

LIFETIME RISK OF AUTOSOMAL RECESSIVE MITOCHONDRIAL DISORDERS CALCULATED FROM GENETIC DATABASES

Jing Tan^{1,2*#}, Matias Wagner, MD^{1,3,4*}, Sarah L. Stenton, MD³, Tim M. Strom, MD^{1,3}, Saskia B. Wortmann, MD, PhD^{1,3,5}, Holger Prokisch, PhD^{1,3}, Thomas Meitinger, MD^{1,3}, Konrad Oexle, MD⁴, Thomas Klopstock, MD^{2,6,7}

¹ Institute of Human Genetics, School of Medicine, Technische Universität München, Munich, Germany

² Department of Neurology, Friedrich-Baur-Institute, University Hospital of the Ludwig-Maximilians-Universität München, Munich, Germany

³ Institute of Human Genetics, Helmholtz Zentrum München, Neuherberg, Germany

⁴ Institute of Neurogenomics, Helmholtz Zentrum München, Neuherberg, Germany

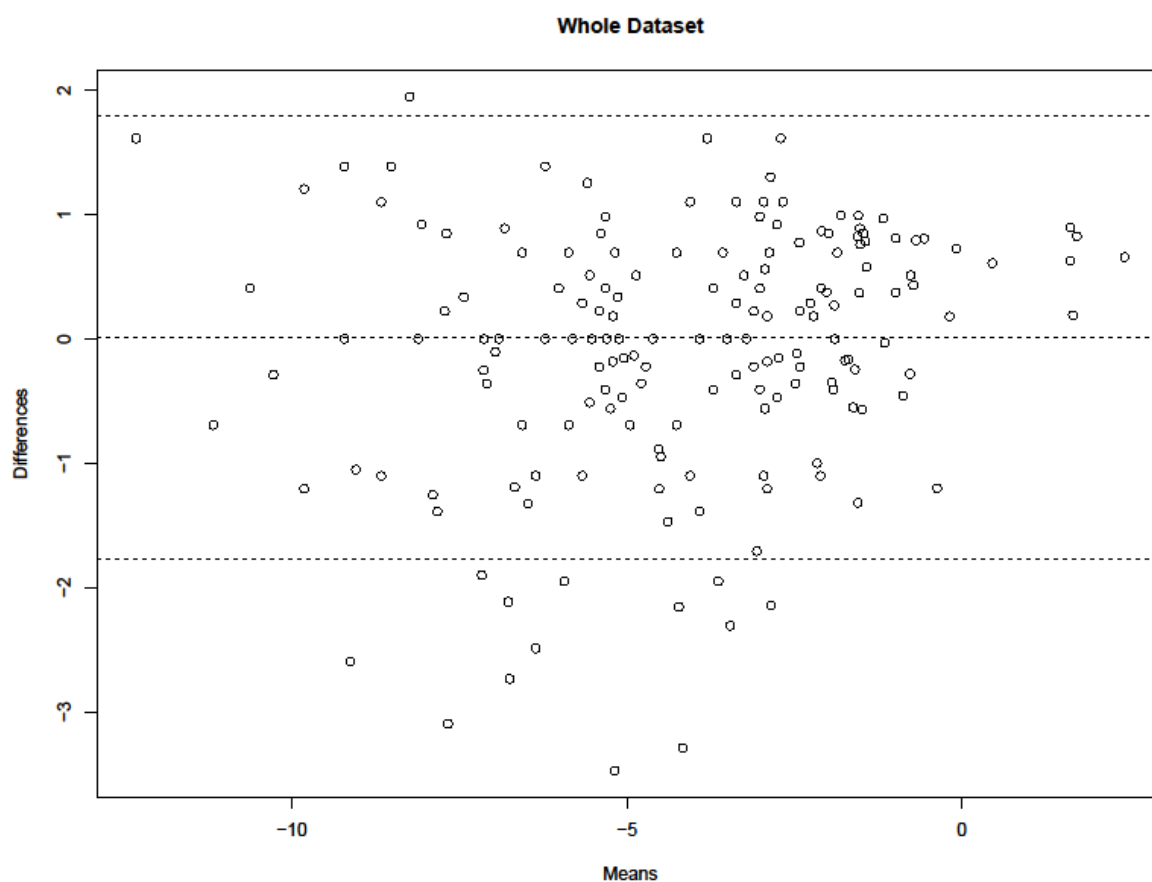
⁵ Department of Pediatrics, University Children's Hospital, Paracelsus Medical University (PMU), Salzburg, Austria

⁶ German Center for Neurodegenerative Diseases (DZNE), Munich, Germany

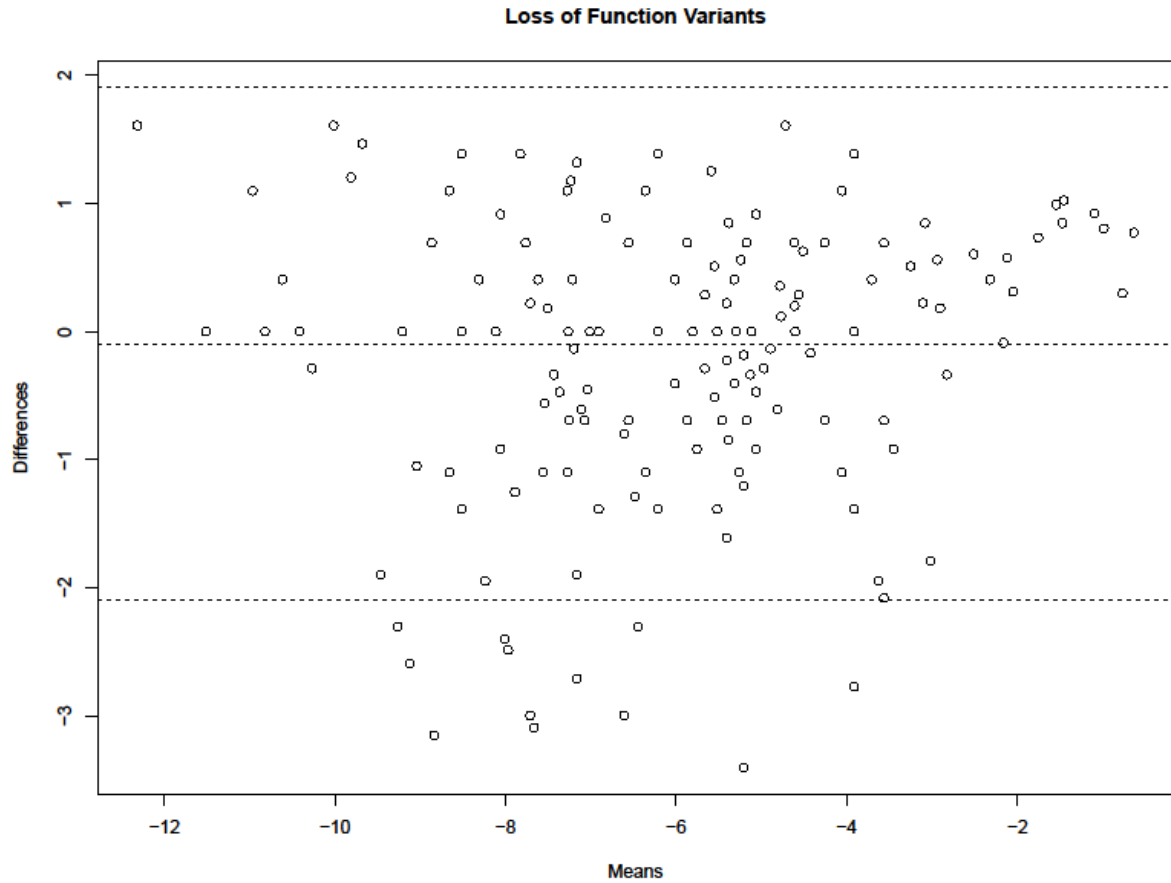
⁷ Munich Cluster for Systems Neurology (SyNergy), Munich, Germany

SUPPLEMENTARY MATERIAL

SUPPLEMENTARY FIGURES



Supplementary Figure 1: Bland-Altman Blot of log-transformed lifetime risk values per monogenic disorder comparing estimates based on worldwide variant data with estimates based European variant data. The differences are plotted on the y-axis shows, the means on the x-axis. The dotted lines represent the mean difference and the 95% confidence intervals. The plot shows a high degree of correlation between the lifetime risks derived from the worldwide and the European gnomAD dataset, respectively (Lin's concordance correlation coefficient of log-transformed lifetime risks $\rho_c = 0.938$).



Supplementary Figure 2: Bland-Altman Blot of log-transformed lifetime risk values per monogenic disorder comparing estimates based on worldwide data and estimates based on European data, using for the estimation Loss-of-Function variants only. The plot is designed as in Suppl. Fig. 1. Although the Loss-of-Function variants are a minor subset (9.4% of the whole dataset) the plot still shows a high degree of correlation ($\rho_c = 0.892$). Of note, the mean $M_{\log}$ of the pairwise differences of the log-transformed risk values as depicted here, is not a faithful representation of the analogous mean M_{nat} of the native, non-transformed values because $\sum_i (\log(a_i) - \log(b_i)) \neq \log(\sum_i ((a_i) - (b_i)))$, so that for two datasets D_1 and D_2 , $M_{\text{nat},1} > M_{\text{nat},2}$ does not generally imply $M_{\log,1} > M_{\log,2}$.