallelic candidate SNP, a binary environmental factor ($E=1$ ($E=0$) in (un)exposed individuals), and the interaction between these components.

For further details see Hein et al. (2008).

How can appropriate values of linkage disequilibrium be determined?

Linkage disequilibrium between disease variant and marker depends on corresponding allele and 2-locus haplotype frequencies (f_T , f_M , f_{TM}). Depending on the selected measure (D or r^2), appropriate values can be calculated by

$$D = f_{TM} - f_T * f_M$$

$$r^2 = D^2 / (f_T * (1 - f_T) * f_M * (1 - f_M))$$

How are the results being presented?

Results of PIAGE are printed on the screen and saved in text-files.

Depending on the hypothesis to be tested, the print-out on the screen includes marker allele frequency, magnitude of the interaction between true disease variant and environmental exposure, measures of linkage disequilibrium (D and r^2), as well as power/sample size required to detect the hypothesised effect. For more information on the content of the result files see the documentation for PIAGE included in PIAGE.zip.

How do I estimate power/sample size for direct association studies?

To estimate power/sample size for direct association studies, you need to specify identical disease and marker allele frequencies and a measure of linkage disequilibrium (D or r^2) which is equal to 1.

Which version of R supports PIAGE?

PIAGE was developed using R 2.5.1 and is supposed to run with older versions as well. You can download the latest version of R using the following link: <http://cran.r-project.org>.

Which operating system can be utilized to estimate power and sample size using PIAGE?

Windows and Unix can be utilized to estimate power and sample size using PIAGE.

References

Gauderman WJ. 2002. Sample size requirements for matched case-control studies of gene-environment interaction. *Stat Med* 21:35-50.

Hein R, Beckmann L, Chang-Claude J. 2008. Sample size requirements for indirect association studies of gene-environment interactions (GxE). *Genet Epidemiol*. Dec 28 2007 [Epub ahead of print].

Kraft P, Yen YC, Stram DO, Morrison J, Gauderman WJ. 2007. Exploiting gene-environment interaction to detect genetic associations. *Hum Hered* 63:1141-119.

R Development Core Team. 2005. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0. URL <http://www.R-project.org>.