

# Whole genome sequences discriminate hereditary hemorrhagic telangiectasia phenotypes by non-HHT deleterious DNA variation

Katie E. Joyce,<sup>1,2</sup> Ebum Onabanjo,<sup>3</sup> Sheila Brownlow,<sup>3</sup> Fadumo Nur,<sup>3</sup> Kike Olupona,<sup>3</sup> Kehinde Fakayode,<sup>3</sup> Manveer Sroya,<sup>4</sup> Geraldine A. Thomas,<sup>4</sup> Teena Ferguson,<sup>3</sup> Julian Redhead,<sup>3</sup> Carolyn M. Millar,<sup>3,5</sup> Nichola Cooper,<sup>3,5</sup> D. Mark Layton,<sup>3,5</sup> Freya Boardman-Pretty,<sup>6</sup> Mark J. Caulfield,<sup>6,7</sup> Genomics England Research Consortium,<sup>6</sup> and Claire L. Shovlin<sup>2,3,8</sup>

<sup>1</sup>Imperial College School of Medicine, Imperial College, London, United Kingdom; <sup>2</sup>Genomics England Respiratory Clinical Interpretation Partnership (GeCIP), London, United Kingdom; <sup>3</sup>West London Genomic Medicine Centre, Imperial College Healthcare NHS Trust, London, United Kingdom; <sup>4</sup>Department of Surgery and Cancer, Imperial College, London, United Kingdom; <sup>5</sup>Centre for Haematology, Department of Immunology and Inflammation, Imperial College, London, United Kingdom; <sup>6</sup>Genomics England Research Consortium, Genomics England, London, United Kingdom; <sup>7</sup>William Harvey Research Institute, Queen Mary University of London, London, United Kingdom; and <sup>8</sup>National Heart and Lung Institute, Imperial College, London, United Kingdom

## Key Points

- We tested 75 HHT-independent genes by whole genome sequencing, integrating clinical severity classifications and population variant burdens.
- We show platelet/coagulation gene deleterious variants to be commonly present and less well tolerated in patients with HHT than the general population.

The abnormal vascular structures of hereditary hemorrhagic telangiectasia (HHT) often cause severe anemia due to recurrent hemorrhage, but HHT causal genes do not predict the severity of hematological complications. We tested for chance inheritance and clinical associations of rare deleterious variants in which loss-of-function causes bleeding or hemolytic disorders in the general population. In double-blinded analyses, all 104 patients with HHT from a single reference center recruited to the 100 000 Genomes Project were categorized on new MALO (more/as-expected/less/opposite) sub-phenotype severity scales, and whole genome sequencing data were tested for high impact variants in 75 HHT-independent genes encoding coagulation factors, or platelet, hemoglobin, erythrocyte enzyme, and erythrocyte membrane constituents. Rare variants (all gnomAD allele frequencies <0.003) were identified in 56 (75%) of these 75 HHT-unrelated genes. Deleteriousness assignments by Combined Annotation Dependent Depletion (CADD) scores >15 were supported by gene-level mutation significance cutoff scores. CADD >15 variants were identified in 38/104 (36.5%) patients with HHT, found for 1 in 10 patients within platelet genes; 1 in 8 within coagulation genes; and 1 in 4 within erythrocyte hemolytic genes. In blinded analyses, patients with greater hemorrhagic severity that had been attributed solely to HHT vessels had more CADD-deleterious variants in platelet (Spearman  $\rho = 0.25$ ;  $P = .008$ ) and coagulation (Spearman  $\rho = 0.21$ ;  $P = .024$ ) genes. However, the HHT cohort had 60% fewer deleterious variants in platelet and coagulation genes than expected (Mann-Whitney test  $P = .021$ ). In conclusion, patients with HHT commonly have rare variants in genes of relevance to their phenotype, offering new therapeutic targets and opportunities for informed, personalized medicine strategies.

Submitted 24 January 2022; accepted 21 February 2022; prepublised online on *Blood Advances* First Edition 22 March 2022; final version published online 7 July 2022. DOI 10.1182/bloodadvances.2022007136.

Presented in abstract form at the British Thoracic Society Winter Meeting, November 2021, virtual.

Original data may be found in a data supplement available with the online version of this article. Primary data from the 100 000 Genomes Project are held in a secure

Research Environment and are available to registered users. Please see <https://www.genomicsengland.co.uk/about-gecip/for-gecip-members/data-and-data-access> for further information.

The full-text version of this article contains a data supplement.

© 2022 by The American Society of Hematology. Licensed under Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0), permitting only noncommercial, nonderivative use with attribution. All other rights reserved.

## Introduction

Hereditary hemorrhagic telangiectasia (HHT) is a well-defined endothelial vasculopathy<sup>1</sup> with established clinical diagnostic criteria, known causal genes, and expert consensus informing clinical practice across general and subspecialty management.<sup>2,3</sup> Estimated to affect approximately 1.5 million individuals worldwide,<sup>4,5</sup> HHT usually results from a heterozygous null allele in *ACVRL1*, *ENG*, or *SMAD4* and is transmissible as an autosomal dominant trait.<sup>6</sup> Evidence from early family-based analyses<sup>7-9</sup> has been supported by decades of studies in patients<sup>10,11</sup> and animal models,<sup>12-17</sup> confirming that loss-of-function causes abnormal vascular structures, including telangiectasia and arteriovenous malformations (AVMs).

Telangiectasia and larger AVMs are characterized by defective vascular walls predisposing to hemorrhage. Individuals affected by HHT experience spontaneous, recurrent nosebleeds due to abnormal nasal vasculature and exhibit small, visible telangiectatic vessels that tend to develop on the lips, oral cavity, and finger pads.<sup>2,3</sup> Many have long-term excellent health with modest hemorrhagic losses, although most experience nosebleeds sufficient to lead to iron deficiency anemia unless dietary iron intake is supplemented,<sup>18</sup> and significant proportions require surgical interventional treatments and/or disease-modifying drugs such as antiangiogenic agents.<sup>2,3,19-21</sup> Patients are commonly found to have silent visceral AVMs, and screening/primary prevention strategies have led to improved life expectancy.<sup>22,23</sup>

As is typical in medicine, it remains impossible to predict which patients will develop complications prior to and/or despite preventative treatments. Vascular malformations are known to differ between affected members of the same family sharing the same HHT causal gene,<sup>24</sup> while similar vascular abnormalities can have profoundly different impacts: hemorrhage that varies in severity of recurrent bleeds from nasal or gastrointestinal telangiectasia<sup>2,3</sup> and rare but life-threatening acute hemorrhages from pulmonary or cerebral AVMs.<sup>25,26</sup> Inflammation is not a prominent feature of HHT,<sup>18</sup> and limited available data suggest that iron handling is appropriate.<sup>18</sup> In one study, however, anemia was out of proportion to hemorrhagic iron losses in more than a quarter of consecutive hospital-reviewed patients, with evidence for reduced red blood cell (erythrocyte) survival in nearly half of these.<sup>27</sup> Counterintuitively, in this hemorrhagic disease,<sup>28</sup> age-adjusted incidence rates of venous thromboembolism (VTE) are several-fold higher in HHT than in the general population.<sup>29,30</sup> Additionally, individuals affected by HHT who have VTE are more likely to have a major deep-seated infection that is a particular risk when venous blood can evade pulmonary capillary processing by passage through right-to-left shunts provided by pulmonary AVMs.<sup>31-35</sup> Socioenvironmental exposures can aggravate HHT, as detailed elsewhere,<sup>27,28,32,36-40</sup> but only a proportion of patients with these added risk factors develop the respective clinical phenotypes. There are hints of familial predispositions<sup>27,29,32,36,37,40</sup> supported by case series in which occasional individuals had coincidental non-HHT genetic contributors to their complications.<sup>27,29</sup>

We hypothesized that rare, high-impact variation in genes beyond the primary HHT-causative gene could play a role in disease manifestation and tested in an essentially unselected cohort of HHT patients who had undergone whole genome sequencing (WGS) through the National Health Service (NHS) 100 000 Genomes Project.

## Methods

### Clinical cohort

Patients attending the Hammersmith Hospital/West London Genomic Medicine Centre for management of HHT were recruited to the 100 000 Genomes Project<sup>41,42</sup> if the family's HHT pathogenic variant was not already known<sup>43,44</sup> and they met the clinical diagnosis of HHT.<sup>45</sup> The maximum number of patients was approached, and there was no bias in recruitment to patients with milder or more severe HHT features. Standardized clinical assessments included a detailed clinical history specifically focusing on aspects of the HHT phenotypes, any disproportionate or unexpected symptomatic manifestations, and hematological/biochemical assessments in hospital service laboratories as described.<sup>6,27,31,46,47</sup>

### WGS

The 100 000 Genomes Project was set up by the UK Department of Health and Social Security in 2013 to sequence whole genomes from NHS patients recruited through 1 of 13 Genomic Medicine Centres. Patient/family DNA sampling was accompanied by phenotypic data entries and integration with NHS health records.<sup>41</sup> HHT was successfully nominated for inclusion in 2015. During data model set up, human phenotypic ontology terms<sup>48</sup> were defined to capture the full range of phenotypes observed in patients with HHT, totaling 301 different terms spanning disease manifestations, accessory, and incidental phenotypes.<sup>6,48</sup> For each patient with HHT recruited between September 2016 and August 2019, relevant human phenotypic ontology terms were submitted to Genomics England after patient recruitment using Open Clinica and Genie.<sup>42</sup>

Genomics England performed all WGS using the Illumina WGS Service Informatics pipeline for sequence alignments to the National Center for Biotechnology Information Genome Reference Consortium Human (GRCh) build 38/hg38<sup>49</sup> and variant calling as described.<sup>42,43</sup> Anonymized outputs from Genomics England internal bioinformatics pipelines and analyses were examined in the Genomics England Research Environment through LabKey.<sup>42</sup>

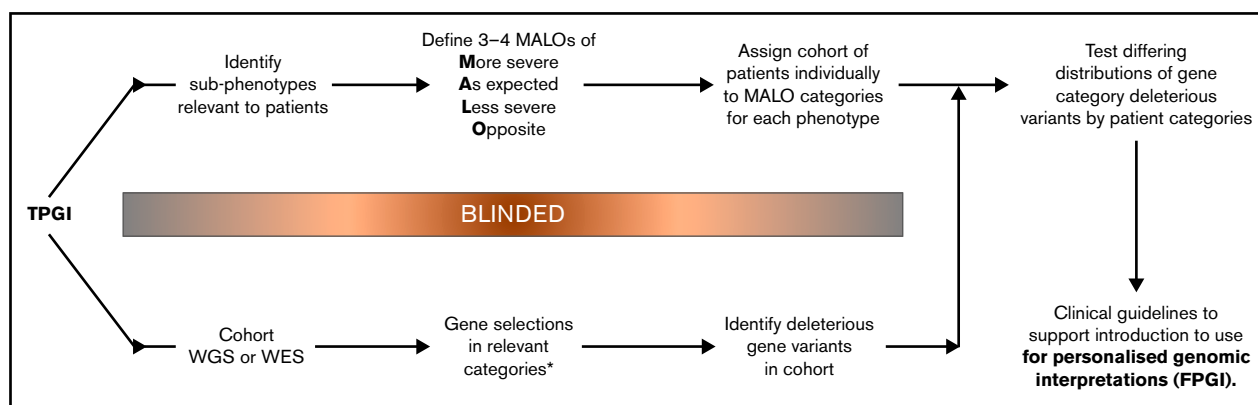
### Analytic framework

To augment numeric and clinical-level confidence, we blindly categorized the 104 patients with HHT and their WGS data in what we term a "toward personalized genomic interpretation" approach (Figure 1). This incorporates phenotypic severity scales to maximize information from small subgroups by generating scales in which individual patients are classified for each subphenotype as **More** severe than expected, **As** expected, and (if feasible) **Less** severe than expected, or **Opposite** phenotype (MALO).

### Disease subphenotype assignments

Based on updated clinical information on the recruited patients by 28 April 2021, continuous variables and severity scales as currently used in our service<sup>6,18,27,46,47</sup> were assigned independently to all recruited patients, blinded to genotypes beyond the causal HHT variants emerging through clinical diagnostic pipelines<sup>43,44</sup>:

- HHT-related bleeding was categorized as H0 (as expected for HHT); H1 (precipitated excessive bleeds, ie, severe bleeds precipitated by a pharmaceutical agent,<sup>21,28,40,51</sup> or pregnancy-associated pulmonary AVM hemorrhage<sup>25,52</sup>); and H2 (spontaneous excessive bleeds). H2 was assigned for severe daily



**Figure 1. Framework for studies “toward personalized genomic interpretations” (TPGI).** MALO: **M**ore severe than expected for disease; **A**s expected for disease; **L**ess severe than expected for disease, and **O**pposite phenotype. \*Gene selection by any broadly relevant set of panels such as PanelApp established for 100 000 Genomes Project interpretations<sup>50</sup> suggested to include additional categories less relevant to phenotypes (as in current study) to control for methodological biases. Where TPGI demonstrates evidence of relevance across a patient cohort, this supports attention to a subsequent “for personalized genomic interpretations” (FPGI) stage to be developed within mainstream medicine. WES, whole exome sequencing; WGS, whole genome sequencing.

nosebleeds or gastrointestinal bleeds resulting in hemorrhage-adjusted iron requirements<sup>18</sup> exceeding replacement feasibility using oral iron supplementation (resulting in intravenous iron and/or blood transfusion dependency)<sup>27</sup> or for a major spontaneous hemorrhage from a pulmonary<sup>53,54</sup> or cerebral<sup>26</sup> AVM.

- Anemia was categorized as A0 (within expected range for hemorrhagic losses and iron intake<sup>18,27</sup>) and A1 (anemia out of proportion to the degree of hemorrhage based on hemorrhage-adjusted iron requirements<sup>18</sup> and total iron intake from the diet and therapeutic oral or intravenous iron, as detailed by Thielemans et al<sup>27</sup>).
- The VTE categories were T0 (no VTE), T1 (drug, perioperative or otherwise precipitated VTE/pulmonary embolus), and T2 (spontaneous VTE/pulmonary embolus, using definitions as in Livesey et al<sup>30</sup>).
- Deep-seated infection categories were I0 or I1, with the deep-seated infection type listed for I1.<sup>31-33</sup>

### Candidate modifier gene selection and analysis

Genes were primarily selected based on causal influences on coagulation, hemorrhage, and/or red blood cell (erythrocyte) survival exposed in the general population by bleeding disorders or congenital hemolytic anemias (supplemental Table 1).<sup>55,56</sup> High-impact deleterious DNA variants are ill-suited to usual genome-wide and statistical associations due to allele rarity.<sup>57-60</sup> To enhance statistical power yet retain a focus on variants most likely to be of clinical significance, the 75 selected genes were categorized by their most recognized role. These spanned the coagulation cascade, platelet biology, hemoglobin production, red cell (erythrocyte) membrane production, and red cell enzymes.

### Variant-level analysis

Genes with variants were plotted using Circos<sup>61</sup> on the Galaxy platform<sup>62</sup> to generate a circular ideogram of GRCh38/hg38.<sup>49</sup> The burden of usual genetic variation was compared between genes and gene categories using the gene damage index<sup>63</sup> and by genetic intolerance measured by the residual variation intolerance score.<sup>64</sup>

Variant deleteriousness was ranked by Combined Annotation Depletion (CADD) scores.<sup>65-67</sup> Before examining the selected gene variants, a CADD score of >15 was selected as a threshold to indicate likely

deleteriousness because it is commonly used in this regard,<sup>65,68</sup> and in other studies, it performed better than other single measures of deleteriousness using the mutation significance cutoff (MSC) gene-level thresholds for variant-level predictions.<sup>68</sup> For the current study, CADD scores were extracted from the University of Washington CADD server.<sup>65-67</sup> Study methodology precluded integration within American College of Medical Genetics clinical care frameworks.<sup>59,60</sup> Generally, variants were categorized by CADD score (>15: likely deleterious, incorporating start loss, stop gain, frameshift, splice donor/acceptor, and some missense and inframe insertion/deletion variants; 5-15: uncertain significance; <5: likely benign, comprising most synonymous, 5' untranslated region, splice region and intronic variants). Gene-level adjustments used the MSC 90% and 95% scores,<sup>68</sup> with outlier gene *SPTA1* tested in both uncertain significance and likely deleterious categories.

Variant number, allele frequencies, and CADD scores and, separately, categorized patient numbers (converted to new codes to prevent inadvertent unblinding) were approved for export through the Research Environment AirLock under subproject RR42 (HHT-Gene-Stop).

### Population-level analysis

As a comparator population, from the gnomAD 3.1.2 database comprising 76 156 genomes and exomes,<sup>69</sup> for each of the 75 study genes we extracted the total number of variants and variants flagged by molecular type as predicted loss-of-function (pLOF). We did not include other variants listed by ClinVar<sup>70</sup> as likely pathogenic/pathogenic. gnomAD and study variant rates were adjusted for the gnomAD-analyzed sequence length of the main protein coding transcript (corresponding to the curated RefSeq transcript from the UCSC Genome Browser<sup>71,72</sup>), adding 75 bp of exon-flanking sequence<sup>69</sup> for each exon. In view of the differing approaches to classifications, it was not appropriate to directly compare rates of pLOF variants in gnomAD 3.1.2 and CADD-defined deleterious variants in the study population. Instead, differential rates were compared between the categories of hematologic genes, enabling less relevant categories to control for methodological sources of bias.

## Data analysis

Summary and comparative statistics were generated using STATA IC versions 15.0 and 16.0 (Statacorp, College Station, TX) and GraphPad Prism 9 (GraphPad Software, San Diego, CA). One-way differences between categories were compared by  $\chi^2$  test and between continuous variables using Kruskal-Wallis with Dunn's multiple comparison test for preselected pairwise comparisons, in which  $P = .0285$  was significant at false discovery rate = 0.05 level.<sup>73,74</sup> Two-way analyses incorporating both data source and gene categories were performed using a 2-way analysis of variance. The strength and direction of association between ranked variables were examined using Spearman's rank-order correlation. Data were also represented using graphics generated in GraphPad Prism 9 (GraphPad Software) and STATA IC version 16.0.(Statacorp).

## Ethical approval

The Project received ethical approval from the HRA Committee East England-Cambridge South (REC Ref 14/EE/1112), and all individuals provided informed, written consent.

## Results

### Cohort demographics

One hundred four individuals with HHT were recruited from a single Genomic Medicine Centre, and all were included in the analyses. They were from 89 families, with 1 to 3 (median 1) affected individuals per family. At the time of recruitment, patient ages ranged from 17 to 87 years (median 50 years) and 66 (64%) were female; 56 of 89 (63%) families were found to have a heterozygous variant in the genes currently on HHT diagnostic testing panels (Figure 2A). In keeping with the "HHT gene negative" status of a subgroup of families on recruitment and ongoing analyses, in 33 of 89 (37%) families, no variants had been identified at the time of the current study.

Blinded to the variants identified in hematological genes, the recruited patients had been assigned to relevant scales for more extreme phenotypes (Figure 2B). Notably, in this group of patients selected purely due to lack of prior genetic test/absence of a known family HHT genotype and attendance at the UK reference center during recruitment to the 100 000 Genomes Project, 36/104 (35%) were categorized with severe hemorrhage, 29/104 (28%) with disproportionate anemia, and 8/104 (8%) had venous thromboembolic disease. Deep-seated infections affected 13/104 (13%).

There was no difference in respective placements on the 3 axes (hemorrhage-thrombosis, anemia or infection) for patients with *ACVRL1* or *ENG* variants (Figure 2C).

### Hematological gene-level variation

Within the cohort of 104 HHT patients, 105 variants were identified in the 75 selected hematological genes (supplemental Table 1). At least 1 variant was present in 52/104 (50%) of the study cohort, and 56/75 (75%) of genes had at least 1 variant identified in the study cohort. Variants were identified in genes on 21 chromosomes and were independent of the causal HHT gene (Figure 3A-B).

The variant burden was not specific to the study cohort because the gene damage index<sup>63</sup> was also greater in the genes harboring variants in the study cohort (Figure 3Ci), providing evidence that these genes are more damaged in the general population. There

was no difference between the variant and nonvariant genes for residual variation intolerance (Figure 3Cii).<sup>64</sup>

### Variant deleteriousness

Variant CADD scores ranged from 1 to 42 (median 15; Figure 4A). Fifty-seven of 105 variants (54.3%) had CADD scores >15, indicating likely deleteriousness. At least 1 variant with a CADD score >15 was present in 38/104 (36.5%) of the study cohort. There was only 1 variant with a CADD score >15 that was identified in 2 members of the same family (an *ACVRL1* family).

Across all genes with variants, the CADD score delimiter of 15 captured the median 90% MSC threshold (Figure 4A), whereas the CADD score delimiter of 5 captured the 90% MSC threshold of all except 1 gene (Figure 4A). Variant overrepresentation in the red cell membrane production category reflected larger genes (supplemental Table 1) and was not apparent when restricting to CADD >15 variants (Figure 4B). All variants were rare, with median gnomAD 3.1.1<sup>69</sup> allele frequency <0.05% for each of the 5 categories and all gnomAD allele frequencies <0.003 (Figure 4B).

### Phenotypic categories

When associations of gene variants with clinical phenotypes were examined, CADD >15 (likely deleterious) variants were identified in all severe hemorrhage, disproportionate anemia, spontaneous thrombosis, and severe deep-seated infection categories. Notably, CADD >15 variants were also present in each of the 4 "no event" categories (Figure 4C). Hemoglobin CADD >15 gene variants tended to distribute toward the "no event" categories (Figure 4C, lower left pie charts). By binary categories that pooled precipitated and spontaneous events, visual inspection suggested that erythrocyte membrane gene variants might be enriched in the 11 patients with severe deep-seated infections (Figures 4C, lower middle pie charts, and 5D). Similarly, visual inspection suggested that erythrocyte enzyme variants might be enriched in the 6 patients with unprovoked VTE (Figure 4C, lower right pie charts, and D). However, these were based on small numbers and unsuited to statistical examination.

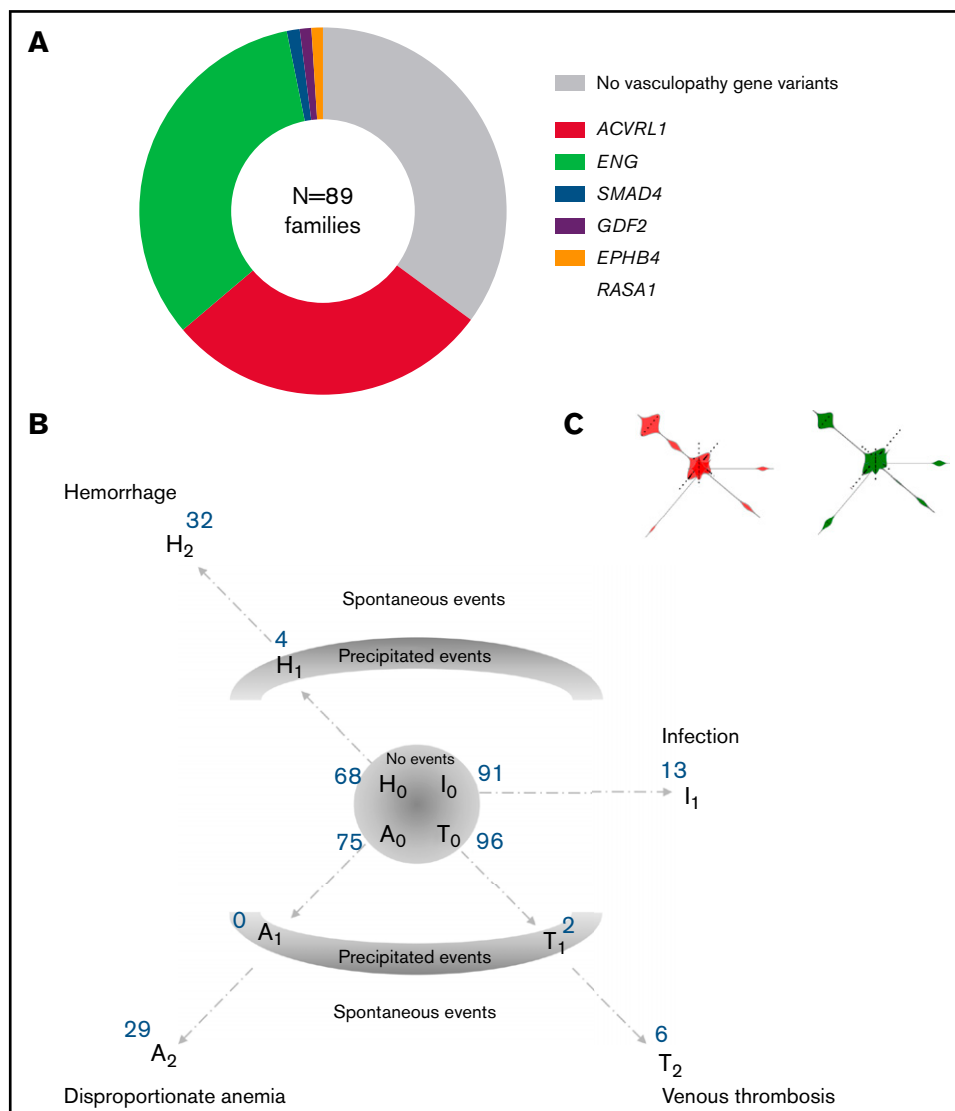
The hemorrhage and thrombosis phenotypes provided greater statistical power because they can be considered to represent opposing ends of a spectrum. There were trends for the higher CADD score variants in both coagulation (Figure 4C, upper middle pie charts) and platelet (Figure 4C, upper right pie charts) genes to be distributed toward the more severe hemorrhagic categories. Examining 3-point scales of severe hemorrhage/usual hemorrhage thrombosis, using nonparametric Spearman's correlations, CADD >15 (likely deleterious) variants were enriched in patients with more severe hemorrhage for both platelet ( $\rho = 0.25$ ;  $P = .008$ ) and coagulation ( $\rho = 0.21$ ;  $P = .024$ ) gene categories (Figure 4D).

### Quantitative trait analyses

While theoretically unsurprising that deleterious variants in platelet and coagulation genes would be associated with more severe HHT hemorrhage, this has not been previously acknowledged in HHT standards of care,<sup>2,3</sup> prompting examination of further supporting evidence.

First, we examined associations between hemorrhage and red cell/coagulation quantitative traits in the cohort, comparing patients with no variants or CADD <5 variants (likely benign); CADD 5 to 15 variants (uncertain significance), and CADD >15 variants (likely





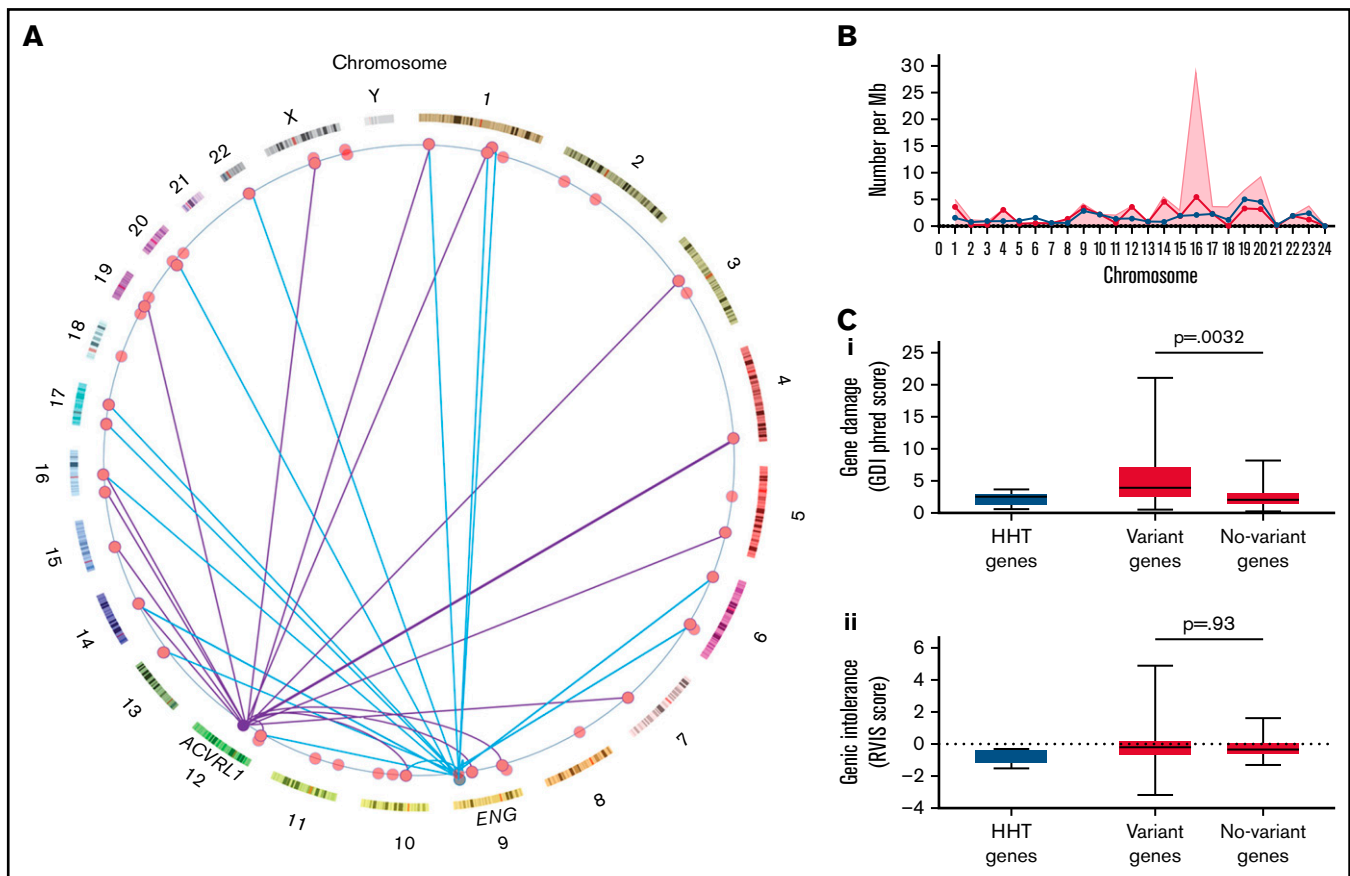
**Figure 2. HHT genes and phenotypic associations.** (A) Endothelial vasculopathy genes and variants identified in current study cohort. Note that the major HHT genes are *ACVRL1* (red) and *ENG* (green).<sup>6-8</sup> *SMAD4*<sup>6,9</sup> is less common, and *GDF2* was more recently described as HHT causal,<sup>44</sup> whereas *EPHB4* and *RASA1* cause separate endothelial vasculopathies (CM-AVM2 and CM-AVM1) that overlap phenotypically with HHT. (B) Schematic of the cohort of 104 patients plotted on 4 separate axes for hemorrhage (H), anemia (A), thrombosis (T), and deep-seated infection (I). Ten deep-seated infections were in association with concurrent and presumed causative pulmonary AVMs (cerebral and spinal abscesses, spinal discitis, and septic arthritis due to polymicrobial flora, particularly anaerobic and aerobic commensals of the gastrointestinal and periodontal spaces<sup>31-33</sup>). Blue numbers indicate the number of the cohort with no events/no excess (H<sub>0</sub>, A<sub>0</sub>, T<sub>0</sub>, and I<sub>0</sub>), precipitated events (H<sub>1</sub>, A<sub>1</sub>, T<sub>1</sub>), and spontaneous events (H<sub>2</sub>, A<sub>2</sub>, T<sub>2</sub>, and I<sub>1</sub>). I<sub>1</sub> were cerebral abscess (N = 6), spinal discitis (N = 2), spinal abscess (N = 1), septic arthritis (N = 1), recurrent bacterial endocarditis (N = 1), osteomyelitis (N = 1), and recurrent sepsis (N = 1). (C) Violin plots for *ACVRL1* (left, red) and *ENG* (right, green) patients orientated on the 3 axes as in (B).

deleterious; Figure 5). For hemoglobin (Figure 5A), red cell membrane (Figure 5B), and red cell enzyme (Figure 5C) genes, there was no discernible difference in any measured red cell index by the presence of variants: the median of each category was within the respective normal range. Similarly, for coagulation genes, there was no discernible difference in prothrombin time or activated partial thromboplastin time according to the presence of a variant (Figure 5D). The median value did lie outside of the normal range for platelet counts in individuals who had variants with platelet gene variants with CADD scores between 5 and 15 (Figure 5Ei), but mean platelet volume did not differ (Figure 5Eii), and the distinction was lost combining with higher impact variants (data not shown). We concluded that the variants

were not clearly reflected in variability within patient quantitative traits at specifically sampled times.

### Variant enrichment analyses

An association of deleterious variants in platelet and coagulation genes with more severe hemorrhage in patients with HHT predicts that deleterious variants in these genes may be less well tolerated and relatively depleted from HHT patient gene pools. To test, we contrasted platelet and coagulation gene variant burdens with the red cell genes in which variants had not displayed clear phenotypic associations and compared across the 5 categories to control for methodological bias.



**Figure 3. Hematological genes and variants.** (A) Genome positions of the variant-containing genes in the study cohort on genome ideogram<sup>61,62</sup> indicating positions of the 56 study genes with variants (red) and the two major HHT genes *ACVRL1* and *ENG*. Variants present in patients with an identified variant in HHT genes are indicated for *ACVRL1* (purple lines) and *ENG* (blue lines). To preserve anonymity, data for the single *SMAD4*, *GDF2*, and *EPHB4* families are not illustrated. (B) Chromosomal distributions of variant-containing genes per 100 Mb of gDNA (blue circles/lines), number of variants per Mb (red shaded background), and variants with CADD score >15 (red circles/lines) per Mb of gDNA by chromosome.<sup>49</sup> (C) Population-level burden of genetic damage in the 6 HHT panel genes (blue) and 75 study hematological genes (red), as detailed in supplemental Table 1. The study genes were subcategorized by variant presence (Variant genes; N = 56) and absence (No-variant genes; N = 19) in the study cohort, and P values were calculated by Dunn's post Kruskal Wallis: (Ci) Gene damage index (GDI) phred scores, a genome-wide, gene-level metric of the mutational damage that has accumulated in the general population that performs well at removing exome variants in genes irrelevant to disease.<sup>63</sup> (Cii) Residual variation intolerance scores (RVIS) that perform better for detection of genes in which newly identified variants are more likely cause a recognized disease.<sup>64</sup>

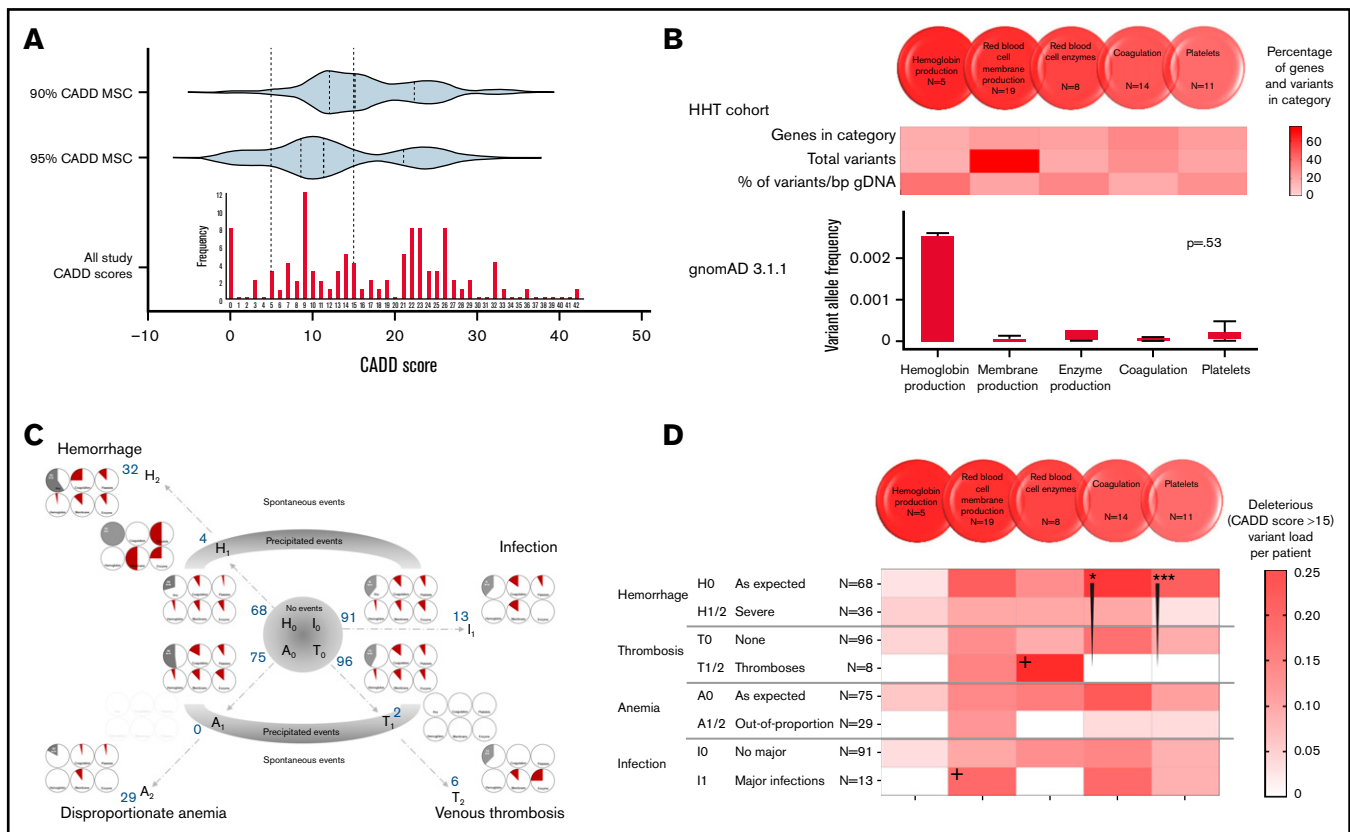
While all variants were identified at similar respective frequencies in both study and gnomAD populations (Figure 6A), there was a median 19-fold (range 10.5- to 28.7-fold) enrichment for study cohort CADD >15 variants compared with the gnomAD pLOF deleterious variants, which exclude missense, inframe indel, splice region, and other likely pathogenic or pathogenic variants (Figure 6B). Enrichments appeared to differ according to gene categories (Figure 6B-C). Restricting to the 36 genes in which deleterious variants were defined in both cohorts and comparing to red cell genes standardized to 1.00 (95% CI, 0.24-1.77), the HHT burden of CADD >15 variants in coagulation and platelet genes was 0.40 (95% CI, 0.22-0.58; Mann-Whitney test  $P = .021$ ; Figure 6D).

## Discussion

We have shown for 75 hematological genes in which variants can cause heritable human disease, that deleterious variants are commonly found within populations. Despite each variant individually being rare (allele frequency <0.3%), overall, more than a third of

unselected patients with HHT, an inherited vasculopathy that places them at risk of hemorrhage and anemia, had at least 1 likely deleterious variant in diverse genes that could augment or modify these disease complications. Patients with more severe hemorrhage that had been attributed solely to their HHT had a higher number of rare, deleterious variants in unrelated platelet and coagulation genes located across the genome. Additionally, there were fewer deleterious variants in platelet and coagulation genes than expected, supporting an interpretation that rare variation leading to a propensity for higher degree of bleeding may not be as well tolerated in HHT as in the general population.

Most attention on the potential of DNA sequencing to personalize medicine focuses on discovery of novel genetic causes of disease<sup>75,76</sup> and development of polygenic risk scores from large-scale studies identifying common variation across blocks of genes co-inherited in linkage disequilibrium<sup>77</sup> that can direct attention to novel *cis*-regulatory elements.<sup>78,79</sup> The presented approach offers an alternate and accelerated route to the clinic by proceeding from genes



**Figure 4. Burden of deleterious variants by categories of genes and phenotypic severity profiles.** (A) The CADD score distribution of variants identified in the study cohort (red) compared with gene-level MSC scores,<sup>68</sup> indicating the lower limit of the 90% and 95% confidence intervals for deleterious CADD scores in individual genes (gray violin plots). Vertical dotted lines represent overall CADD scores 5 and 15 and the violin plot median and interquartile ranges. Note that CADD scores rank the deleteriousness for all 9 billion single nucleotide variants, and millions of small indels and splice site variants, based on machine learning trained on diverse genomic features derived from surrounding sequence context, gene model annotations, evolutionary constraint, epigenetic measurements, and functional predictions.<sup>65-67</sup> (B) The 5 process-level categories, respective category contributions in the study cohort, are shown for the total number of genes, and all variants, with each box representing the category's percentage for the row, as indicated by the heat scale. (C) The burden of deleterious (CADD >15) variants in the 5 categories of genes for each predefined phenotypic subcohort. Filled pie-chart regions quantify the number of CADD >15 variants per patient in the category (gray if pooling across all gene categories; red if gene category specific; fully filled if 1 variant per patient). The 10 variants with CADD scores of 1 to 5 (likely benign) and 36 with CADD scores between 5 and 15 (of uncertain significance) are effectively included in the remaining white pie chart regions to enhance clarity. For each set of 6 pie charts, gene category placements are consistent. The upper row indicates total number of CADD >15 variants across all categories, and the remaining upper row pie charts focus on hemorrhage (upper middle for coagulation genes, upper right for platelet genes). The lower row presents the red cell gene categories (lower left for hemoglobin genes, lower middle for erythrocyte [red cell] membrane genes, and lower right for erythrocyte enzyme genes). The set of 6 gray-lined empty pie charts indicate there were no patients in the predefined A1 category. (D) Heat maps for number of variants per patient in each phenotypic subcohort, in which the precipitated and spontaneous H1/H2, T1/T2, and A1/A2 subcohorts have been pooled. Trends highlighted in the text are denoted by a white cross or black star. \* $P < .0286$ ; \*\* $P < .01$ .

already known to have functional impact in processes potentially relevant to the patient. Although a small proportion of rare deleterious variants causes recognized human diseases, most do not, either because they are not present at sufficient zygosity or because they are not in an appropriate patient/environment milieu to be exposed above an essentially compensated perturbation to normal physiology. Major databases emphasize the very high rates of rare deleterious variants in the human general population<sup>69</sup> as well as the necessity to use novel analytic methods<sup>57</sup> in comparison with more common variant analyses that more readily meet genome-wide significance scores in genomes and transcriptomes.<sup>78,79</sup> The current study offers an indicative example of how aberrant physiology imposed by an endothelial vasculopathy appears to augment the impact of variants in separate systems that might otherwise be

clinically insignificant. More generally, the approach can be adopted in a framework of personalized genomic interpretations (PGIs) either in "toward PGI" studies as in the current manuscript to provide statistical support or "for PGI" studies to deliver what is genuinely personalized (N = 1 medicine) (Figure 1).

Strengths of the study include WGS of a respectable number of patients with HHT with detailed phenotypic characterization, spanning a range of severity phenotypes. Unusually for a hospital-recruited population, not all were referred because of personally severe features, with some presenting through family-based screening programs. Nevertheless, a weakness of the study is that the group did have a bias toward a hospital presentation, and we cannot rule out that the burden of deleterious rare variants would be lower in a community-recruited population, potentially with milder HHT phenotypic spectra. That said,

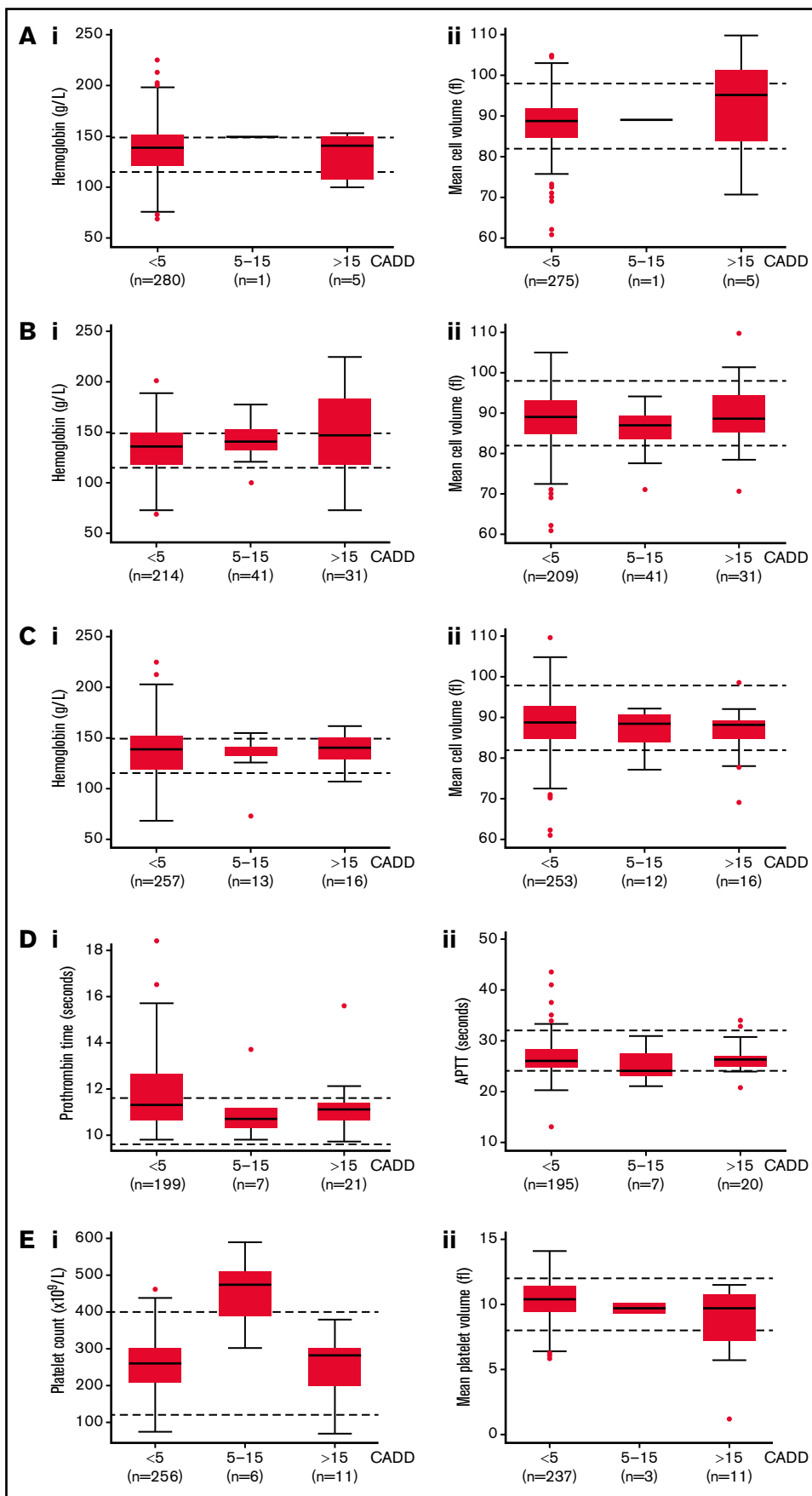
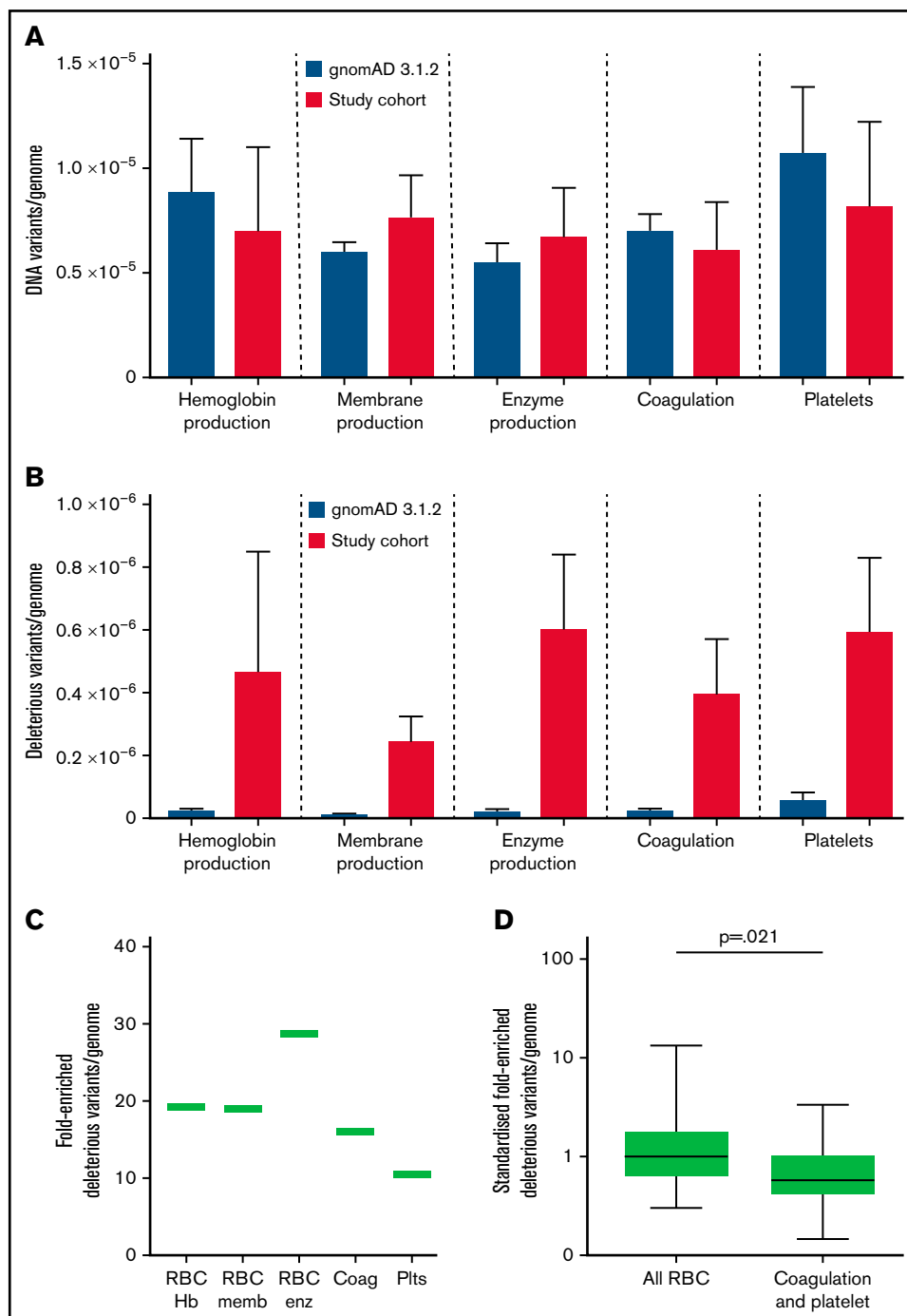


Figure 5.





**Figure 6. Variant burden in HHT cohort compared with general population.** (A) Number of variants in the 75 study genes by category, within the gnomAD 3.1.2 dataset (blue) and current cohort (red) per genome. (B) gnomAD 3.1.2 dataset loss of function (pLOF) variants compared with CADD >15 genes in study cohort per genome, with mean and standard error illustrated. (C) Fold-enriched total number of deleterious variants per category, in all 75 genes, for HHT CADD >15 variants compared with pLOF variants in gnomAD 3.1.2 dataset.<sup>69</sup> (D) Comparison of fold-enriched deleterious variants limiting to HHT CADD >15 variants and gnomAD pLOF variants in the same genes. *P* value calculated by Mann-Whitney test.

**Figure 5 (continued) HHT patient subcategories and variation.** Quantitative phenotypic measurements categorized by presence, absence, and CADD score of variants. Data points represent all datasets captured in the 104 patients across their clinical assessments, plotted according to the presence or absence of variants in the genes for (A) hemoglobin production, (B) red blood cell membrane production, (C) red blood cell enzymes, (D) coagulation and (E) platelets as listed in supplemental Table 1. Two indices are provided for each category of genes, but there were also no differences observed between categories for other indices examined (data not shown).

the cohort was not recruited to the 100 000 Genomes Project because of a bleeding, anemia, thrombotic, or infective phenotype but instead because they met diagnostic criteria for a separate disease, the endothelial vascular dysplasia of HHT. As such, they provide an important window on the potential for similar scale, WGS-based high-definition analyses in other diseases in which phenotypes could be modified by concurrent, otherwise minor diatheses.

All variants identified were rare, with minor allele frequencies <0.5%. Figure 3A indicates that with 1 exception (*AKI*), there was no reason for variants to segregate with HHT because almost all variant-containing genes were widely separated from the patient's HHT gene, usually on a different chromosome. Study findings could be extended by incorporating additional genes and categories, and in later individual-level considerations, it may be helpful to distinguish genes associated with dominant phenotypes (such as *ANK1* and *SPTB*) from others in which heterozygous variants are recessive. Further, it is not the place of the current study to explore how these and the wider variants might change HHT clinical practice. Such considerations with patient-level integrations are underway through the NHS Rare Disease Collaborative Network for HHT and Genomics England Clinical Interpretation Partnership (Figure 1). That said, some overview comments are appropriate:

- 11/104 patients with HHT were heterozygous for a total of 11 deleterious variants (10 unique variants, 10 different families), in genes in which deficiency is known to impact on platelet function leading to bleeding disorders. Furthermore, these variants were found more commonly in patients defined as having more severe hemorrhage either spontaneously or precipitated, for example, by pregnancy, antiplatelet agents, anticoagulants, or iron. Currently no attention in HHT management focuses on platelet disorders, yet the current study suggests that such considerations alongside modifications to commonly prescribed drugs may be relevant to at least 10% of patients with HHT.
- Similarly, 13/104 patients with HHT were heterozygous for a total of 14 deleterious variants (all unique variants and in different families) in genes in which deficiency is known to impact on the coagulation cascade. The data suggest that this may be relevant to >12% of patients with HHT.
- 24/104 patients with HHT were heterozygous for a total of 32 rare deleterious variants (all unique) in genes in which deficiency is associated with inherited hemolytic anemias due to an increased rate of red cell destruction. Again, currently no attention in HHT anemia management focuses on extending the lifespan of endogenous erythrocytes because they suffer additional stresses on passage through HHT vasculature. Cross-over strategies with hemolytic anemias may be relevant to 1 in 4 patients with HHT, with particular attention directed to the subgroup of patients with HHT currently receiving intense, intravenous iron regimens on a weekly or near-weekly basis.<sup>2,16,19,20</sup>

The data suggest that there was no a priori reason why the study cohort had their rare, deleterious variants. Instead, comparison with population-level databases predict comparable burdens in other patient populations, drawing attention to new foci for research. For example, whereas genetic variants predisposing to pathogenic venous thromboses are usually assumed to encode components and mediators of the coagulation cascade, the current cohort suggested a possible trend with erythrocyte enzyme

variants, and there has been some discussion of thrombosis in red cell enzymopathy fields.<sup>80</sup> Opportunities to test current associations will be enhanced if enzymopathy genes are included in gene panels for VTE and further scientific examination appears warranted. Similarly, the observed enrichment of deleterious erythrocyte membrane gene variants in patients experiencing severe deep-seated infection warrants further study in the setting of the common (daily) bacteremias that can seed abscesses by currently unknown mechanisms.<sup>31,81</sup> Mechanistic examination of whether erythrocyte membrane changes modify macrophage opsonization rates<sup>82</sup> may influence clinical risk assessments, resulting in extension or restriction of HHT patient cohorts for whom prophylactic antibiotics are recommended.<sup>3,32,33,53</sup> Notably, in both the study cohort and the gnomAD 3.1 population, the greatest proportion of variants was for the red cell membrane category, in which relevance to evolutionary fitness could be postulated.

We conclude that the potential to unmask pathophysiologically relevant processes is augmented by categorizing patients using severity scales that additionally influence risk-benefit considerations inherent in therapeutic provision. Accelerated pathways to more informed clinical trial design and personalized medical care can be the expected outcome of WGS when analytics focus on rare deleterious variants for secondary clinical phenotypes that matter to patients and health care services.

## Acknowledgments

The authors thank the National Health Service staff of the West London Genomic Medicine Centre and the families for their willing participation.

This research was made possible through access to the data and findings generated by the 100 000 Genomes Project. The 100 000 Genomes Project is managed by Genomics England Limited (a wholly owned company of the Department of Health and Social Care). The 100 000 Genomes Project is funded by the National Institute for Health Research (NIHR) and NHS England. The Wellcome Trust, Cancer Research UK, and the Medical Research Council have also funded research infrastructure. The 100 000 Genomes Project uses data provided by patients and collected by the National Health Service as part of their care and support.

This research was co-funded by the NIHR Imperial Biomedical Research Centre. The views expressed are those of the authors and not necessarily those of funders, the NHS, the NIHR, or the Department of Health and Social Care.

## Authorship

Contribution: K.E.J. co-developed the project design, selected gene categories and genes, performed all data analysis within the Genomics England Research Environment (including variant frequency assignments in gnomAD 3.1), exported anonymized datasets for numerical interrogation, retrieved all CADD scores, performed quantitative trait association studies, contributed content to Figures 4 and 5, and contributed to manuscript writing; E.O. and S.B. assisted with patient evaluations; F.N., K.O., and K.F. recruited patients; M.S. and G.A.T. processed patient samples; G.A.T., T.F., J.R., and C.L.S. contributed to project set up at West London Genomic Medicine Centre and patient recruitment; C.M.M. contributed to coagulation discussions; N.C. and D.M.L. contributed to red cell discussions; M.J.C. set

up and F.B.-P. facilitated anonymized analyses within the Genomics England Research environment; Genomics England Research Consortium performed WGS, sequence alignments, and variant calling; C.L.S. set up the HHT-specific projects with Genomics England, the Respiratory Clinical Interpretation Partnership (GeCIP), and at West London Genomic Medicine Centre, reviewed the patients, categorized phenotypic severities, co-developed the project design, extracted gnomAD 3.1.2 and UCSC data, performed all remaining data analyses, generated the figures, and wrote the manuscript; K.E.J., C.M.M., N.C., D.M.L., and C.L.S. contributed to manuscript revisions; all authors reviewed and approved the final manuscript; and C.L.S. is responsible for the overall content as guarantor.

The complete list of the members of the Genomics England Research Consortium appears in "Appendix."

Conflict-of-interest disclosure: The authors declare no competing financial interests.

ORCID profiles: K.E.J., 0000-0002-5951-3383; K.F., 0000-0003-4031-6029; M.S., 0000-0003-3251-2986; T.F., 0000-0001-5319-3318; J.R., 0000-0002-7102-0071; F.B.-P., 0000-0001-8459-4739; M.J.C., 0000-0001-9295-3594; C.L.S., 0000-0001-9007-5775.

## References

1. Shovlin CL. Hereditary haemorrhagic telangiectasia: pathophysiology, diagnosis and treatment. *Blood Rev.* 2010;24(6):203-219.
2. Faughnan ME, Mager JJ, Hetts SW, et al. Second international guidelines for the diagnosis and management of hereditary hemorrhagic telangiectasia. *Ann Intern Med.* 2020;173(12):989-1001.
3. Shovlin CL, Buscarini E, Sabbà C, et al. The European Rare Disease Network for HHT Frameworks for management of hereditary haemorrhagic telangiectasia in general and speciality care. *Eur J Med Genet.* 2022;65(1):104370.
4. Bideau A, Plauchu H, Brunet G, Robert J. Epidemiological investigation of Rendu-Osler disease in France: its geographical distribution and prevalence. *Popul.* 1989;44(1):3-22.
5. Kjeldsen AD, Vase P, Green A. Hereditary haemorrhagic telangiectasia: a population-based study of prevalence and mortality in Danish patients. *J Intern Med.* 1999;245(1):31-39.
6. Shovlin CL, Simeoni I, Downes K, et al. Mutational and phenotypic characterization of hereditary hemorrhagic telangiectasia. *Blood.* 2020;136(17):1907-1918.
7. McAllister KA, Grogg KM, Johnson DW, et al. Endoglin, a TGF- $\beta$  binding protein of endothelial cells, is the gene for hereditary haemorrhagic telangiectasia type 1. *Nat Genet.* 1994;8(4):345-351.
8. Johnson DW, Berg JN, Baldwin MA, et al. Mutations in the activin receptor-like kinase 1 gene in hereditary haemorrhagic telangiectasia type 2. *Nat Genet.* 1996;13(2):189-195.
9. Gallione CJ, Repetto GM, Legius E, et al. A combined syndrome of juvenile polyposis and hereditary haemorrhagic telangiectasia associated with mutations in MADH4 (SMAD4). *Lancet.* 2004;363(9412):852-859.
10. Bourdeau A, Cymerman U, Paquet ME, et al. Endoglin expression is reduced in normal vessels but still detectable in arteriovenous malformations of patients with hereditary hemorrhagic telangiectasia type 1. *Am J Pathol.* 2000;156(3):911-923.
11. The HHT Mutation Database: <http://www.arup.utah.edu/database/HHT>. Accessed 23 December 2021.
12. Bourdeau A, Dumont DJ, Letarte M. A murine model of hereditary hemorrhagic telangiectasia. *J Clin Invest.* 1999;104(10):1343-1351.
13. Park SO, Wankhede M, Lee YJ, et al. Real-time imaging of de novo arteriovenous malformation in a mouse model of hereditary hemorrhagic telangiectasia. *J Clin Invest.* 2009;119(11):3487-3496.
14. Jin Y, Muhl L, Burmakin M, et al. Endoglin prevents vascular malformation by regulating flow-induced cell migration and specification through VEGFR2 signalling. *Nat Cell Biol.* 2017;19(6):639-652.
15. Ola R, Künzel SH, Zhang F, et al. SMAD4 prevents flow induced arteriovenous malformations by inhibiting casein kinase 2. *Circulation.* 2018;138(21):2379-2394.
16. Bernabeu C, Bayrak-Toydemir P, McDonald J, Letarte M. Potential second-hits in hereditary hemorrhagic telangiectasia. *J Clin Med.* 2020;9(11):3571.

Correspondence: Claire L. Shovlin, National Heart and Lung Institute, Imperial Centre for Translational and Experimental Medicine, Imperial College London, Hammersmith Campus, Du Cane Rd, London W12 0NN, UK; e-mail: c.shovlin@imperial.ac.uk

## Appendix

The members of the Genomics England Research Consortium are: J. C. Ambrose, P. Arumugam, R. Bevers, C. R. Boustred, H. Brittain, G. C. Chan, T. Fowler, A. Giess, A. Hamblin, T. J. P. Hubbard, R. Jackson, M. Kayikci, A. Kousathanas, L. Lahnstein, S. E. A. Leigh, I. U. S. Leong, F. J. Lopez, F. Maleady-Crowe, M. McEntagart, F. Minnici, N. Murugaesu, P. O'Donovan, C. A. Odhams, D. Perez-Gil, M. B. Pereira, J. Pullinger, T. Rahim, A. Rendon, T. Rogers, K. Savage, K. Sawant, R. H. Scott, A. Siddiq, A. Sieghart, S. C. Smith, A. Stuckey, M. Tanguy, A. L. Taylor Tavares, S. R. Thompson, M. J. Welland, E. Williams (Genomics England, London, UK); F. Boardman-Pretty, M. A. Brown, M. J. Caulfield, S. Henderson, L. J. Jones, D. Kasperaviciute, L. Moutsianas, M. Mueller, A. C. Need, C. Patch, A. Sosinsky, E. R. A. Thomas, A. Tucci, K. Witkowska, S. M. Wood (Genomics England, London, UK); and William Harvey Research Institute, Queen Mary University of London, UK).

17. Snodgrass RO, Chico TJA, Arthur HM. Hereditary haemorrhagic telangiectasia, an inherited vascular disorder in need of improved evidence-based pharmaceutical interventions. *Genes (Basel)*. 2021;12(2):174.
18. Finnamore H, Le Couteur J, Hickson M, Busbridge M, Whelan K, Shovlin CL. Hemorrhage-adjusted iron requirements, hematinics and hepcidin define hereditary hemorrhagic telangiectasia as a model of hemorrhagic iron deficiency. *PLoS One*. 2013;8(10):e76516.
19. Al-Samkari H, Kasthuri RS, Parambil JG, et al. An international, multicenter study of intravenous bevacizumab for bleeding in hereditary hemorrhagic telangiectasia: the InHIBIT-Bleed study. *Haematologica*. 2021;106(8):2161-2169.
20. Al-Samkari H. Hereditary hemorrhagic telangiectasia: systemic therapies, guidelines, and an evolving standard of care. *Blood*. 2021;137(7):888-895.
21. Buscarini E, Botella LM, Geisthoff U, et al; VASCERN-HHT. Safety of thalidomide and bevacizumab in patients with hereditary hemorrhagic telangiectasia. *Orphanet J Rare Dis*. 2019;14(1):28.
22. Kjeldsen A, Aagaard KS, Tørring PM, Möller S, Green A. 20-year follow-up study of Danish HHT patients: survival and causes of death. *Orphanet J Rare Dis*. 2016;11(1):157.
23. de Gussem EM, Edwards CP, Hosman AE, et al. Life expectancy of parents with hereditary haemorrhagic telangiectasia. *Orphanet J Rare Dis*. 2016;11(1):46.
24. Shovlin CL, Hughes JM, Tuddenham EG, et al. A gene for hereditary haemorrhagic telangiectasia maps to chromosome 9q3. *Nat Genet*. 1994;6(2):205-209.
25. Dupuis O, Delagrangue L, Dupuis-Girod S. Hereditary haemorrhagic telangiectasia and pregnancy: a review of the literature. *Orphanet J Rare Dis*. 2020;15(1):5.
26. Eker OF, Boccardi E, Sure U, et al. European Reference Network for Rare Vascular Diseases (VASCERN) position statement on cerebral screening in adults and children with hereditary haemorrhagic telangiectasia (HHT). *Orphanet J Rare Dis*. 2020;15(1):165.
27. Thielemans L, Layton DM, Shovlin CL. Low serum haptoglobin and blood films suggest intravascular hemolysis contributes to severe anemia in hereditary hemorrhagic telangiectasia. *Haematologica*. 2019;104(4):e127-e130.
28. Devlin HL, Hosman AE, Shovlin CL. Antiplatelet and anticoagulant agents in hereditary hemorrhagic telangiectasia. *N Engl J Med*. 2013;368(9):876-878.
29. Shovlin CL, Sulaiman NL, Govani FS, Jackson JE, Begbie ME. Elevated factor VIII in hereditary haemorrhagic telangiectasia (HHT): association with venous thromboembolism. *Thromb Haemost*. 2007;98(5):1031-1039.
30. Livesey JA, Manning RA, Meek JH, et al. Low serum iron levels are associated with elevated plasma levels of coagulation factor VIII and pulmonary emboli/deep venous thromboses in replicate cohorts of patients with hereditary haemorrhagic telangiectasia. *Thorax*. 2012;67(4):328-333.
31. Shovlin C, Bamford K, Sabbà C, et al; VASCERN HHT. Prevention of serious infections in hereditary hemorrhagic telangiectasia: roles for prophylactic antibiotics, the pulmonary capillaries-but not vaccination. *Haematologica*. 2019;104(2):e85-e86.
32. Boothe EJ, Brownlow S, Tighe HC, Bamford KB, Jackson JE, Shovlin CL. Cerebral abscess associated with odontogenic bacteremias, hypoxemia, and iron loading in immunocompetent patients with right-to-left shunting through pulmonary arteriovenous malformations. *Clin Infect Dis*. 2017;65(4):595-603.
33. Mathis S, Dupuis-Girod S, Plauchu H, et al. Cerebral abscesses in hereditary haemorrhagic telangiectasia: a clinical and microbiological evaluation. *Clin Neurol Neurosurg*. 2012;114(3):235-240.
34. Dupuis-Girod S, Giraud S, Decullier E, et al. Hemorrhagic hereditary telangiectasia (Rendu-Osler disease) and infectious diseases: an underestimated association. *Clin Infect Dis*. 2007;44(6):841-845.
35. Topiwala KK, Patel SD, Saver JL, Streib CD, Shovlin CL. Ischemic stroke and pulmonary arteriovenous malformations: a review. *Neurology*. 2022;98(5):188-198.
36. Silva BM, Hosman AE, Devlin HL, Shovlin CL. Lifestyle and dietary influences on nosebleed severity in hereditary hemorrhagic telangiectasia. *Laryngoscope*. 2013;123(5):1092-1099.
37. Finnamore H, Silva BM, Hickson BM, Whelan K, Shovlin CL. 7-day weighed food diaries suggest patients with hereditary hemorrhagic telangiectasia may spontaneously modify their diet to avoid nosebleed precipitants. *Orphanet J Rare Dis*. 2017;12(1):60.
38. Chamali B, Finnamore H, Manning R, et al. Dietary supplement use and nosebleeds in hereditary haemorrhagic telangiectasia: an observational study. *Intractable Rare Dis Res*. 2016;5(2):109-113.
39. Cavalcoli F, Gandini A, Matelloni IA, et al. Dietary iron intake and anemia: food frequency questionnaire in patients with hereditary hemorrhagic telangiectasia. *Orphanet J Rare Dis*. 2020;15(1):295.
40. Shovlin CL, Gilson C, Busbridge M, et al. Can iron treatments aggravate epistaxis in some patients with hereditary hemorrhagic telangiectasia? *Laryngoscope*. 2016;126(11):2468-2474.
41. 100,000 Genomes Project Pilot Investigators, Smedley D, Smith KR, et al. 100,000 genomes pilot on rare-disease diagnosis in health care: preliminary report. *N Engl J Med*. 2021;385(20):1868-1880.
42. Genomics England Confluence. Research environment user guide: Genomics England Research Environment. <https://research-help.genomicsengland.co.uk/display/GERE/Research+Environment+User+Guide>. Accessed 10 August 2021.
43. Clarke JM, Alikian M, Xiao S, et al; Genomics England Research Consortium. Low grade mosaicism in hereditary haemorrhagic telangiectasia identified by bidirectional whole genome sequencing reads through the 100,000 Genomes Project clinical diagnostic pipeline. *J Med Genet*. 2020;57(12):859-862.

44. Balachandar S, Graves TJ, Shimonty A, et al; Genomics England Research Consortium. Identification and validation of a novel pathogenic variant in GDF2 (BMP9) responsible for hereditary hemorrhagic telangiectasia and pulmonary arteriovenous malformations. *Am J Med Genet A*. 2022; 188(3):959-964.
45. Shovlin CL, Guttmacher AE, Buscarini E, et al. Diagnostic criteria for hereditary hemorrhagic telangiectasia (Rendu-Osler-Weber syndrome). *Am J Med Genet*. 2000;91(1):66-67.
46. Shovlin CL, Chamali B, Santhirapala V, et al. Ischaemic strokes in patients with pulmonary arteriovenous malformations and hereditary hemorrhagic telangiectasia: associations with iron deficiency and platelets. *PLoS One*. 2014;9(2):e88812.
47. Gawecki F, Strangeways T, Amin A, et al. Exercise capacity reflects airflow limitation rather than hypoxaemia in patients with pulmonary arteriovenous malformations. *QJM*. 2019;112(5):335-342.
48. Köhler S, Doelken SC, Mungall CJ, et al. The Human Phenotype Ontology project: linking molecular biology and disease through phenotype data. *Nucleic Acids Res*. 2014;42(database issue):D966-D974.
49. National Center for Biotechnology Information. Genome reference consortium human build 38/hg38. [https://www.ncbi.nlm.nih.gov/assembly/GCF\\_000001405.26/](https://www.ncbi.nlm.nih.gov/assembly/GCF_000001405.26/). Accessed 23 December 2021.
50. Martin AR, Williams E, Foulger RE, et al. PanelApp crowdsources expert knowledge to establish consensus diagnostic gene panels. *Nat Genet*. 2019;51(11):1560-1565.
51. Shovlin CL, Millar CM, Droegge F, et al; VASCERN-HHT. Safety of direct oral anticoagulants in patients with hereditary hemorrhagic telangiectasia. *Orphanet J Rare Dis*. 2019;14(1):210.
52. Shovlin CL, Sodhi V, McCarthy A, Lasjaunias P, Jackson JE, Sheppard MN. Estimates of maternal risks of pregnancy for women with hereditary haemorrhagic telangiectasia (Osler-Weber-Rendu syndrome): suggested approach for obstetric services. *BJOG*. 2008;115(9):1108-1115.
53. Shovlin CL, Condliffe R, Donaldson JW, Kiely DG, Wort SJ; British Thoracic Society. British Thoracic Society Clinical Statement on pulmonary arteriovenous malformations. *Thorax*. 2017;72(12):1154-1163.
54. Anderson E, Sharma L, Alsafi A, Shovlin CL. Pulmonary arteriovenous malformations can be the only feature of genetically-confirmed hereditary haemorrhagic telangiectasia which commonly displays few clinical manifestations [published online ahead of print 14 February 2022]. *Thorax*. doi: 10.1136/thoraxjnl-2021-218332.
55. Kim Y, Park J, Kim M. Diagnostic approaches for inherited hemolytic anemia in the genetic era. *Blood Res*. 2017;52(2):84-94.
56. Downes K, Megy K, Duarte D, et al; NIHR BioResource. Diagnostic high-throughput sequencing of 2396 patients with bleeding, thrombotic, and platelet disorders. *Blood*. 2019;134(23):2082-2091.
57. Ferraro NM, Strober BJ, Einson J, et al; GTE Consortium. Transcriptomic signatures across human tissues identify functional rare genetic variation. *Science*. 2020;369(6509):eaaz5900.
58. Bernabeu-Herrero ME, Patel D, Bielowska A, et al. Heterozygous transcriptional and nonsense decay signatures in blood outgrowth endothelial cells from patients with hereditary haemorrhagic telangiectasia [published online ahead of print 6 December 2021]. *Biorxiv* 471269. doi: 10.1101/2021.12.05.471269
59. Richards S, Aziz N, Bale S, et al; ACMG Laboratory Quality Assurance Committee. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med*. 2015;17(5):405-424.
60. Davieson CD, Joyce KE, Sharma L, Shovlin CL. DNA variant classification-reconsidering "allele rarity" and "phenotype" criteria in ACMG/AMP guidelines. *Eur J Med Genet*. 2021;64(10):104312.
61. Krzywinski M, Schein J, Birol I, et al. Circos: an information aesthetic for comparative genomics. *Genome Res*. 2009;19(9):1639-1645.
62. Jalili V, Afgan E, Gu Q, et al. The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2020 update. *Nucleic Acids Res*. 2020;48(W1):W395-W402.
63. Itan Y, Shang L, Boisson B, et al. The human gene damage index as a gene-level approach to prioritizing exome variants. *Proc Natl Acad Sci USA*. 2015;112(44):13615-13620.
64. Petrovski S, Wang Q, Heinzen EL, Allen AS, Goldstein DB. Genic intolerance to functional variation and the interpretation of personal genomes. *PLoS Genet*. 2013;9(8):e1003709.
65. Kircher M, Witten DM, Jain P, O'Roak BJ, Cooper GM, Shendure J. A general framework for estimating the relative pathogenicity of human genetic variants. *Nat Genet*. 2014;46(3):310-315.
66. Rentzsch P, Witten D, Cooper GM, Shendure J, Kircher M. CADD: predicting the deleteriousness of variants throughout the human genome. *Nucleic Acids Res*. 2019;47(D1):D886-D894.
67. Rentzsch P, Schubach M, Shendure J, Kircher M. CADD-Splice-improving genome-wide variant effect prediction using deep learning-derived splice scores. *Genome Med*. 2021;13(1):31.
68. Itan Y, Shang L, Boisson B, et al. The mutation significance cutoff: gene-level thresholds for variant predictions. *Nat Methods*. 2016;13(2):109-110.
69. Karczewski KJ, Francioli LC, Tiao G, et al; Genome Aggregation Database Consortium. The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature*. 2020;581(7809):434-443.
70. Landrum MJ, Chitipiralla S, Brown GR, et al. ClinVar: improvements to accessing data. *Nucleic Acids Res*. 2020;48(D1):D835-D844.
71. Kent WJ, Sugnet CW, Furey TS, et al. The human genome browser at UCSC. *Genome Res*. 2002;12(6):996-1006.



72. Lee BT, Barber GP, Benet-Pagès A, et al. The UCSC Genome Browser database: 2022 update. *Nucleic Acids Res.* 2022;50(D1):D1115-D1122.
73. Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: a practical and powerful approach to multiple testing. *J R Stat Soc B.* 1995;57(1):289-300.
74. Shovlin CL, Jackson JE, Bamford KB, et al. Primary determinants of ischaemic stroke/brain abscess risks are independent of severity of pulmonary arteriovenous malformations in hereditary haemorrhagic telangiectasia. *Thorax.* 2008;63(3):259-266.
75. Precision medicine needs an equity agenda. *Nat Med.* 2021;27(5):737.
76. Rehm HL, Berg JS, Brooks LD, et al; ClinGen. ClinGen: the clinical genome resource. *N Engl J Med.* 2015;372(23):2235-2242.
77. Epi25 Collaborative. Sub-genic intolerance, ClinVar, and the epilepsies: a whole-exome sequencing study of 29,165 individuals. *Am J Hum Genet.* 2021;108(6):965-982.
78. Keramati AR, Chen MH, Rodriguez BAT, et al; NHLBI Trans-Omics for Precision (TOPMed) Consortium. Genome sequencing unveils a regulatory landscape of platelet reactivity. *Nat Commun.* 2021;12(1):3626.
79. Hu Y, Stilp AM, McHugh CP, et al; NHLBI Trans-Omics for Precision Medicine (TOPMed) Consortium. Whole-genome sequencing association analysis of quantitative red blood cell phenotypes: the NHLBI TOPMed program. *Am J Hum Genet.* 2021;108(5):874-893.
80. Zaninoni A, Fermo E, Vercellati C, Marcello AP, Barcellini W, Bianchi P. Congenital hemolytic anemias: is there a role for the immune system? *Front Immunol.* 2020;11:1309.
81. Limeres Posse J, Álvarez Fernández M, Fernández Feijoo J, et al. Intravenous amoxicillin/clavulanate for the prevention of bacteraemia following dental procedures: a randomized clinical trial. *J Antimicrob Chemother.* 2016;71(7):2022-2030.
82. Kanoi BN, Nagaoka H, White MT, et al. Global repertoire of human antibodies against *Plasmodium falciparum* RIFINs, SURFINs, and STEVORs in a malaria exposed population. *Front Immunol.* 2020;11:893.