Received: 3 September 2025 | Revised: 11 December 2025 | Accepted: 22 December 2025

https://doi.org/10.1016/j.jtha.2025.12.025

ORIGINAL ARTICLE

Global prevalence of hereditary hemorrhagic telangiectasia-associated variants estimated by analysis of large-scale genomic databases

Eleonora Gaetani[1,2,3] | Agnese Giovannetti[4] | Luigi Di Martino[1,3] | Niccolò Liorni[4] | Viviana Caputo[5] | Antonio Gasbarrini[1,2,6] | Roberto Pola[1,7] | Tommaso Mazza[4]

[1]Hereditary Hemorrhagic Telangiectasia Center, Fondazione Policlinico Universitario Agostino Gemelli Istituto di ricovero e cura a carattere scientifico (IRCCS), Rome, Italy

[2]Department of Translational Medicine and Surgery, Università Cattolica del Sacro Cuore, Rome, Italy

[3]Unità Operativa Complessa di Medicina Interna, Ospedale Cristo Re, Rome, Italy

[4]Unità Operativa Semplice Computational Biology and Bioinformatics, Fondazione Policlinico Universitario Agostino Gemelli Istituto di ricovero e cura a carattere scientifico (IRCCS), Rome, Italy

[5]Department of Experimental Medicine, La Sapienza University of Rome, Rome, Italy

[6]Unità Operativa Complessa of Internal Medicine and Gastroenterology, Fondazione Policlinico Universitario Agostino Gemelli Istituto di ricovero e cura a carattere scientifico (IRCCS), Rome, Italy

[7]Percorso Trombosi, Department of Aging, Orthopedic, and Rheumatologic Sciences, Fondazione Policlinico Universitario Agostino Gemelli Istituto di ricovero e cura a carattere scientifico (IRCCS), Rome, Italy

Correspondence
Roberto Pola, Room Q812, Policlinico A. Gemelli, L.go A. Gemelli 8 00168 Rome, Italy.
Email: roberto.pola@unicatt.it

Abstract

Background: Hereditary hemorrhagic telangiectasia (HHT) is an autosomal dominant disorder with an overwhelming hemorrhagic phenotype. It is mainly caused by variants in the *ENG* and *ACVRL1* genes. HHT prevalence is currently estimated to be 1 in 5000 individuals, but the disease is likely underdiagnosed due to variable clinical presentation, misdiagnosis, and delayed recognition.

Objectives: To estimate the global genetic prevalence of HHT-associated variants in *ENG* and *ACVRL1*.

Methods: We analyzed 3 large population-scale genomic databases: gnomAD, All of Us, and Regeneron Genetics Center–Million Exome. We considered known pathogenic and likely pathogenic variants of *ENG* and *ACVRL1* and extended the analysis to potentially pathogenic variants passing the pathogenic criteria established by the guidelines for HHT of the American College of Medical Genetics and Genomics/Association for Molecular Pathology.

Results: The genetic prevalence of HHT ranged from 1.753 to 2.555 in 5000 individuals, when considering only pathogenic and likely pathogenic variants, and from 2.874 to 4.327 in 5000 individuals, when also potentially pathogenic variants were considered.

Conclusion: This study assesses the prevalence of HHT-associated variants in the general population. Our unbiased approach demonstrates that the genetic prevalence of the disease is substantially higher than currently estimated.

KEYWORDS
genetic analysis, global prevalence, hereditary hemorrhagic telangiectasia

Manuscript handled by: Suzanne Cannegieter

Final decision: Suzanne Cannegieter, 22 December 2025

1 | INTRODUCTION

Hereditary hemorrhagic telangiectasia (HHT) is a rare autosomal dominant disease characterized by recurrent spontaneous epistaxis, mucocutaneous telangiectasias, and visceral arteriovenous malformations (AVMs) [1,2]. It is caused by loss-of-function pathogenic variants in genes encoding proteins in the bone morphogenetic protein signaling pathway. Most of the disease-causing variants are observed in *ENG* and *ACVRL1*, with *SMAD4* and *GDF2* less frequently responsible. The most frequent HHT manifestations in adults are recurrent epistaxis and bleeding from gastrointestinal telangiectasias, with iron deficiency and anemia. Major complications may occur because of the presence of AVMs in the lungs, liver, and central nervous system, with ischemic stroke or cerebral abscesses caused by right-to-left vascular shunting through pulmonary AVMs, intracranial hemorrhages due to brain AVMs, and high cardiac output with heart failure due to left-to-right shunting through systemic AVMs. Management objectives of HHT include control of epistaxis, screening and treatment of iron deficiency and anemia, screening and treatment of gastrointestinal telangiectasias, management of AVMs in the lung and the central nervous system, and genetic counseling for patients and family members. Novel therapeutic approaches, such as systemic anti-angiogenic treatments, are being actively investigated [3]. Although HHT is associated with increased morbidity, appropriate screening and treatment of visceral AVMs, effective management of anemia, and management of overt or occult bleeding sites have the potential to improve overall survival and quality of life (QoL) [4].

Currently, the estimated prevalence of HHT is approximately 1 in 5000 individuals [5]. However, HHT is often not detected until it reaches an advanced stage, or the full clinical picture becomes evident in adulthood [6]. In a German analysis, it has been reported that an average of 18 years elapses between the first signs of the disease and the diagnosis of HHT [7]. In addition, HHT is extremely heterogeneous in terms of clinical presentation, even among members of the same family that carry the same mutations [8]. This is another reason why the disease is often underdiagnosed or diagnosed later. When a new patient is diagnosed, it is common to identify previously undiagnosed relatives. A population-based study conducted in the United Kingdom in 2014 concluded that HHT was significantly underdiagnosed in subjects from lower socioeconomic groups, likely due to lower rates of consultation with primary care services [9]. It is plausible to assume that the same issue exists in rural and underdeveloped areas worldwide. Therefore, it is likely that the actual prevalence of HHT is substantially underestimated at the global level.

In recent years, the public availability of large-scale genomic databases has allowed for unbiased estimation of rare disease prevalence, based on the allele frequency of known or predicted disease-causing variants. Recent examples of such approach are the analyses conducted by Seidizadeh et al. [10] to determine the actual global prevalence of hereditary thrombotic thrombocytopenic purpura and platelet-type von Willebrand disease [11]. In this context, a very recent study using a research-based approach, estimated HHT prevalence to be higher than current estimates, ranging from 2.1 to 11.9 in 5000 individuals [12].

Here, using a medical genetics-oriented approach, we estimated the genetic prevalence of HHT through the analysis of 3 population-scale genomic databases, gnomAD, All of Us, and the Regeneron Genetics Center–Million Exome (RGC-ME). To this aim, we considered HHT disease-causing variants in *ENG* and *ACVRL1* including pathogenic and likely pathogenic variants, coming from both clinical databases and passing the pathogenic criteria established by the Variant Curation Guidelines for HHT of the American College of Medical Genetics and Genomics/Association for Molecular Pathology (ACMG/AMP) [13]. Despite the more conservative approach, the results of these analyses further showed that the genetic prevalence of HHT is higher than currently estimated, supporting the concept that the disease might be underdiagnosed and that many individuals around the world are not receiving appropriate care.

2 | METHODS

Pathogenic variant selection

Pathogenic (P) and likely pathogenic (LP) variants in *ENG* and *ACVRL1* genes were retrieved from 3 clinical databases: the electronic repository curated by the ClinGen HHT Expert Panel (ClinGen) (accessed in February 2025); "disease causing" ("DM") and "likely disease-causing" ("DM?") variants from the Human Gene Mutation Database (HGMD), a manually curated collection of published germline mutations (accessed in December 2024); "pathogenic" and "likely pathogenic" variants from the ClinVar database, a publicly accessible archive of human variations associated with diseases (accessed in February 2025). Variants with conflicting clinical interpretations among the 3 databases were excluded from the analysis.

Potentially pathogenic (PP) variants were retrieved from population databases as those that passed the pathogenic criteria generally applicable to variants without *a priori* knowledge of the clinical information of an individual, that is PVS1, PS1, PM1, PM2, PM4, PM5, and PP3, according to the Variant Curation Guidelines for HHT of the ACMG/AMP (Supplementary Table S1) [13]. Any of these variants were considered if at least one of the "strong" or "very strong" criteria was met; variants meeting "moderate" criteria were retained only if at least a "supporting" criterion was also met. We excluded variants for which only "supporting" criteria (PM2 or PP3) could be attributed.

Variant annotation

When available, the genomic coordinates were aligned with the human genome assembly GRCh38/hg38 for all the collected variants. Missing coordinates were retrieved using Variant Validator [14], a

software package that precisely maps genomic variants to the desired human genome reference. Subsequently, all variants were annotated using the Ensembl Variant Effect Predictor tool [15] version 111, with RefSeq version 109 as the reference transcript set. Human Genome Variation Society (HGVS) nomenclature was provided based on the MANE transcripts NM_001114753.3 (*ENG*) and NM_000020.3 (*ACVRL1*). The variants were annotated using a custom pipeline that integrates annotations from dbNSFP 4.1a [16], a database of functional predictions and annotations of all potential nonsynonymous single-nucleotide variants in the human genome, and a wealth of other specific annotations, such as dosage sensitivity, loss-of-function putative effects, and impairment of the nonsense-mediated decay biological process. This integrated functional annotation allowed us to identify PP variants according to the established ACMG/AMP guidelines for HHT [13].

Population-scale databases

To estimate the genetic prevalence of HHT, we considered single nucleotide variants (SNVs) and small insertions and deletions (small indels, <50 bp) in coding exons, introns, UTRs, and splicing regions of *ENG* and *ACVRL1* genes (MANE transcripts NM_001114753.3 and NM_000020.3, respectively) from the following databases: gnomAD 4.1 (accessed in December 2024), a database of 807,162 unrelated individuals which were not affected by pediatric diseases and that were collected from various disease-specific and population genetic studies [17]; All of Us (Public Tier accessed in February 2025), a database that compiles data from 414,840 genome sequences from volunteer participants, resident in the USA, to focus on populations historically underrepresented in biomedical research [18]; and the RGC-ME (accessed in March 2025), which contains genomic variants from 821,979 adult and unrelated individuals sequenced through exome sequencing and coming from dozens of collaborations, including large biobanks and health systems [19]. High-quality *ENG* and *ACVRL1* variants were retrieved from gnomAD and RGC-ME. Because the All of Us database did not publicly indicate quality flags, we treated all retrieved variants from this database as high quality.

Prevalence estimation

To assess the genetic prevalence of HHT, we used genotype data from gnomAD, All of Us, and RGC-ME population databases. Initially, we searched for *ENG* and *ACVRL1* variants that are known to be associated with HHT (P or LP), as documented in clinical databases, in gnomAD, All of Us, and RGC-ME databases. Following this, we searched for PP variants in the same large-scale databases, guided by the ACMG/AMP Variant Curation Guidelines for HHT [13]. We identified individuals carrying at least one pathogenic allele and calculated the prevalence as the proportion of heterozygous variant carriers among the total number of genotyped individuals. When only allele counts were available and genotype counts were not directly reported, as for RGC-ME, we resorted to the Hardy-Weinberg equilibrium ($p^2 + 2pq + q^2$), where p represents the population frequency of the major allele and q denotes the population frequency of the minor allele. Given that HHT is a dominant disorder, our prevalence estimation focused on the frequency of heterozygous ($2pq$) and homozygous individuals (q^2). Statistical analyses were performed using a two-sided binomial test.

TABLE 1 Effects of the 1239 known pathogenic and likely pathogenic nonredundant variants of *ENG* and *ACVRL1* found in the ClinGen, Human Gene Mutation Database, and ClinVar databases. The effects were reported based on the sequence ontology nomenclature (http://www.sequenceontology.org/). Variants generally defined as "protein-altering" cause the deletion and subsequent insertion of one or more amino acids, and therefore cannot be classified as inframe or frameshift variants.

	ENG		ACVRL1	
	Total number	Frequency	Total number	Frequency
Missense	115	17.7%	263	44.6%
Frameshift	305	47.0%	179	30.3%
Start lost	7	1.1%	0	0.0%
Stop gained	84	12.9%	69	11.7%
Intronic	3	0.5%	8	1.4%
Splice donor/acceptor[a]	81	12.5%	42	7.1%
Splice region[a]	28	4.3%	9	1.5%
Inframe indel[a]	19	2.9%	17	2.9%
Synonymous	1	0.2%	1	0.2%
UTR[a]	3	0.5%	1	0.2%
Protein-altering	3	0.5%	1	0.2%

[a] Some effects were grouped as follows:
1) splice donor/acceptor: splice_donor_variant and splice_acceptor_variant
2) splice region: splice_donor_5th_base_variant, splice_donor_region_variant, splice_region_variant and splice_polypyrimidine_tract_variant
3) inframe indel: inframe_deletion and inframe_insertion
4) UTR: 5_prime_UTR_variant and 3_prime_UTR_variant

RESULTS

The search of the clinical databases led to the identification of 17 P or LP variants from the electronic repository curated by the ClinGen HHT Expert Panel, 869 disease-causing ("DM") or likely disease-causing ("DM?") mutations in HGMD, and 769 P or LP variants in ClinVar, for a total of 1239 nonredundant variants (Figure). The effects of the 1239 P and LP variants, based on the sequence ontology (SO) nomenclature, are shown in Table 1. Of note, using the ACMG/AMP guidelines for HHT [13], we were able to classify 609 of these variants as LP. For this latter analysis, we modified the criteria outlined in the Supplementary Table S1. Since PS1 is a "strong" criterion

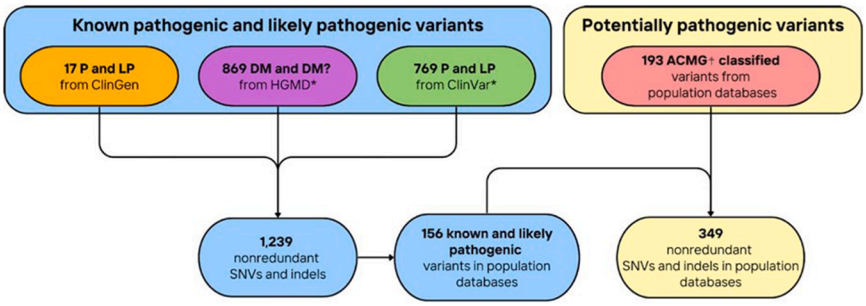

FIGURE Known pathogenic (P) and likely pathogenic (LP) variants of *ENG* and *ACVRL1* were retrieved from the ClinGen (n=17), HGMD (n=869), and ClinVar (n=769) databases; 1239 of these variants were not nonredundant among the 3 clinical databases; of these, 156 variants were found in population databases and used for a first estimation of HHT prevalence. Potentially pathogenic variants of *ENG* and *ACVRL1* were retrieved from population databases using the criteria of the American College of Medical Genetics and Genomics/Association for Molecular Pathology guidelines for HHT (n=193). These potentially pathogenic variants were added to the 156 pathogenic and likely pathogenic variants outlined above, resulting in 349 variants which were used for a further estimation of HHT prevalence. *Variants collected from the HGMD and ClinVar databases were considered only if reported to be associated with HHT. ‡Only automatically applicable criteria were used (see "Methods" section for specification, Supplementary Table S1). HGMD, Human Gene Mutation Database; HHT, hereditary hemorrhagic telangiectasis; P, pathogenic variants; LP, likely pathogenic variants; DM, disease-causing variant from HGMD; DM?, likely disease-causing variant from HGMD.

that can significantly contribute to a likely pathogenic classification and is applicable when a variant results in the "same amino acid change as a previously established pathogenic variant," we aimed to prevent the recursive classification of HGMD and ClinVar variants as likely pathogenic. Therefore, we considered only pathogenic amino acid changes if the HHT VCEP guidelines indicated them. Of the 1239 P or LP variants identified in the clinical databases, a total of 156 nonredundant were found in the genomic-population databases: 87 in gnomAD, 58 in All of Us, and 80 in RGC-ME (Figure 1). Based on these data, the estimated prevalence of HHT ranged from 1.753 in 5000 individuals in gnomAD, 1.788 in 5000 in RGC-ME, and 2.555 in 5000 in All of Us (Table 2 and Supplementary Table S2). When considering the different genetic ancestry groups available in gnomAD, the estimated genetic prevalence of HHT was significantly different from the current prevalence in African/African Americans (6.526/5000) and in Europeans (1.551/5000) (Table 3).

Then, to identify PP variants in the genomic-population databases gnomAD, All of Us, and RGC-ME, we used the criteria established by the ACMG/AMP guidelines for HHT, as outlined in the Methods section. This research led to the identification of 193 PP variants, out of the 30,629 variants occurring in the *ENG* and *ACVRL1* genes and reported in genomic databases (Table 4, Figure, and Supplementary Table S3; see Methods for details on the inclusion criteria). These 193 PP variants were added to the 156 P and LP variants identified through the search of the clinical databases, resulting in a total of 349 variants that were used for further prevalence estimation (Figure). The HHT estimated genetic prevalence based on these 349 variants was 2.874 in 5000 individuals in gnomAD, 3.120 in 5000 in RGC-ME, and 4.327 in 5000 in All of Us (Table 5 and Supplementary Table S4). When we considered the different genetic ancestry groups available in gnomAD, the estimated genetic prevalence of HHT was again significantly different from the current estimated prevalence, especially in African/African Americans (9.455/5000) and in Middle Easterners (6.598/5000) (Table 6).

DISCUSSION

HHT is a rare autosomal genetic disorder with an estimated prevalence of 1 in 5000 individuals. However, such estimates are derived from a combination of clinical observations, analysis of representative primary and secondary care databases, and cross-sectional studies based on existing or new diagnoses of HHT in dedicated centers [5]. Thus, the actual prevalence of HHT might be substantially higher, as the signs and symptoms of the disease are heterogeneous and may be misinterpreted, the typical clinical features might become evident only in adulthood, and many potential patients fail to be screened because of low access to medical services. In this study, we estimated the prevalence of HHT using genetic analyses. We used 3 distinct databases and focused firstly on P and LP variants in the *ENG* and *ACVRL1* genes and extended the analysis to PP variants passing the pathogenic criteria established by the ACMG/AMP guidelines for HHT [13]. The findings revealed a higher than currently known prevalence. Depending on the methodology used, prevalence ranged from 1.753 to 4.327 per 5000 individuals. This result is in line with the results coming from a very recent work, in which, using a different methodological approach, the authors found that HHT prevalence ranged from 2.1 to 11.9 in 5000 individuals [12].

It is important to note that these are genetic, rather than clinical, estimates. Nonetheless, HHT is generally considered a condition with incomplete, age-related penetrance and variable expressivity [13]. This means that carrying a P, LP, or PP variant does not necessarily correspond to having clinically evident disease. However, since penetrance increases with age, most people with a pathogenic

TABLE 2 Prevalence of hereditary hemorrhagic telangiectasia estimated considering the 156 pathogenic and likely pathogenic nonredundant variants of *ENG* and *ACVRL1* found in population databases. The proportion of subjects carrying pathogenic and likely pathogenic variants was directly observed in the gnomAD and All of Us databases. In contrast, in the Regeneron Genetics Center–Million Exome database, this proportion was estimated using the Hardy-Weinberg equation.

	Total number of sequenced individuals	Total number of heterozygous individuals	Total number of homozygous individuals	Estimated prevalence	Prevalence in 5000 individuals	*P* value
gnomAD	807,162	283	0	3.51E-04	1.753	5.48E-18
All of Us	414,840	212	0	5.11E-04	2.555	2.35E-32

	Total number of sequenced alleles	Total number of alternative alleles	Estimated frequency of alternative alleles	Estimated prevalence	Prevalence in 5000 individuals	*P* value
Regeneron Genetics Center–Million Exome	1,643,958	294	1.79E-04	3.58E-04	1.788	0.057

mutation in one of the HHT-associated genes will develop clinical signs of the disease over time. Moreover, due to the fact that the severity and combination of symptoms can vary widely, many affected individuals remain undiagnosed. Based on this, the results of our study support the concept that, at the global level, there is a conspicuous number of subjects who have the clinical signs of HHT, or will develop them over time, who are unknown to health services and do not receive appropriate screening and therapy.

Our study also provides evidence that the prevalence of HHT differs among various genetic ancestry groups in the gnomAD database. Looking only at the already known pathogenic and likely pathogenic variants, we found the highest prevalence in Africans/African Americans (6.526/5000). Regarding the Europeans (and Finnish), the estimated prevalence was 1.551/5000 (more than 55% higher than currently estimated). In contrast to the findings reported by Anzell et al. [12], our results suggest that HHT prevalence may vary across genetic ancestries, further supporting the hypothesis that the Afro-Caribbean population might have the highest prevalence [20].

In this analysis, we did not identify any homozygotes in either gnomAD or All of Us. This suggests that either, as supported by Hardy-Weinberg expectations, such homozygous individuals would occur at a frequency of roughly 1 in 5.3 million, far below the detection threshold of our cohorts, or that, a homozygous pathogenic phenotype would be extremely severe, and therefore potentially lethal [21]; nonetheless, a novel homozygous missense variant has been recently reported [21].

To classify variants, we adopted the ACMG/AMP classification framework specific for HHT [13], applied here with the sole purpose of identifying those that could be defined as potentially pathogenic. Interestingly, among the 193 variants considered across the population databases, as many as 37 could already be classified as LP. By

TABLE 3 Prevalence of hereditary hemorrhagic telangiectasia in different genetic ancestry groups of the gnomAD database estimated considering the 156 pathogenic and likely pathogenic nonredundant variants of *ENG* and *ACVRL1* found in population databases. Genetic ancestry groups follow the definitions used in gnomAD, as described on the gnomAD website.

	Total number of sequenced individuals	Total number of heterozygous individuals	Total number of homozygous individuals	Estimated prevalence	Prevalence in 5000 individuals	*P* value
African/African American	37,545	49	0	1.31E-03	6.526	8.30E-24
Admixed American	30,019	5	0	1.67E-04	0.833	0.839
Ashkenazi Jewish	14,804	3	0	2.03E-04	1.013	0.773
East Asian	22,448	5	0	2.23E-04	1.114	0.810
European (and Finnish)	622,057	193	0	3.10E-04	1.551	1.09E-08
Middle Eastern	3031	2	0	6.60E-04	3.299	0.124
South Asian	45,546	10	0	2.20E-04	1.098	0.738
Remaining (and Amish)[a]	31,712	16	0	5.05E-04	2.523	9.03E-04

[a] Remaining individuals mostly accounted for individuals with no ethnicity assigned by gnomAD 4.1.

TABLE 4 Effects of the 156 pathogenic and likely pathogenic nonredundant variants (identified by analysis of ClinGen, Human Gene Mutation Database, and ClinVar and found in population databases) (left) and the 193 potentially pathogenic variants of *ENG* and *ACVRL1* (found in population databases according to the American College of Medical Genetics and Genomics/Association for Molecular Pathology guidelines for hereditary hemorrhagic telangiectasia) (right). Effects are reported based on the sequence ontology nomenclature.

	Known pathogenic and likely pathogenic (n=156)				Potentially pathogenic (n=193)			
	ENG		ACVRL1		ENG		ACVRL1	
	Total number (n=66)	Frequency	Total number (n=90)	Frequency	Total number (n=77)	Frequency	Total number (n=116)	Frequency
Missense	24	36.4%	42	46.7%	16	20.8%	51	44.0%
Frameshift	14	21.2%	19	21.1%	25	32.5%	15	12.9%
Start lost	2	3.0%	0	0.0%	0	0.0%	2	1.7%
Stop gained	10	15.2%	15	16.7%	1	1.3%	6	5.2%
Stop lost	0	0.0%	0	0.0%	3	3.9%	1	0.9%
Intronic	1	1.5%	4	4.4%	0	0.0%	0	0.0%
Splice donor/acceptor[a]	10	15.2%	7	7.8%	7	9.1%	8	6.9%
Splice region[a]	3	4.5%	2	2.2%	0	0.0%	1	0.9%
Inframe indel[a]	1	1.5%	1	1.1%	22	28.6%	9	7.8%
Synonymous	0	0.0%	0	0.0%	3	3.9%	23	19.8%
UTR[a]	1	1.5%	0	0.0%	0	0.0%	0	0.0%

[a] Some effects are grouped:
1) Splice donor/acceptor: splice_donor_variant and splice_acceptor_variant
2) Splice region: splice_donor_5th_base_variant, splice_donor_region_variant, splice_region_variant and splice_polypyrimidine_tract_variant
3) Inframe indel: inframe_deletion and inframe_insertion
4) UTR: 5_prime_UTR_variant and 3_prime_UTR_variant

extending the analysis, we aimed to approximate an upper-bound estimate of potential genotype prevalence. Importantly, while this approach involved certain simplifications, all variant classifications were performed in accordance with established ACMG/AMP guidelines specific to HHT [13].

Our study presents multiple limitations: first, we did not consider variants in the *SMAD4* and *GDF2* genes. We are aware that the identification of these variants is crucial for accurate diagnosis of HHT. However, while *ENG* and *ACVRL1* are responsible for the vast majority of HHT cases, *SMAD4* and *GDF2* mutations are detected in only a small percentage of patients, often around 2%. The analysis of *SMAD4* and *GDF2* variants could only have led to a further increase in the estimate of the genetic prevalence of HHT. Moreover, there is an emerging hypothesis suggesting that somatic mutations in the second allele of the affected gene contribute to the development of vascular malformations [22]. This "two-hit" mechanism could further explain

TABLE 5 Prevalence of hereditary hemorrhagic telangiectasia estimated considering 349 pathogenic, likely pathogenic, and potentially pathogenic nonredundant variants of *ENG* and *ACVRL1* according to clinical databases and the American College of Medical Genetics and Genomics/Association for Molecular Pathology guidelines.

	Total number of sequenced individuals	Total number of heterozygous individuals	Total number of homozygous individuals	Estimated prevalence	Prevalence in 5000 individuals	*P* value
gnomAD	807,162	464	0	5.75E-04	2.874	1.19E-83
All of Us	414,840	359	0	8.65E-04	4.327	7.70E-111

	Total number of sequenced alleles	Total number of alternative alleles	Estimated frequency of alternative alleles	Estimated prevalence	Prevalence in 5000 individuals	p-value
Regeneron Genetics Center–Million Exome	1,643,958	513	3.12E-04	6.24E-04	3.120	6.17E-21

In the gnomAD and All of Us databases, the proportion of subjects carrying *P* variants was directly observed. In contrast, in the Regeneron Genetics Center–Million Exome database, this proportion was estimated using the Hardy-Weinberg equation.

TABLE 6 Prevalence of hereditary hemorrhagic telangiectasia in different genetic ancestry groups of the gnomAD database, estimated considering 349 pathogenic, likely pathogenic, and potentially pathogenic nonredundant variants of *ENG* and *ACVRL1* according to clinical databases and American College of Medical Genetics and Genomics/Association for Molecular Pathology guidelines.

	Total number of sequenced individuals	Total number of heterozygous individuals	Total number of homozygous individuals	Estimated prevalence	Prevalence in 5000 individuals	*P* value
African/African American	37,545	71	0	1.89E-03	9.455	1.00E-43
Admixed American	30,019	18	0	6.00E-04	2.998	5.72E-05
Ashkenazi Jewish	14,804	4	0	2.70E-04	1.351	0.549
East Asian	22,448	14	0	6.24E-04	3.118	2.45E-04
European (and Finnish)	622,057	301	0	4.84E-04	2.419	7.93E-41
Middle Eastern	3031	4	0	1.32E-03	6.598	0.003
South Asian	45,546	25	0	5.49E-04	2.744	1.05E-05
Remaining (and Amish)[a]	31,712	27	0	8.51E-04	4.257	9.48E-10

Genetic ancestry groups follow the definitions used in gnomAD, as described on the gnomAD website.

[a] Remaining individuals were mostly accounted for by individuals with no ethnicity assigned by gnomAD 4.1.

cases that remain undetected by germline variant analysis alone, as individuals carrying a P variant may not manifest HHT-associated vascular malformations until a secondary, somatic event occurs within affected tissues. Another possible limitation in the prevalence estimation through population-databases is that gnomAD explicitly excludes severely affected individuals, which may result in under-representation of clinical phenotypes. Moreover, regarding the different prevalence of HHT among various genetic ancestry groups, it should be noted that certain groups are over- or under-represented within this database. This is therefore a potential bias in terms of generalizability of our prevalence estimates. A further limitation is that gnomAD, All of Us, and RGC-ME do not provide per-individual genotypes or phasing data, therefore, we cannot determine whether different heterozygous variants occur within the same individual. However, the co-occurrence of two heterozygous pathogenic alleles is expected to be negligible given the stringent criteria used to select pathogenic, likely pathogenic and potentially pathogenic variants. Finally, considering the PP variants, we deliberately excluded broader, nonspecific criteria, such as general indicators of allele rarity (eg, PM2) or *in silico* functional predictions (eg, PP3), which, although informative in other settings, are not HHT-specific and may inflate pathogenicity estimates.

In this context, a recent study has shown that for missense variants in *ENG*, Alpha Missense could be a better predictor to identify missense PP variants compared with REVEL [12]. Therefore, considering this scoring system could lead to identifying a greater number of potentially pathogenic *ENG* missense variants and, consequently, to an increased prevalence. However, the currently used Variant Curation Guidelines for HHT of the ACMG/AMP [13] only indicate REVEL to attribute the PP3 criterion. In the same study, the authors estimated HHT prevalence using a research-based approach, using recently developed and widely used deleteriousness predictors, such as CADD and Alpha Missense, and developing a machine-learning-based classification system to improve the classification of HHT-related missense variants, the authors found a higher prevalence for HHT [12]. Here, we used a medical genetics-oriented approach, focusing on HHT-related variants already defined as "pathogenic" or "likely pathogenic" from highly reliable databases, including the electronic repository curated by the ClinGen HHT Expert Panel, and extending our prevalence estimation to variants that can be defined as "potentially pathogenic," guided by the ACMG/AMP Variant Curation Guidelines for HHT [13]. Although more stringent and almost complementary to the method presented by Anzell et al. [12], our approach further suggests that HHT is an underdiagnosed disease and that clinical criteria currently used to diagnose affected individuals lack sensitivity [13].

Nonetheless, the clinical implications of these findings are multiple. First, the true prevalence of HHT at the global level may be significantly higher than currently thought. Second, there is a high number of individuals at potentially increased risk of developing the disease. This is certainly important for early diagnosis and screening. Third, many subjects around the world may not have been diagnosed yet, and thus may not be receiving appropriate therapy. Additionally, identifying subjects carrying HHT-associated mutations—even if they are healthy and/or asymptomatic—would potentially enhance the early identification of other family members at risk. Finally, patient outcomes and QoL might be positively affected by broad genetic approaches to rare diseases.

In conclusion, this study attempts to determine the genetic prevalence of HHT using large-scale genetic databases. This

population-based analysis further indicates a substantially higher than currently estimated prevalence of variants associated with HHT at the global level. Strategies based on genetic screenings might facilitate timely clinical diagnosis and potentially improve patient outcomes and QoL.

AUTHOR CONTRIBUTIONS

E. Gaetani, A. Giovannetti, R. Pola, and T. Mazza designed the study; L. Di Martino, N. Liorni, and V. Caputo collected data; A. Giovannetti, N. Liorni, V. Caputo e T. Mazza performed the prevalence analyses; E. Gaetani, A. Giovannetti, A. Gasbarrini, R. Pola, and T. Mazza supervised all the aspects of the study; E. Gaetani, A. Giovannetti, R. Pola, and T. Mazza wrote the manuscript; all authors reviewed the final version.

DECLARATION OF COMPETING INTERESTS

The authors declare they have no conflict of interest.

ORCID

Roberto Pola https://orcid.org/0000-0001-5224-2931

REFERENCES

[1] Kritharis A, Al-Samkari H, Kuter DJ. Hereditary hemorrhagic telangiectasia: diagnosis and management from the hematologist's perspective. *Haematologica*. 2018;103:1433–43.

[2] Al Tabosh T, Al Tarrass M, Tourvieilhe L, Guilhem A, Dupuis-Girod S, Bailly S. Hereditary hemorrhagic telangiectasia: from signaling insights to therapeutic advances. *J Clin Invest*. 2024;134.

[3] Al-Samkari H, Kasthuri RS, Parambil JG, Albitar HA, Almodallal YA, Vázquez C, Serra MM, Dupuis-Girod S, Wilsen CB, McWilliams JP, Fountain EH, Gossage JR, Weiss CR, Latif MA, Issachar A, Mei-Zahav M, Meek ME, Conrad M, Rodriguez-Lopez J, Kuter DJ, et al. An international, multicenter study of intravenous bevacizumab for bleeding in hereditary hemorrhagic telangiectasia: the InHIBIT-Bleed study. *Haematologica*. 2021;106:2161–9.

[4] Zarrabeitia R, Fariñas-Álvarez C, Santibáñez M, Señaris B, Fontalba A, Botella LM, et al. Quality of life in patients with hereditary haemorrhagic telangiectasia (HHT). *Health Qual Life Outcomes*. 2017;15:19.

[5] Faughnan ME, Mager JJ, Hetts SW, Palda VA, Lang-Robertson K, Buscarini E, Deslandres E, Kasthuri RS, Lausman A, Poetker D, Ratjen F, Chesnutt MS, Clancy M, Whitehead KJ, Al-Samkari H, Chakinala M, Conrad M, Cortes D, Crocione C, Darling J, et al. Second International Guidelines for the Diagnosis and Management of Hereditary Hemorrhagic Telangiectasia. *Ann Intern Med*. 2020;173:989–1001.

[6] Danesino C, Cantarini C, Olivieri C. Hereditary hemorrhagic telangiectasia in pediatric age: focus on genetics and diagnosis. *Pediatr Rep*. 2023;15:129–42.

[7] Droege F, Kuerten CHL, Kaiser C, Dingemann J, Kaster F, Dahlfrancis PM, Lueb C, Zioga E, Thangavelu K, Lang S, Geisthoff U. Hereditäre hämorrhagische Teleangiektasie: Symptome und diagnostische Latenz [Hereditary hemorrhagic telangiectasia: symptoms and diagnostic latency] *Laryngorhinootologie*. 2021;100:443–52.

[8] Shovlin CL, Simeoni I, Downes K, Frazer ZC, Megy K, Bernabeu-Herrero ME, et al. Mutational and phenotypic characterization of hereditary hemorrhagic telangiectasia. *Blood*. 2020;136:1907–18.

[9] Donaldson JW, McKeever TM, Hall IP, Hubbard RB, Fogarty AW. The UK prevalence of hereditary haemorrhagic telangiectasia and its association with sex, socioeconomic status and region of residence: a population-based study. *Thorax*. 2014;69:161–7.

[10] Seidizadeh O, Cairo A, Mancini I, George JN, Peyvandi F. Global prevalence of hereditary thrombotic thrombocytopenic purpura determined by genetic analysis. *Blood Adv*. 2024;8:4386–96.

[11] Seidizadeh O, Cairo A, Othman M, Peyvandi F. Global prevalence of platelet-type von Willebrand disease. *Res Pract Thromb Haemost*. 2025;9:102682.

[12] Anzell AR, M White C, Diergaarde B, Carlson JC, Roman BL. Hereditary Hemorrhagic Telangiectasia Prevalence Estimates Calculated From GnomAD Allele Frequencies of Predicted Pathogenic Variants in ENG and ACVRL1. *Circ Genom Precis Med*. 2025 Oct;18(5):e005061. https://doi.org/10.1161/circgen.124.005061

[13] DeMille D, McDonald J, Bernabeu C, Racher H, Olivieri C, Cantarini C, Sbalchiero A, Thompson BA, Jovine L, Shovlin CL, Dupuis-Girod S, Lesca G, Tusseau M, Ganguly A, Kasthuri RS, Jessen J, Massink MPG, Ichikawa S, Bayrak-Toydemir P. Specifications of the ACMG/AMP Variant Curation Guidelines for Hereditary Hemorrhagic Telangiectasia Genes-ENG and ACVRL1. *Hum Mutat*. 2024;2024:3043736.

[14] Freeman PJ, Wagstaff JF, Fokkema IFAC, Cutting GR, Rehm HL, Davies AC, den Dunnen JT, Gretton LJ, Dalgleish R. Standardizing variant naming in literature with VariantValidator to increase diagnostic rates. *Nat Genet*. 2024;56:2284–6.

[15] McLaren W, Gil L, Hunt SE, Riat HS, Ritchie GR, Thormann A, Flicek P, Cunningham F. The ensembl variant effect predictor. *Genome Biol*. 2016;17:122.

[16] Liu X, Li C, Mou C, Dong Y, Tu Y. dbNSFP v4: a comprehensive database of transcript-specific functional predictions and annotations for human nonsynonymous and splice-site SNVs. *Genome Med*. 2020;12:103.

[17] Chen S, Francioli LC, Goodrich JK, Collins RL, Kanai M, Wang Q, Alföldi J, Watts NA, Vittal C, Gauthier LD, Poterba T, Wilson MW, Tarasova Y, Phu W, Grant R, Yohannes MT, Koenig Z, Farjoun Y, Banks E, Donnelly S, et al. A genomic mutational constraint map using variation in 76,156 human genomes. *Nature*. 2024;625:92–100.

[18] All of Us Research Program Investigators, Denny JC, Rutter JL, Goldstein DB, Philippakis A, Smoller JW, Jenkins G, Dishman E. The "All of Us" Research Program. *N Engl J Med*. 2019;381:668–76.

[19] Sun KY, Bai X, Chen S, Bao S, Zhang C, Kapoor M, Backman J, Joseph T, Maxwell E, Mitra G, Gorovits A, Mansfield A, Boutkov B, Gokhale S, Habegger L, Marcketta A, Locke AE, Ganel L, Hawes A, Kessler MD, et al. A deep catalogue of protein-coding variation in 983,578 individuals. *Nature*. 2024;631:583–92.

[20] Westermann CJ, Rosina AF, De Vries V, de Coteau PA. The prevalence and manifestations of hereditary hemorrhagic telangiectasia in the Afro-Caribbean population of the Netherlands Antilles: a family screening. *Am J Med Genet A*. 2003;116A:324–8.

[21] Mathavan A, Mathavan A, Krekora U, Rao A, Zumberg MS, Justice J, Bayrak-Toydemir P, McDonald J, Ataya A. Case report: Novel homozygous ACVRL1 missense variant in a family with hereditary hemorrhagic telangiectasia and pulmonary arterial hypertension: findings suggest a hypomorphic allele. *Front Genet*. 2025;16:1554624.

[22] DeBose-Scarlett E, Ressler AK, Gallione CJ, Sapisochin Cantis G, Friday C, Weinsheimer S, Schimmel K, Spiekerkoetter E, Kim H, Gossage JR, Faughnan ME, Marchuk DA. Somatic mutations in arteriovenous malformations in hereditary hemorrhagic telangiectasia support a bi-allelic two-hit mutation mechanism of pathogenesis. *Am J Hum Genet*. 2024;111:2283–98.

SUPPLEMENTARY MATERIAL

The online version contains supplementary material available at https://doi.org/10.1016/j.jtha.2025.12.025.