Impact of gnomAD v4 on BS1 Application

Primary Author: Michelle E. Duffe, Department of Genetic Counseling, MGH Institute of Health Professions, Boston, MA, USA
Author 2: Samantha M. Baxter, Broad Institute of MIT and Harvard, Cambridge, MA, USA
Author 3: Grace E. VanNoy, Ambry Genetics, Aliso Viejo, CA, USA
Author 4: Brooklynn Gasser, Ambry Genetics, Aliso Viejo, CA, USA
Author 5: Danielle R. Azzariti, Broad Institute of MIT and Harvard, Cambridge, MA, USA

Abstract

The Genome Aggregation Database (gnomAD) is a population database composed of exome and genome data from around the world. The database is used by many in the genomics field, including clinical genetic laboratories. Since gnomAD aggregates data from existing research studies, groups that are underrepresented in research are also underrepresented in gnomAD. The latest version of the database (v4) is almost 5 times larger than previous versions, with a greater number of non-European genetic ancestry samples. BS1 is a population data criterion that can be applied during variant classification as benign evidence, as described by Richards et al. (2015). BS1 application suggests that a variant is likely too common to be associated with a particular genetic disease. This retrospective study aimed to quantitatively assess the impact of gnomAD v4 on the application of the BS1. This study used a review of pediatric clinical exome cases, each with at least one variant classified as a VUS, from a single laboratory between 2019 to 2023. The dataset included demographic information and genomic data. For each gene in the dataset, the BS1 threshold utilized by the clinical laboratory at initial classification was provided. Maximum allele frequencies from gnomAD v2 and v4.1 were compared with gene-specific BS1 thresholds for each variant in the dataset to evaluate if the increased size of gnomAD v4.1 resulted in BS1 applications for more variants. Descriptive statistics were then used to contextualize the results. A total of 754 variants were identified within the cohort of 590 patients. Compared to gnomAD v2, more variants in our analysis were present in gnomAD v4.1 (47.3% in v4.1 vs. 31.7% in v2), but the percentage of variants present in gnomAD with an allele frequency high enough to apply BS1 decreased by about half (8.7% in v4.1 vs. 15.3% in v2). This study identified just six variants for which the BS1 criterion could be newly applied when v4.1 was utilized as compared to v2, demonstrating that the use of gnomAD v4.1 had a limited effect on BS1 application in this study. In all six cases, the variant was absent from gnomAD v2 and above the BS1 threshold in gnomAD v4.1. Importantly, for all six cases, the variant classification changed from uncertain to likely benign when the BS1 criterion was applied. In conclusion, the findings from this study highlight some of the complexities of the use of population data in variant classification. As demonstrated by the case examples, allele frequency may change over time as new versions of population databases are released, which may impact variant classification.