

endogamous/consanguineous populations, and that it can be improved by making use of precise estimates of the inbreeding coefficient based on the analysis of genomic patterns of homozygosity.⁷

CONFLICT OF INTEREST

The authors declare no conflict of interest.

Alessandro Gialluisi¹, Tommaso Pippucci² and Giovanni Romeo^{*,2,3}

¹Language and Genetics Department, Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands;

²Unità Operativa di Genetica Medica, Policlinico Sant'Orsola Malpighi, Bologna, Italy;

³European Genetics Foundation, Bologna, Italy
E-mail: romeo@eurogene.org

Reconciling clinical importance and statistical significance

European Journal of Human Genetics (2014) **22**, 158–159;
doi:10.1038/ejhg.2013.110; published online 22 May 2013

Upon revisiting our published work on the genetic architecture of adult height,¹ we noted the following sentence: ‘Height was measured with a stadiometer to the nearest 0.1 cm.’ Genetic association studies generally do not utilize information regarding precision of phenotypic measurements. We suggest that doing so addresses a long-standing issue between clinical importance and statistical significance. In this specific instance, we suggest that effects smaller than 0.1 cm are clinically insignificant, as they are not measurable; the issue then becomes how to incorporate this limit into statistical approaches in human genetics investigations such as genome-wide association studies (GWAS).

One solution to this problem is based on classical measurement error. Assume that the observed value of the outcome (dependent variable) y^* is equal to the true, underlying value of y plus a random component e . In ordinary least squares (OLS), the true model $y = X\beta + \varepsilon$ becomes $y^* = X\beta + e + \varepsilon$. Assuming that (1) the two errors are uncorrelated, (2) the expected values of both errors are 0, and (3) both error terms are uncorrelated with the independent variable, the OLS estimate $\hat{\beta}$ is a consistent and unbiased estimate of β . However, the variance of $\hat{\beta}$ increases from $(X'X)^{-1}\text{var}(\varepsilon)$ to $(X'X)^{-1}[\text{var}(e) + \text{var}(\varepsilon)]$. Consequently, test statistics accounting for measurement error will be smaller. Critically, claims of statistical significance will be limited by the precision of the measurement of the outcome.

A second solution to this problem is to use protected inference.² Classical inference is based on testing what is called a point null hypothesis: in GWAS, a normally distributed test statistic can be formulated as $(\hat{\beta} - \beta_0)/\text{SE}(\hat{\beta})$ with the point null value $\beta_0 = 0$. In slightly simplified terms, $\hat{\beta}$ is called consistent if it converges to the true value of β with enough data, ie, a suitably large sample size. However, the probability mass of any point in a continuous

- 1 Ten Kate LP, Teeuw M, Henneman L, Cornel MC: Comment on Gialluisi *et al*. *Eur J Hum Genet* 2014; **22**: 157.
- 2 Gialluisi A, Pippucci T, Anikster Y *et al*: Estimating the allele frequency of autosomal recessive disorders through mutational records and consanguinity: the homozygosity index (HI). *Ann Hum Genet* 2012; **76**: 159–167.
- 3 Gialluisi A, Incollu S, Pippucci T *et al*: The homozygosity index (HI) approach reveals high allele frequency for Wilson disease in the Sardinian population. *Eur J Hum Genet* 2013; **21**: 1308–1311.
- 4 Cavalli-Sforza LL, Moroni A, Zei G: *Consanguinity, Inbreeding and Genetic Drift in Italy*. Princeton, NJ: Princeton University Press, 2004.
- 5 Gialluisi A, Dedi D, Francks C, Fisher SE: Persistence and transmission of recessive deafness and sign language: new insights from village sign languages. *Eur J Hum Genet* 2013; **21**: 894–896.
- 6 Zappu A, Magli O, Lepori MB *et al*: High incidence and allelic homogeneity of Wilson disease in 2 isolated populations: a prerequisite for efficient disease prevention programs. *J Pediatr Gastroenterol Nutr* 2008; **47**: 334–338.
- 7 Pemberton TJ, Absher D, Feldman MW, Myers RM, Rosenberg NA, Li JZ: Genomic patterns of homozygosity in worldwide human populations. *The Am J Hum Genet* 2012; **91**: 275–292.

distribution is 0. Consequently, statistical significance can be attained for trivial effects by simply increasing the sample size (Figure 1a). Under protected inference, the null hypothesis is an interval rather than a point. We suggest two ways to implement protected inference. We can control the false positive error rate at the borders of the null interval, allowing the test to become overly conservative within the null interval (Figure 1b). Alternatively, we can control the false positive error rate at a fixed value across the entire null interval (Figure 1c). We recommend this latter approach,

formulating the test statistic as $\begin{cases} \frac{|\hat{\beta}| - \beta_0}{\text{SE}(\hat{\beta})} & \text{if } |\hat{\beta}| \geq \beta_0 \\ 0 & \text{otherwise} \end{cases}$, with β_0 now

corresponding to the limit of precision, eg, 0.1 cm for our stadiometer. Either way, protected inference prevents arbitrarily large sample sizes from yielding statistical significance for trivial effects.

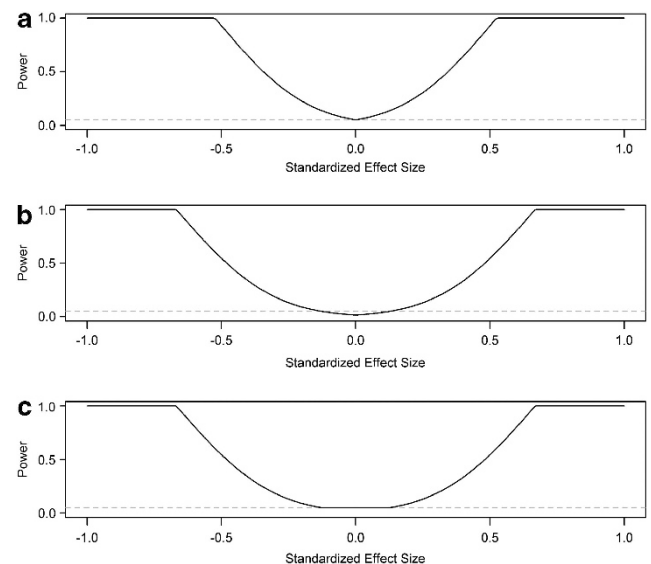


Figure 1 Power in classical vs protected inference. The gray dashed lines represent the significance level $\alpha = 0.05$. (a) Classical inference involves testing the point null hypothesis that the effect size is 0. (b) Protected inference involves testing a null interval hypothesis. In this implementation of protected inference, the false positive error rate is controlled at the borders of the null interval. (c) In this alternative implementation of protected inference, the false positive error rate is fixed across the entire null interval.

Another way to address measurement error is through use of repeated measures. In a repeated measures design, the dependent variable is measured multiple times for each individual to account for within-individual variability. However, each measurement is still limited by the precision of the instrument, so that repeated measures designs will not protect against trivial effects.

Returning to our example, we standardized our measurement error (0.1 cm corresponded to 0.0103 SD) and compared it with the standardized effect sizes reported for 180 loci for human height identified by meta-analysis of GWAS.³ The smallest effect reported in their follow-up (Stage 2) analysis was 0.010 SD, suggesting that the detection of genetic loci influencing human height is near its limit; increasing the sample size will have limited utility until the precision of the measurement of height improves.

Clinical importance ought to have much to do with statistical significance. The use of statistical models that account for the precision of measured traits, either through measurement error models or null interval testing, will eliminate the problem of trivial effects, thereby helping to reconcile clinical importance and statistical significance.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGEMENTS

The contents of this publication are solely the responsibility of the authors and do not necessarily represent the official view of the National Institutes of Health. This research was supported by the Intramural Research Program of the Center for Research on Genomics and Global Health (CRGGH). The CRGGH is supported by the National Human Genome Research Institute, the National Institute of Diabetes and Digestive and Kidney Diseases, the Center for Information Technology, and the Office of the Director at the National Institutes of Health (Z01HG200362).

Daniel Shriner*, Adebowale Adeyemo and Charles N Rotimi
*Center for Research on Genomics and Global Health,
National Human Genome Research Institute, Bethesda, MD, USA*
E-mail: shrinerda@mail.nih.gov

1 Shriner D, Adeyemo A, Gerry NP *et al*: Transferability and fine-mapping of genome-wide associated loci for adult height across human populations. *PLoS One* 2009; **4**: e8398.

2 Friston K: Ten ironic rules for non-statistical reviewers. *NeuroImage* 2012; **61**: 1300–1310.

3 Lango Allen H, Estrada K, Lettre G *et al*: Hundreds of variants clustered in genomic loci and biological pathways affect human height. *Nature* 2010; **467**: 832–838.