The ADAMTS13-VWF axis is dysregulated in chronic thromboembolic pulmonary hypertension.

Newnham, Michael; South, Kieron; Bleda, Marta; Auger, William; Barberà, Joan ; Bogaard, Harm; Bunclark, Katherine ; Cannon, John; Delcroix, Marion ; Hadinnapola, Charaka; Howard, Luke; Jenkins, David; Mayer, Eckhard; Ng, Choo; Rhodes, Christopher; Screaton, Nicholas; Sheares, Karen; Simpson, Michael; Southwood, Mark; Su, Li

DOI:
10.1183/13993003.01805-2018

Document Version
Peer reviewed version

Citation for published version (Harvard):
Newnham, M, South, K, Bleda, M, Auger, W, Barberà, J, Bogaard, H, Bunclark, K, Cannon, J, Delcroix, M, Hadinnapola, C, Howard, L, Jenkins, D, Mayer, E, Ng, C, Rhodes, C, Screaton, N, Sheares, K, Simpson, M, Southwood, M, Su, L, Taboada, D, Traylor, M, Trembath, R, Villar, S, Wilkins, M, Wharton, J, Gräf, S, Pepe-Zaba, J, Laffan, M, Lane, D, Morrell, N & Toshner, M 2019, 'The ADAMTS13-VWF axis is dysregulated in chronic thromboembolic pulmonary hypertension.', *European Respiratory Journal*, vol. 53, no. 3, 1801805. https://doi.org/10.1183/13993003.01805-2018

SUPPLEMENTAL MATERIAL

Supplementary Methods

Sample size calculations

The effect size for a difference in ADAMTS13 between other thrombotic diseases and healthy controls was calculated from a study of myocardial infarction and stroke (Andersson). An estimated 142 individuals would be required for a small-medium effect size at a power of 80% (supplementary figure S7a). However, to establish what effect size is appropriate for CTEPH a pilot study was performed in 2016 to refine the sample size estimates. ADAMTS13 plasma antigen levels were measured in 94 CTEPH patients and 29 healthy controls. This confirmed a larger effect size between CTEPH and healthy controls than had previously been reported for other thrombotic diseases. Consequently, an estimated 30 individuals (per group) would be required for a large effect size in ADAMTS13 with 80% power (supplementary figure S7b). Furthermore, as CTEPH, CTED and IPAH are rare diseases with limited sample numbers available, the study was powered for larger effect sizes.

Study samples and participants

The 208 CTEPH patients represented 40% (208/514) of all the patients diagnosed with CTEPH during the same period at Royal Papworth Hospital, UK. The healthy controls (n=68) were volunteers bio-banked from Royal Papworth Hospital and Hammersmith Hospital, UK. Propensity matching for age and sex was explored for CTEPH, healthy

controls and disease comparators (CTED, IPAH and PE). However, given the known demographic differences across these disease groups and limited sample availability for rare diseases, it was not possible to fully propensity match. Therefore, healthy controls and disease comparators were selected for the closest possible age- and sex- matching to the CTEPH group. The closest matched (age and sex) CTED patients (n=35) from Royal Papworth Hospital with available bio-banked samples were selected. The definition used for CTED patients has been previously described [1]. IPAH patients (n=30) from Royal Papworth Hospital were also selected for the closest age/sex matching and additionally all IPAH patients were matched for anticoagulation therapy usage but had not had a venous thromboembolism. The PE group were sampled from a specialist PE follow-up service (Hammersmith Hospital) at a median of 220 (interquartile range (IQR) 218) days following an acute PE. All consecutive patients (n=28) with bio-banked samples available (November 2013 to October 2014) were included in the study.

As the CTEPH, healthy control and disease comparator groups could not be fully matched for age and sex, these variables were included in multivariable linear models described in the statistical analysis section.

ADAMTS13 and VWF plasma concentrations

Plasma ethylenediaminetetraacetic acid (EDTA) samples were used to measure ADAMTS13 and VWF antigen (Ag) levels by enzyme-linked immunosorbent assays

(ELISA). Normal human control plasma (Technoclone, Vienna, Austria) with known concentrations of ADAMTS13 and VWF was used as the reference [2].

ADAMTS13

ADAMTS13 plasma antigen levels were quantified using a polyclonal rabbit anti-ADAMTS13 antibody (5µg/mL, anti-TSP2–4 depleted) as previously described [2, 3]. The antibody was immobilised in 96-well microplates (Nunc, Rochester, USA) in 50mM carbonate buffer, pH 9.6 at 4°C overnight. Washes were performed with phosphate buffered saline (PBS) + 0.1% Tween-20 (PBST), and this was repeated between each step. Wells were blocked with 1% bovine serum albumin (BSA) in PBS for 1 hour. Plasma samples were diluted 1:20 using 1% BSA in PBS and added to the wells in duplicate for 2 hours. A standard curve of 0-126 ng/mL was made with normal human control plasma (NHP) (Technoclone, Vienna, Austria). Bound ADAMTS13 was detected with biotinylated anti-TSP2–4 polyclonal antibody for 2 hours followed by incubation of wells with streptavidin-horseradish peroxidase (HRP) (GE Healthcare, UK) for 1 hour. Plates were developed with a peroxidase substrate (o-phenylenediamine dihydrochloride (OPD); Sigma-Aldrich, Darmstadt, Germany) for 5 minutes and the reaction was stopped with 65µL/well of 2.5M H_2SO. Absorbance was read at 492nm (FLUOstar Omega plate reader, BMG Labtech). ADAMTS13 concentrations were obtained by interpolating from the standards fitted with a four-parameter logistic curve. The intra- and inter-assay coefficients of variation were 8 and 12% respectively.

To enable a comparison with other published thrombotic diseases, each ADAMTS13 plasma antigen level was divided by the median of the healthy control group and expressed as a percentage. The CTEPH group was then divided into quartiles of the ADAMTS13 distribution of the healthy control group. The odds ratios for a combination of ADAMTS13 and VWF antigen levels. The quartile thresholds were used to stratify CTEPH patients and healthy controls into groups using a combination of ADAMTS13 and VWF levels. Odds ratios for the different groups were then assessed using logistic regression adjusted for age, sex, ethnicity and experimental batch.

VWF

VWF plasma antigen levels were quantified in a similar well-described manner using a polyclonal rabbit anti-VWF antibody (3.1 µg/mL; Dako, Santa Clara, USA) [3]. After overnight antibody immobilisation and washes, wells were blocked with 1% BSA in PBST for 1 hour. Plasma samples were diluted 1:400 in PBST 1% BSA and a standard curve of 0-125 ng/L was made with NHP. VWF was detected with 1.1µg/mL Polyclonal Rabbit Anti-Human VWF/HRP (Dako) followed by plate development with OPD for 3 minutes. The intra- and inter-assay coefficients of variation were 5 and 8% respectively.

The ADAMTS13 and VWF ELISAs were performed for all groups in 2016 (batch1) and additional CTEPH samples (n=115) were included in 2017 (batch2). Batch variation in ADAMTS13 and VWF was adjusted using replicate samples (ADAMTS13: n=24, VWF:

n=12) and additionally batch was included as a covariate in multivariable linear regression models.

ADAMTS13 activity, D-dimer, anti-ADAMTS13 autoantibodies and VWF multimeric size

Additional experiments were performed on a subset of the CTEPH (n=23) and healthy control (n=14) groups to identify potential mechanisms for any dysregulation of the ADAMTS13-VWF axis. Plasma lithium heparin samples were used to measure ADAMTS13 activity, D-dimer concentrations and anti-ADAMTS13 autoantibodies. The CTEPH sample subset were those with the lowest ADAMTS13 antigen levels (below the first quartile of the CTEPH group) and the healthy controls were those with ADAMTS13 antigen levels closest to the median of the control group. An estimate of VWF multimeric size was made by measuring VWF collagen binding (VWF:CBA) and comparing this with VWF antigen levels in the CTEPH (n=21) samples with the highest VWF antigen concentrations (above the third quartile of the CTEPH group) and the same healthy control subset.

ADAMTS13 Activity

ADAMTS13 activity was measured with a fluorescence resonance energy transfer (FRETS) assay using a short synthetic VWF peptide (VWF73: PeptaNova, Sandhausen, Germany) containing the ADAMTS13 cleavage site for VWF [4]. Plasma samples and NHP were diluted to 1:10 in reaction buffer (5 mM Bis-Tris, 25 mM $CaCl_2$ and 0.005% Tween-20 at pH 6.0) in 96-well plates (Nunc, Rochester, USA). FRETS-

VWF73 substrate (an equal volume of 4μM) was added and fluorescence was recorded at 1-minute intervals for 1 hour (FLUOstar Omega plate reader) to monitor substrate proteolysis. Assays were repeated 3 times to obtain the mean fluorescence and ADAMTS13 activity was normalised to NHP, which was defined as 100%.

D-Dimer plasma levels

Plasma D-Dimer levels were quantified using an ELISA kit (ab196269, abcam, Cambridge, USA) according to the manufacturer's instructions. Plasma lithium heparin samples from CTEPH patients and healthy controls were used at a dilution of 1:1000.

Anti-ADAMTS13 autoantibodies

ADAMTS13 was immobilised directly on Maxisorp plates (Nunc) in carbonate buffer overnight at 4°C. After blocking with 1% BSA, patient plasma samples (1:10 dilution in 1% BSA) were applied to the plates for 1 hour. Bound antibodies were detected and quantified using a rabbit anti-human IgG polyclonal antibody conjugated with HRP. A standard curve of normal human plasma, containing known concentrations of the purified, patient derived anti-ADAMTS13 monoclonal antibody (II-1), was used as a positive control for the binding and detection of autoantibody [5]. Anti-ADAMTS13 autoantibodies were normalised to NHP, which was defined as 100%.

VWF multimeric size

VWF multimeric size was evaluated with a collagen binding assay (CBA) which utilises the increased collagen binding of higher multimeric VWF. Human type III placental

collagen (5µg/mL) was immobilised in 96-well microplates plates (Nunc) in 50mM carbonate buffer, pH 9.6 at 4°C overnight. After washes with PBST, wells were blocked with 2% BSA in PBST for 1 hour. Plasma lithium heparin samples were diluted 1:100 in PBST 1% BSA and a standard curve of 0-1000ng/mL was made with NHP. VWF was detected with 1.1µg/mL polyclonal rabbit anti-human VWF/HRP (Dako) followed by plate development with OPD for 3 minutes. Collagen binding is reported as a ratio over the total plasma VWF antigen.

Clinical phenotype data

Phenotype data for the CTEPH, CTED and IPAH groups were recorded closest to the time of diagnosis and pre-operatively for the CTEPH and CTED patients undergoing PEA. This included demographics, haemodynamics, WHO functional class, 6-minute walk distance (6mwd), clinical blood tests, smoking history and anticoagulation therapy usage. Additionally, post-operative haemodynamics were recorded within 1 year of surgery for the CTEPH and CTED patients that underwent PEA, as part of routine care. Haemodynamics were evaluated by right heart catheterisation according to international guidelines and PEA was performed as previously described [6, 7]. The PE group had phenotype data recorded at a follow-up visit (median 220 days) after their acute PE, which also included a ventilation perfusion (VQ) scan to assess residual perfusion defects.

The relationship between the ADAMTS13-VWF axis and the amount of chronic thromboembolism was not evaluated as a validated radiological classification system does not exist.

Genotype data and genome-wide association study

To date, 1457 Caucasian CTEPH patients have been recruited from 8 European and US specialist pulmonary hypertension centres. The participating centres were:

- Royal Papworth Hospital, Cambridge, UK
- Hammersmith Hospital, London, UK
- University of California, San Diego, USA
- Hospital Clínic - IDIBAPS-CIBER Enfermedades Respiratorias, University of Barcelona, Spain
- KU Leuven – University of Leuven
- Kerckhoff Heart and Lung Centre, Bad Nauheim, Germany
- VU University Medical Centre, Amsterdam, Netherlands
- Medical University, Vienna, Austria

They have been compared to 1536 healthy Caucasian controls from the Wellcome Trust Case Control Consortium (WTCCC) [8].

Genomic DNA was extracted and quantified from whole blood or buffy coat fractions (LGC, Hoddesdon, Herts, UK). Genotyping was performed using the Illumina HumanOmniExpressExome-8 v1.2 BeadChip containing 964,193 single-nucleotide

polymorphism (SNP) markers (Kings College, London, UK). The Genome Reference Consortium human genome (build 37) (GRCh37) was used for genomic positions. Three batches have been genotyped from 2014-16 (batch1: 2014, batch2: 2015, batch3: 2016), with all WTCCC controls genotyped in batch1.

Each batch of micro-array intensity data was normalised, clustered and called independently using Illumina GenomeStudio [9]. Samples containing more than 1% of missing genotypes were removed and SNPs were then re-clustered to remove confounding from differential SNP exclusions between batches. SNPs with a low GenTrain score (<0.7) or clustering separation score (<0.5) were excluded [9, 10].

Sample/SNP QC

Divergent ancestry was assessed using principal component analysis (PCA). Each batch was merged with the 1000 Genomes data using an intersecting set of SNPs [11]. A robust set of independent SNPs (n=30,609) was used for PCA, selected by the following criteria: genotype missingness < 5%, SNPs in Hardy-Weinberg equilibrium (HWE) ($p>1x10^{-5}$), minor allele frequency (MAF) > 5%, non-independent SNPs (pairwise $R^2 < 0.2$). Furthermore, SNPs in several regions with long-range linkage disequilibrium (LD) were excluded [12]. PCA against all populations in the 1000 genomes data was conducted and samples not clustering with European population were excluded. A second PCA was then performed with the remaining samples against European samples only, and outlying samples were also excluded. Thresholds were set by visual inspection of plots.

Additionally, individual samples were removed due to: outlying heterozygosity (3 standard deviations from the mean), duplicates or relatedness (Identity by descent > 0.1875) and missing genotypes (>1%; as described). SNP markers were removed due to: missing genotypes (>1%), deviation from HWE ($p<1x10^{-6}$), differential missingness rate between cases and controls ($p<1x10^{-5}$). and multi-allelic SNPs.

The 3 batches were then merged using an intersecting set of SNPs and further PCA was performed using the robust SNP set previously used for divergent ancestry quality control, to check for batch and recruiting centre effects.

Phasing/imputation

After quality control exclusions there were 1250 CTEPH cases, 1492 healthy controls and 915,999 SNPs. Phasing and imputation were performed using EAGLE 2 (v2.0.5) and positional Burrows–Wheeler transform (PBWT) software via the Sanger imputation service (https://imputation.sanger.ac.uk/) [13, 14]. The reference panel was the Haplotype Reference Consortium (release 1.1), containing 39 million biallelic SNPs from 32,470 individuals [15]. Following imputation, SNPs were excluded if they had a low minor allele frequency (<1%) or if they were poorly imputed (info score < 0.5), with 7,675,738 SNPs remaining for association testing.

208 CTEPH patients with ADAMTS/VWF antigen levels and 28 patients with CTED were also included in the CTEPH GWAS. Genotypes were available for 207 (187

CTEPH; 23 CTED) after GWAS quality control exclusions. Matched genotypes and ADAMTS / VWF antigen levels were not available for the healthy control, IPAH or PE groups.

Genetic *ABO* groups

The tagging SNPs used to reconstruct the genetic *ABO* groups A1, A2, B and O from phased haplotypes were: rs8176746, rs8176704, rs687289 and rs507666 [16]. The genetic *ABO* groups were compared to the available ABO antigen groups measured by serology (n=1490 Healthy control group) to confirm the accuracy of this method (98% concordance; n=29 were unable to be classified). The 10 genetic *ABO* groups were converted into A, B, AB and O groups using the following criteria: A = A1A1, A1A2, A2A2, A1O, A2O; B = BB, BO; AB = A1B, A2B; O = OO.

Protein quantitative trait loci

The post-imputation allelic dosages (0-2) of individual SNPs in the *ADAMTS13* gene ± 40kb flanking region (n=396 variants) were tested against log transformed ADAMTS13 protein levels (dependent variable) and adjusted for age, sex and ADAMTS13 plasma antigen experimental batch. This region included the ADAMTS13 cis-pQTLs that have previously been described [17-19]. A Bonferroni *p*-value threshold $<1.26\times10^{-4}$ (0.05/396 variants) was used to denote statistical significance. Partitioning of the variance explained by each variable within the models was performed by averaging over orders using the R package 'relaimpo' [20].

Linkage disequilibrium was quantified from 1000 Genomes project data using LDlink (https://analysistools.nci.nih.gov/LDlink/, accessed 22/01/2018) in all European populations [21].

Statistical analysis

The differences in categorical variables between groups were assessed using Chi-squared or Fisher's exact test. The differences in continuous variables were assessed using the Mann-Whitney *U* test and the Kruskal-Wallis test.

Post-hoc pairwise diagnostic group comparisons were performed using Dunn's test with false discovery rate (FDR) adjustment for multiple testing. For matched values pre- and post-PEA Wilcoxon signed-rank test was used. *P*-values are reported to 3 decimal places and experimental data are reported to 3 significant figures. Data averages are described as median ± interquartile range unless specified.

Multivariable linear regression was performed using log-transformed ADAMTS13 or VWF protein levels as the dependent (outcome) variables after assessing log-likelihoods using the Box-Cox power transformation. Log-transformed ADAMTS13 and VWF were used in all multivariable linear regression models (supplementary tables S3, S4, S6, S7 and S8). The models were adjusted for age, sex, experimental batch (supplementary tables S7 and S8) and additionally ethnicity (supplementary tables S3, S4 and S6), VWF (supplementary tables S3 and S8) and 5 ancestry informative principal components (supplementary tables S7 and S8). The β coefficients and

confidence intervals are presented as percentage change ($(\exp^{\beta}-1) \times 100$) to enable clinical interpretation of the log-transformed values. Models were checked for normality of residuals, homoscedasticity and multicollinearity (variance inflation factor), with additional checks performed using the R package 'gvlma' [22]. We investigated interacting effects between the variables that were used in supplementary tables S4 and S5. The significant ($p<0.05$) and informative interactions were included in additional multivariable linear regression models (supplementary table S9).

Spearman's rank correlation coefficients were used to describe associations between ADAMTS13 or VWF protein levels and clinical phenotypes associated with disease severity (pulmonary vascular resistance (PVR), 6mwd and N-terminal pro b-type natriuretic peptide (NT-proBNP)) and blood markers of inflammation (white cell count (WCC), C-reactive protein (CRP), neutrophil and lymphocyte percentages). *P*-values from correlation testing were corrected for multiple testing using false discovery rate (FDR) adjustment.

Software

The analysis was performed using the following software: Illumina Genomestudio [9], PLINK (v1.90beta) [23], bcftools (v1.4.1) [24], LDLink [21], R (3.4.3) [25] and RStudio (1.1.414) [26]. The R packages used included: MASS [27], coin [28], gvlma [22], PMCMR [29], SNPRelate [30], relaimpo [20], jtools [31], forestmodel [32] and the tidyverse suite [33].

Supplementary Results

Data missingness

All individuals with CTEPH, CTED, PE, IPAH and healthy controls included in the analyses had ADAMTS13 and VWF antigen levels available (missingness: n=0/369, 0%). The core variables used in multivariable linear models had the following missingness: there was no missing data for age and sex (n=0/369, 0% missing) and ethnicity was missing in 7% (n=26/369). Data missingness for other variables are displayed in the figure and table legends. Analyses were performed using complete data. As missingness was low in core variables, imputation was not performed and is unlikely to have altered the results.

ADAMTS13 plasma concentrations

We were 95% powered to detect the difference in ADAMTS13 levels between CTEPH and healthy controls (healthy controls(n1)=68, CTEPH(n2)=208, effect size=0.5, two-sided t-test of means) and 99% powered to detect the difference between the CTED group and healthy controls (healthy controls(n1)=68, CTED (n2)=35, effect size=1.12, two-sided t-test of means).

When the analysis is confined to batch 1 samples with a closer 1:1 matching between CTEPH (n=94) and healthy controls (n=68), the markedly reduced ADAMTS13 in the CTEPH group remains in a multivariable linear regression model (β (95% CI) (% change) = -24.8 (-32.5, -16.2)%, $p<0.001$).

We investigated interaction effects for the variables used in supplementary tables S3 and S4. For ADAMTS13 antigen levels, there was a significant interaction between age and CTEPH (p=0.007) and additionally between age and sex (p=0.019) (supplementary table S9 and supplementary figure S8). This suggests that the reduction in ADAMTS13 levels with increasing age is of more relevance within the CTEPH group. Consideration of the interaction terms is most relevant for the extreme values. For example, there is less difference between a 30-year-old Caucasian female with CTEPH and a 30-year-old Caucasian male healthy control (predicted ADAMTS13: 1.07 vs. 1.28 µg/mL, 16% reduction) than an 80-year-old Caucasian male with CTEPH and an 80-year-old Caucasian female healthy control (0.688 vs. 1.18 µg/mL, 42% reduction). There were no significant interaction effects for a separate model of VWF antigen levels using the variables in supplementary table 4.

Converting ADAMTS13 antigen levels to a percentage of the median value of the healthy control group (set at 100%) allowed comparisons with thrombotic diseases in other studies. The majority of the CTEPH group (n=136, 65%) were in the lowest quartile (Q1<88% ADAMTS13) (supplementary table S10).

There was a modest negative correlation between ADAMTS13 and VWF plasma levels in CTEPH (rho = -0.164, p=0.018) but they were not correlated in healthy controls (rho = -0.0622, p=0.614) (supplementary figure S9). Furthermore, adjusting the multivariable linear regression model of ADAMTS13 antigen levels (supplementary table S3) by VWF

had minimal effect, suggesting that the associations are not mediated by VWF antigen levels.

ADAMTS13 and VWF: Pre- and post-pulmonary endarterectomy

ADAMTS13 and VWF levels did not change pre- and post-PEA and this also applied when limited to patients with normal post-operative haemodynamics (mPAP <25mmHg) (n=7, ADAMTS13: p=0.742, VWF: p=0.195).

ADAMTS13 activity, D-dimer, anti-ADAMTS13 autoantibodies and VWF multimeric size measurements

ADAMTS13 activity and D-dimer concentrations were measured in a subset of patients with CTEPH (n=23) with the lowest ADAMTS13 protein concentrations (below the 1st Quartile of the CTEPH group) (0.556 ± 0.130 µg/mL) and compared to a subset of healthy controls (n=14, ADAMTS13: 1.03 ± 0.284 µg/mL).

ADAMTS13 activity

ADAMTS13 activity that is not adjusted for ADAMTS13 antigen levels is reduced (supplementary figure S10a). The increased specific ADAMTS13 activity (Act:Ag ratio) reflects the greater decrease in ADAMTS13:Ag compared to ADAMTS13:Act. Specific ADAMTS13 activity (Act:Ag) is not correlated with VWF:Ag in either CTEPH or healthy controls (supplementary figure S10b).

VWF multimeric size

The CTEPH (n=21) and healthy control subset (n=14) that were used to assess VWF multimeric size had a VWF plasma antigen level of 32.5 ± 6.80 µg/mL and 9.97 ± 4.99 µg/mL respectively. VWF:CBA was not correlated with ADAMTS13:Ag in CTEPH or healthy controls (supplementary figure S10c). VWF:CBA was correlated with VWF:Ag in healthy controls but not in CTEPH (supplementary figure S10d).

Clinical phenotype associations with ADAMTS13 and VWF

There were no correlations with blood markers of inflammation (CRP, WCC, neutrophil and lymphocyte %) (supplementary figure S5). including when confining the analysis to samples that were taken on the same day as ADAMTS13 and VWF sampling (n=81, for WCC, neutrophil and lymphocyte %; n= 77 for CRP).

As proximal operable CTEPH has different risk associations to distal inoperable CTEPH and thus potentially different pathophysiological mechanisms we investigated the disease sub-types [34]. There was no difference in ADAMTS13 (p=0.070) or VWF (p=0.253) between the different sub-diagnostic categories of CTEPH (supplementary figures S11a and S11b). Furthermore, there was no difference in ADAMTS13 (p=0.366) or VWF (p=0.078) in those with and without post-operative residual pulmonary hypertension (mPAP $\geq$ 25mmHg) (n=83, 63%), which is a potential marker of distal vasculopathy (supplementary figures S11c and S11d) [35].

CTEPH is a potential severe consequence of acute PE, however there are a spectrum of changes following PE (post-PE syndrome) that may have differing pathobiology

[36]. We evaluated whether there was a difference in ADAMTS13 or VWF antigen levels depending on the degree of post-PE perfusion defects on available VQ scans (n=20). There was no difference in ADAMTS13 (p=0.812) or VWF (p=0.678) levels in those that had residual perfusion defects post-PE (n=12) compared with those with no perfusion defects (n=8) (supplementary figures S12a and S12b). Furthermore, there was no difference in ADAMTS13 (p=0.938) or VWF (p=0.427) levels when the PE group was stratified into provoked PE (n=8) and idiopathic PE (n=12) (supplementary figures S12c and S12d).

Dysregulation of the ADAMTS13-VWF axis is associated with stroke and myocardial infarction. Where co-morbidities were available for the CTEPH group, 5/74 (6%) had cerebrovascular disease and 20/124 (14%) had coronary artery disease. This is unlikely to have been a major confounder in the CTEPH group.

***ABO* groups and ADAMTS13-VWF**

Reconstructing genetic *ABO* groups allowed us to explore more complex associations within the *ABO* subgroups. Whilst the A1 and A2 groups would be classified as non-O on serological testing, they have been associated with different effects on VWF levels and VTE risk [37].

The small sample size of some non-O A*BO* groups (n: B=12, AB=13) limits the power to detect associations of smaller effect.

Protein quantitative trait loci for ADAMTS13

In a model adjusted for age, sex and batch, the lead SNP (rs3739893) explained 7.7% of the variance in ADAMTS13 levels within the CTEPH group (supplementary table S8). However, as only 10 CTEPH patients had the rs3739893 effect allele, this accounts for a small proportion of the ADAMTS13 antigen level reduction observed in CTEPH.

The 4 other significantly associated SNPs were highly correlated with the lead SNP (R^2=0.91-1.00, p<0.001). Additional analysis correcting for the first 5 ancestry informative principal components and VWF antigen levels did not alter the results.

In the linear regression model of ADAMTS13 antigen levels and rs3739893, the most variance in ADAMTS13 antigen levels was attributable to age (16%) (supplementary table S7), which is higher than reported in healthy cohorts [18]. In the CTEPH GWAS the effect allele frequency for rs3739893 was 0.0128 in CTEPH cases and 0.0158 in healthy controls and this was similar to a European (non-Finnish) reference population in gnomAD (0.0160) (http://gnomad.broadinstitute.org/, accessed Feb 2018). The results were unchanged when the analysis was confined to the CTEPH group.

Supplementary Figures

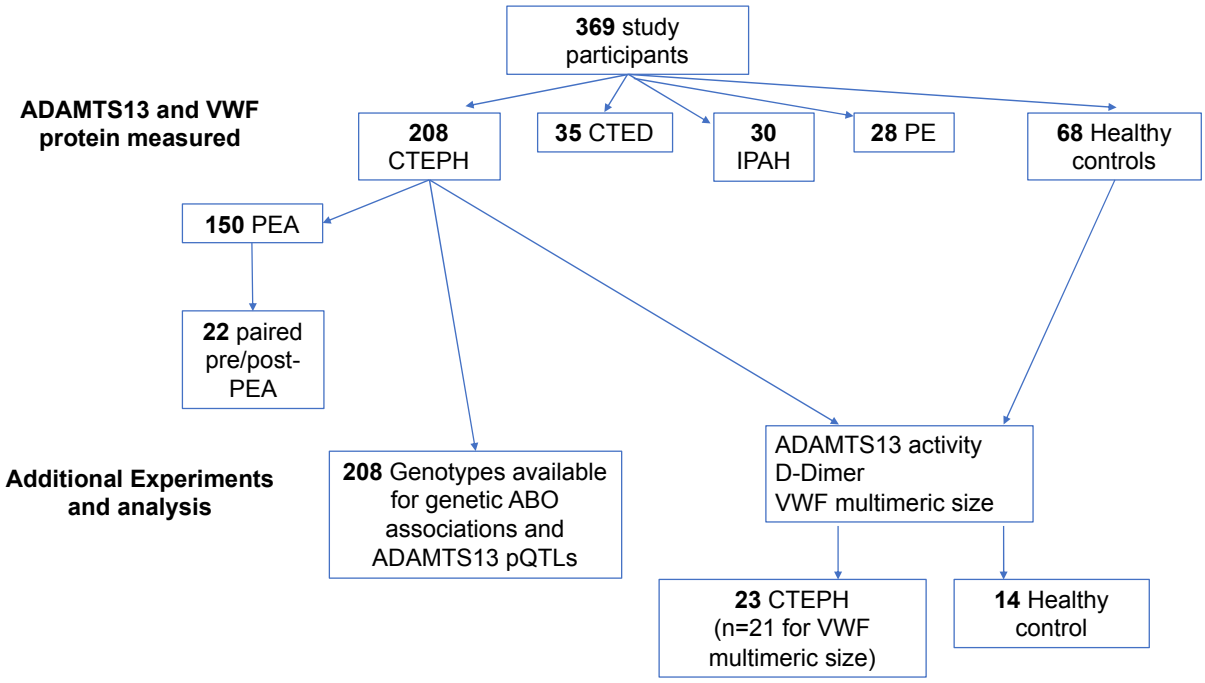

FIGURE S1

Flow chart of study design and study participant numbers. PEA (pulmonary endarterectomy), pQTL (protein quantitative trait loci). The number in each group are shown in bold.

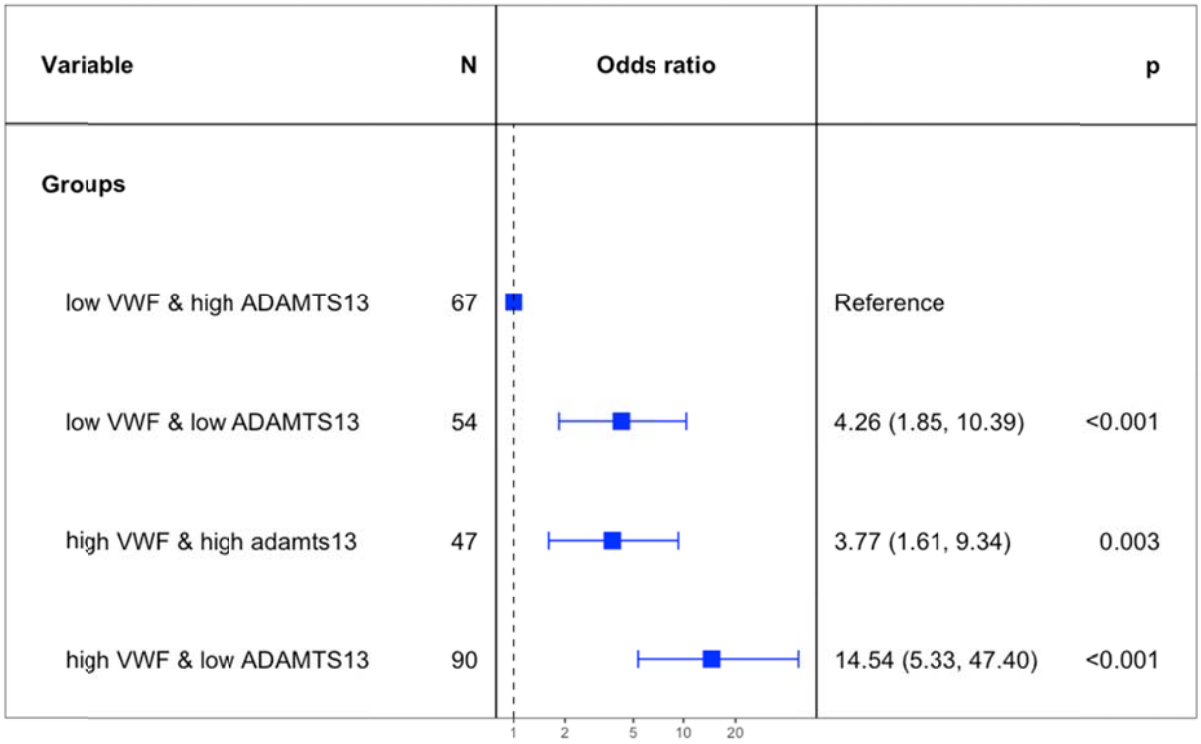

FIGURE S2

The odds ratios of CTEPH in relation to healthy controls for combined ADAMTS13 and VWF groups. The ADAMTS13 and VWF group ORs are adjusted for age, sex, ethnicity and batch in a logistic regression model. N represents the total for both the CTEPH and healthy control groups. Threshold criteria and n (%) within the CTEPH and healthy control groups are shown in supplementary table S5. Forest plot generated with the R package 'forestmodel' [32].

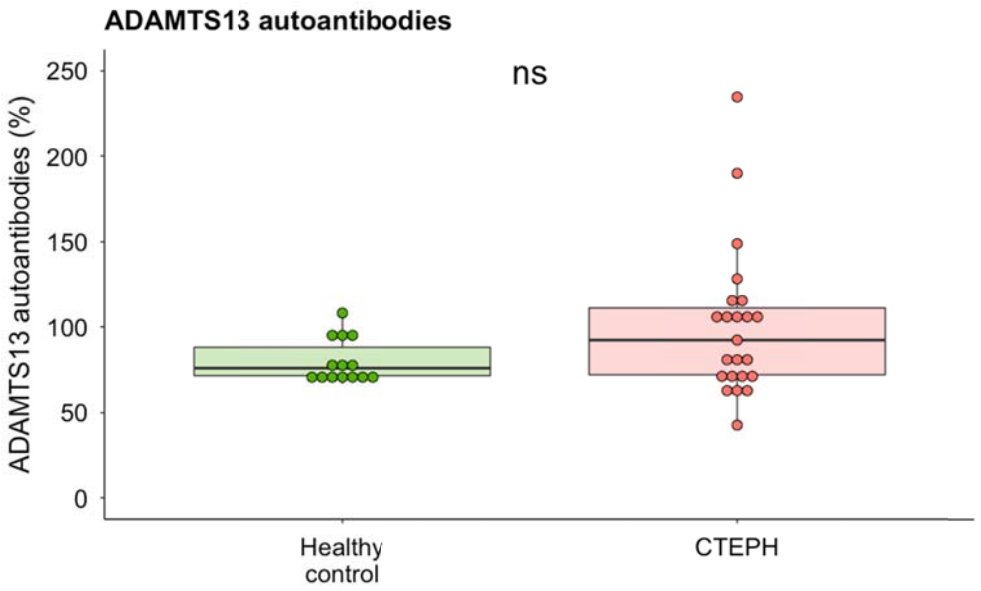

FIGURE S3

Anti-ADAMTS13 autoantibody percentage in CTEPH and healthy controls. Anti-ADAMTS13 autoantibodies were normalised to normal human control plasma, which was defined as 100%

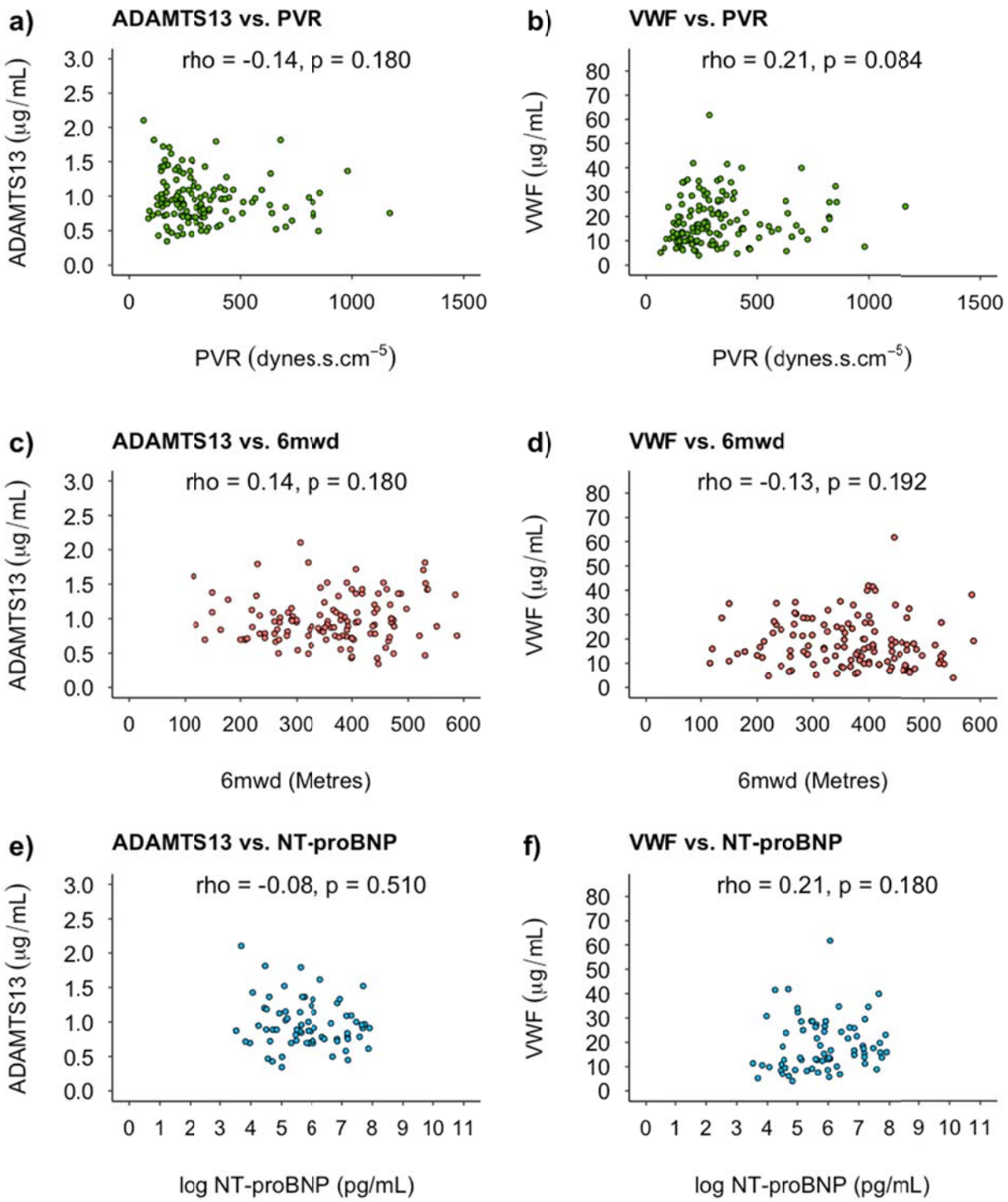

FIGURE S4

Correlation of ADAMTS13 and VWF antigen levels with markers of disease severity in CTEPH. Correlation was assessed by Spearman's rank test. *P*-values adjusted for the

number of statistical tests performed using FDR correction. 6mwd (6-minute walk distance), NT-proBNP (N-terminal pro b-type natriuretic peptide), PVR (pulmonary vascular resistance). NT-proBNP log-transformed to improve visualisation. Numbers in each group: PVR = 169, 6mwd = 165, NT-proBNP = 144.

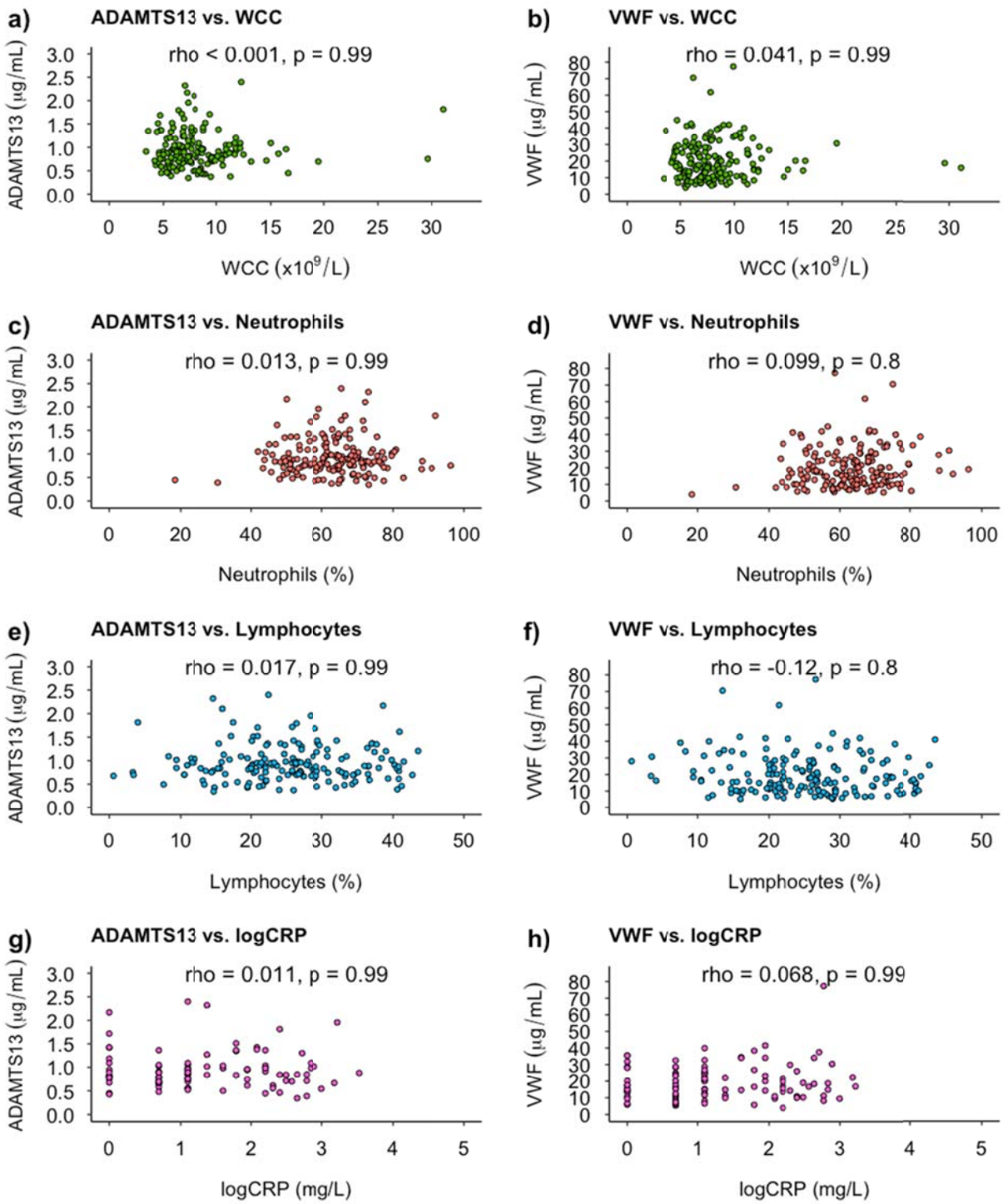

FIGURE S5

Correlation of ADAMTS13 and VWF antigen levels with blood markers of inflammation. Correlation was assessed by Spearman's rank test. *P*-values adjusted for the number

of statistical tests performed using FDR correction. CRP log-transformed to improve visualisation. CRP (C-reactive protein), WCC (white cell count). Numbers in each group: WCC = 169, Neutrophils = 169, Lymphocytes = 168, CRP = 95.

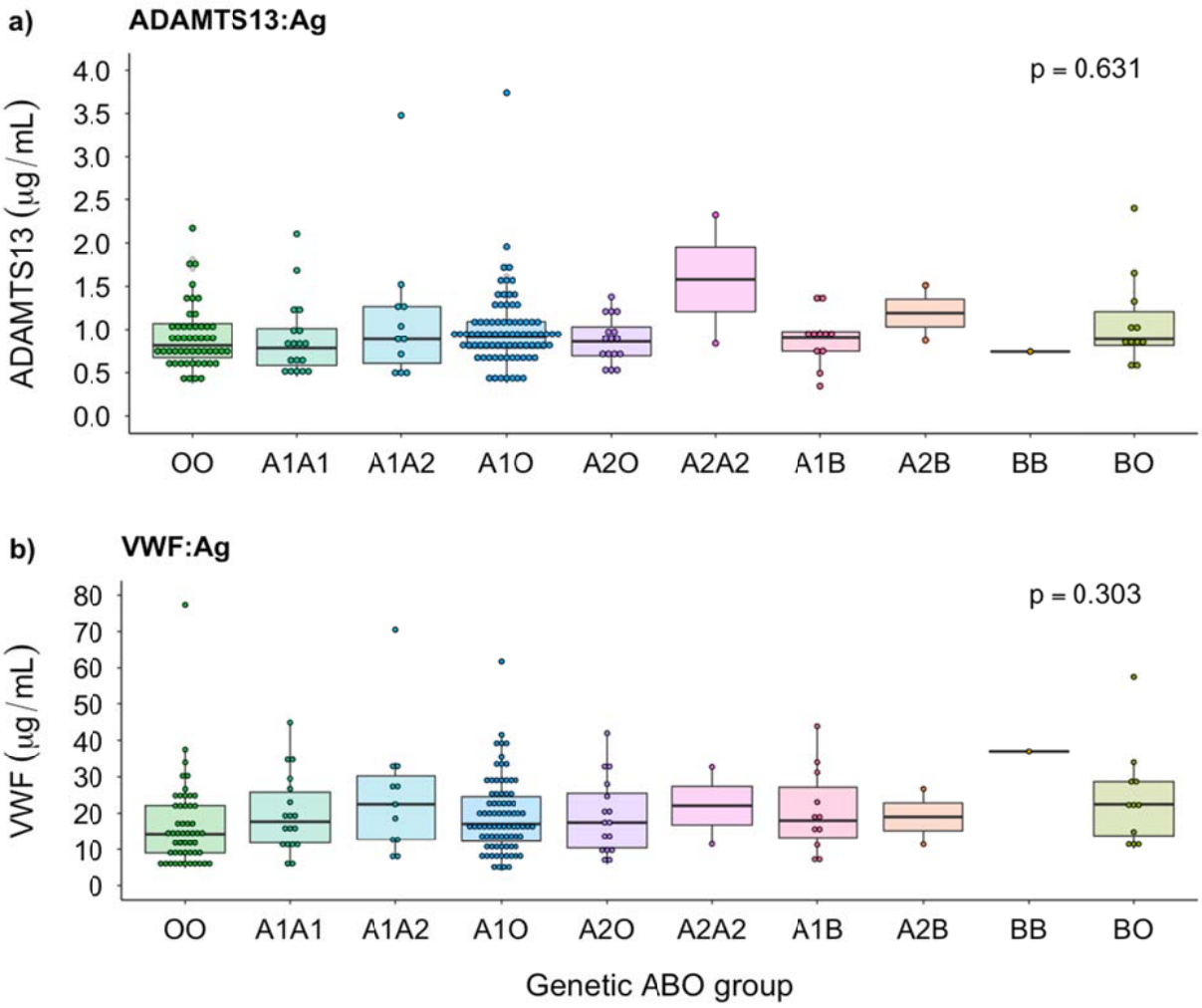

FIGURE S6

ADAMT13 and VWF antigen levels by comprehensive *ABO* genetic groups. The group differences were assessed using the Kruskal-Wallis test. Numbers in each group: OO = 51, A1A1 = 18, A1A2 = 11, A1B = 11, A1O = 81, A2A2 = 2, A2B = 2, A2O = 16, BB = 1, BO = 11.

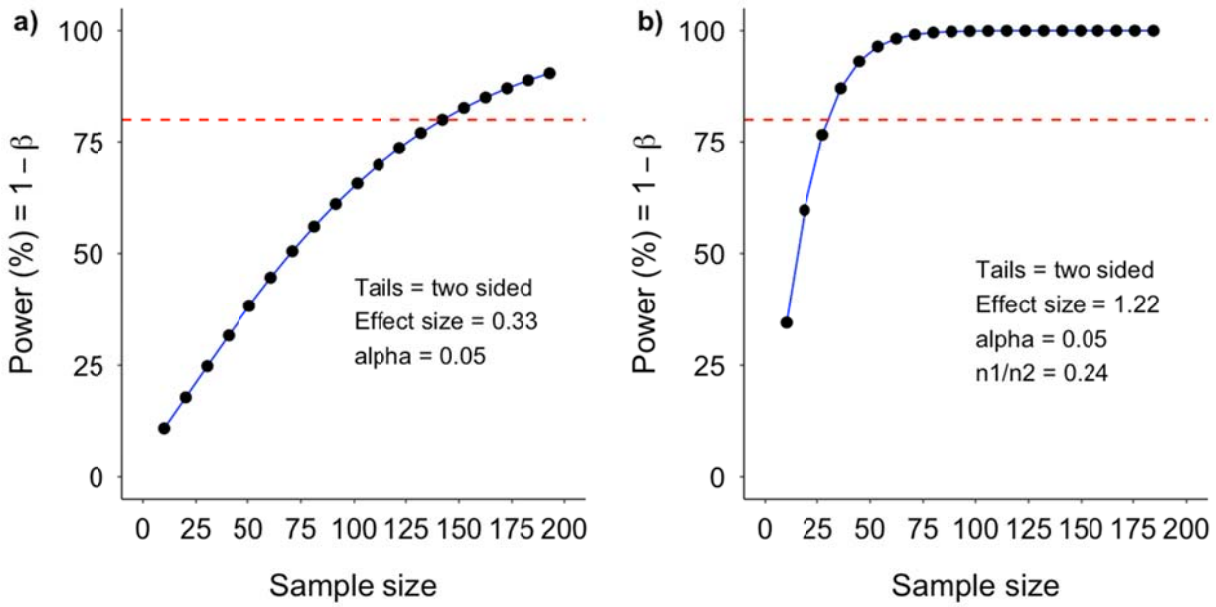

FIGURE S7

Sample size calculations for ADAMTS13 plasma concentration. Sample sizes calculated from two samples using t-test of mean (R package `pwr`).

a) Sample size estimation for a small-medium effect size. An estimated 142 individuals (per group) would be required for a small-medium effect size at a power of 80%.

b) Sample size estimation for large effect size (determined by pilot data) with unequal sample size (n: healthy controls 29, CTEPH 94).

n1/n2 (ratio of healthy controls to CTEPH), alpha (type 1 error probability), β (type 2 error probability).

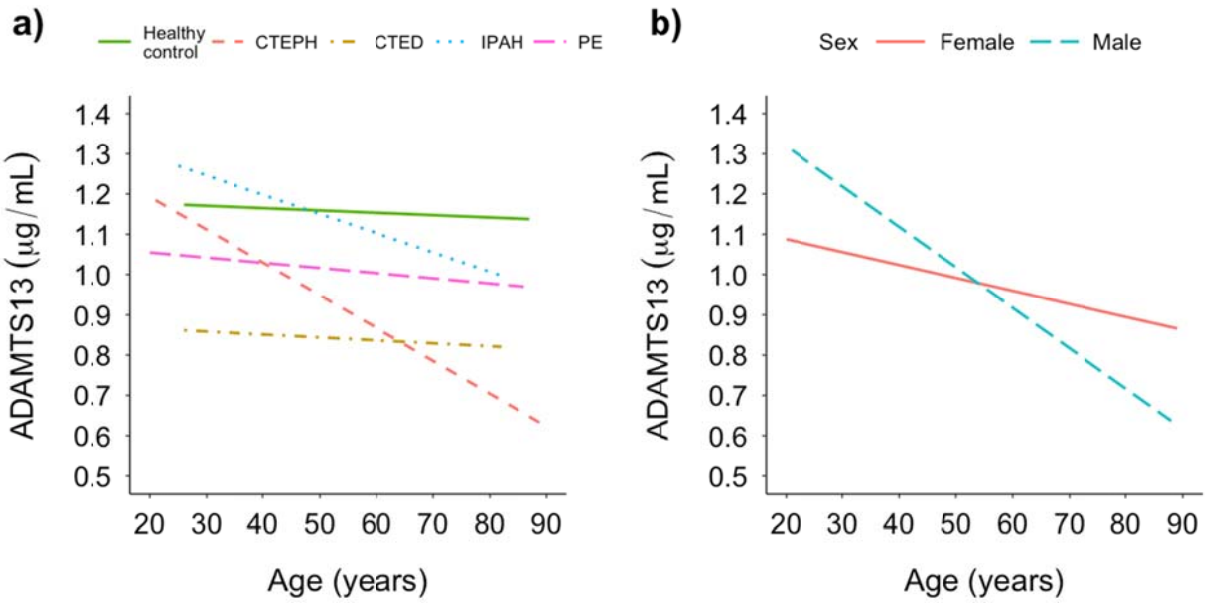

FIGURE S8

Interaction effects for the multivariable linear model shown in supplementary table S9. The predicted ADAMTS13 values are plotted with the interaction terms:

a) Age:Diagnostic group. The ADAMTS13 difference between CTEPH patients and healthy controls is more pronounced for older patients than for younger patients. ADAMTS13 levels remain lower in CTEPH patients across all ages compared with healthy controls. The rate of ADAMTS13 reduction in the CTED groups is similar to healthy controls and is also lower than healthy controls across all ages.

b) Sex:Age. The rate of ADAMTS13 reduction with age is more pronounced for males than for females.

n=343 individuals included in the models.

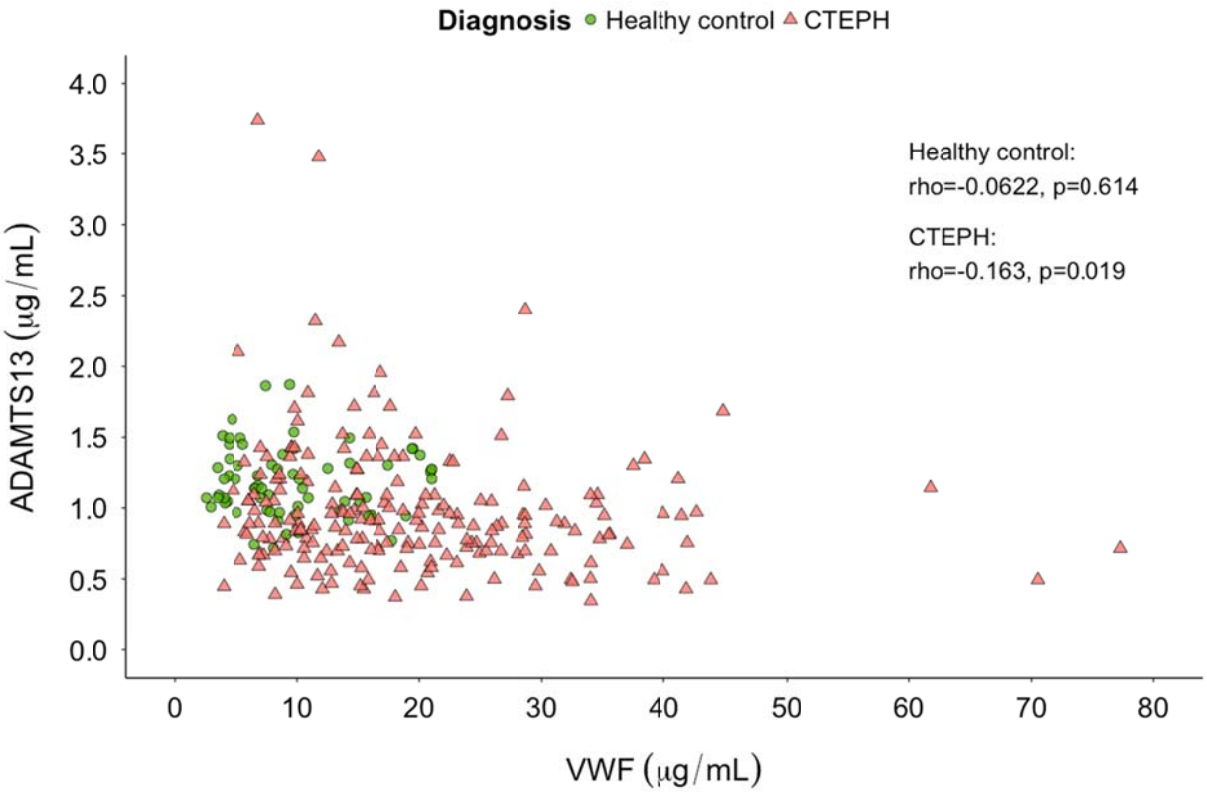

FIGURE S9

Correlation of ADAMTS13 with VWF antigen levels in CTEPH (n=208) and healthy controls (n=68). Spearman's rank correlation in other diagnostic groups: CTED (rho = -0.161, *p* = 0.354), IPAH (rho = 0.329, *p* = 0.076), PE (rho = -0.0504, *p* = 0.799).

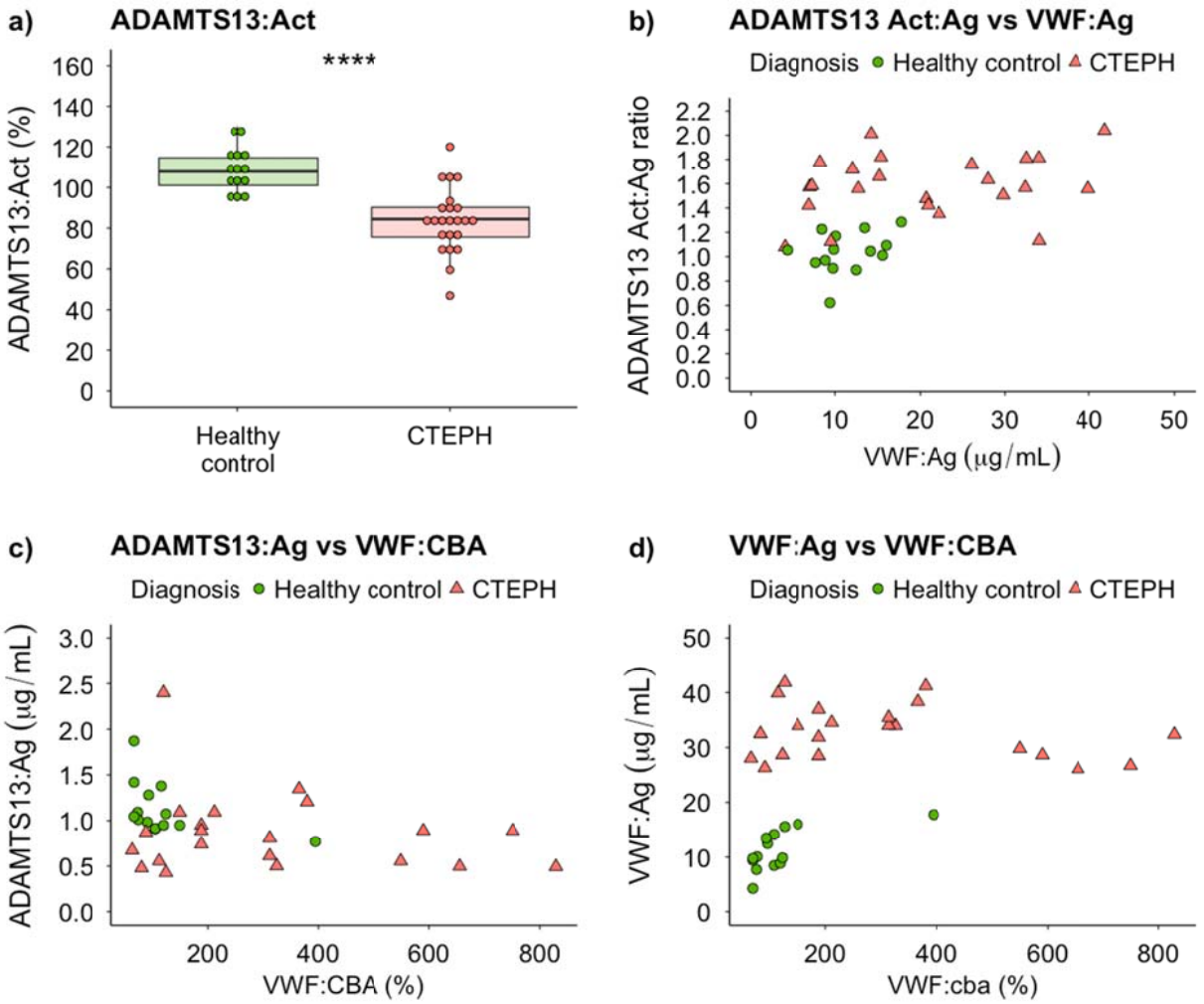

FIGURE S10

ADAMTS13 activity, D-dimer and VWF multimeric size in CTEPH and healthy controls. Additional analysis using the data presented in figure 3 (See figure 3 for group details).

a) Unadjusted ADAMTS13:Act (%) for CTEPH (84.4 ± 14.6 %) and healthy control (107 ± 13.5 %) groups. The Mann-Whitney *U* test was used to calculate group differences.

b) ADAMTS13 Act:Ag ratio vs. VWF:Ag. Healthy control correlation: rho=0.36, *p*=0.210; CTEPH correlation: rho=0.24, *p*=0.270.

c) ADAMTS13:Ag ratio vs. VWF:CBA. Healthy control correlation: rho=-0.61, *p*=0.022; CTEPH correlation: rho=-0.02, *p*=0.930.

d) VWF:Ag vs. VWF:CBA. Healthy control correlation: rho=0.670, *p*=0.008; CTEPH correlation: rho=-0.088, *p*=0.700.

Correlation was assessed with Spearman's rank correlation coefficients for both the healthy control (green circles) and the CTEPH (red triangles) groups (b, c, d).

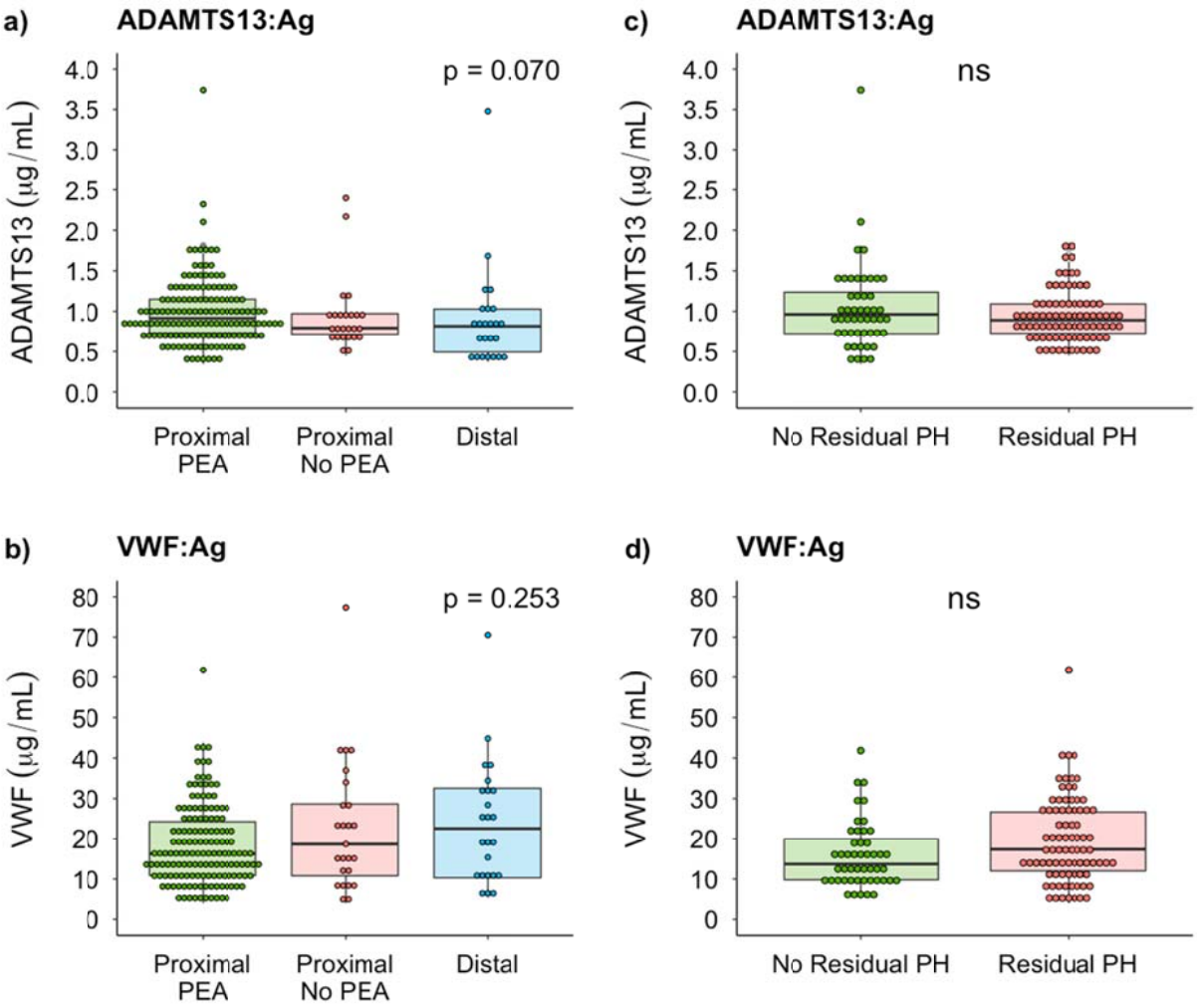

FIGURE S11

ADAMTS13 and VWF antigen levels in CTEPH sub-diagnostic and post-PEA residual pulmonary hypertension groups. a) and b) ADAMTS13 and VWF antigen levels in CTEPH diagnostic sub-groups. Numbers in each group: Proximal PEA = 150, Proximal no PEA = 25, Distal (surgically inaccessible) = 24, insufficient clinical data in 9 patients to classify them. c) and d) ADAMTS13 and VWF antigen levels in CTEPH post-PEA residual pulmonary hypertension (mPAP ≥ 25mmHg) groups. Numbers in each group: No residual PH = 49, Residual PH = 83, insufficient clinical data in 18 patients to

classify them. The group differences were assessed using the Kruskal-Wallis test (a, b) and the Mann-Whitney *U* test (c, d).

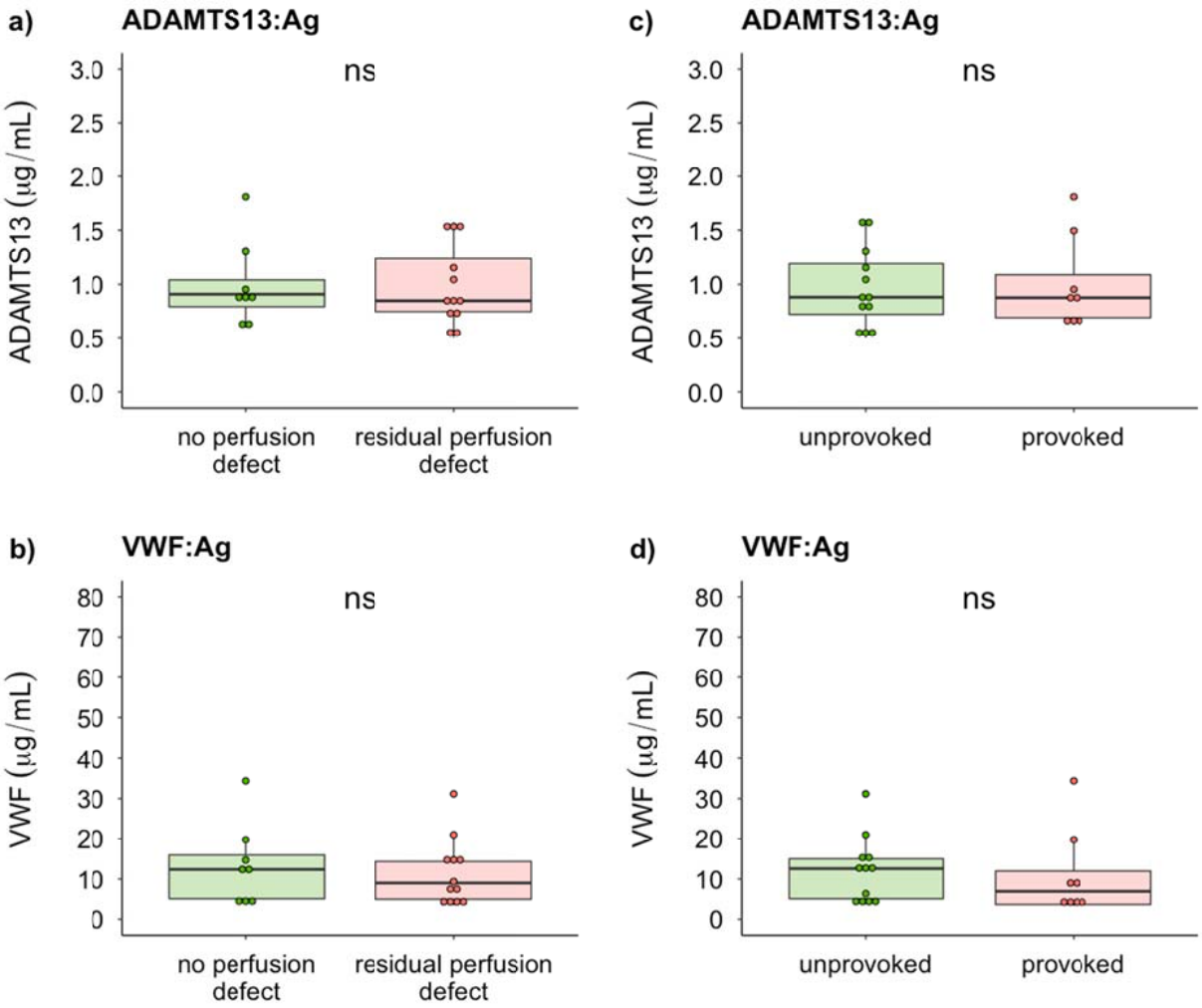

FIGURE S12

ADAMTS13 and VWF antigen levels in PE stratified by residual perfusion defects and provoked PE. The group differences were assessed using the Mann-Whitney *U* test. Numbers in each group: no perfusion defect on VQ scan = 8, residual perfusion defect on VQ scan = 12, unprovoked (no VTE risk factors) = 12, provoked (VTE risk factors) = 8. Of those with residual perfusion defects, the majority were minor (n=10).

Supplementary tables

	n (%)
CTEPH	
Disease distribution	
Proximal	176 (87)
Distal	25 (13)
PEA	150 (72)
Residual PH (>25mmHg)	83 (63)
Co-morbidities	
IHD	20 (14)
DM	19 (13)
Malignancy	19 (13)
Thrombophilia	9 (6)
Splenectomy	9 (6)
Systemic hypertension	48 (34)
Atrial fibrillation / flutter	14 (10)
COPD	8 (6)
PE	
VQ defects post PE	
None	8 (40)
Minor	10 (50)
Major	2 (10)
Idiopathic PE	8 (40)

TABLE S1

Additional clinical phenotype data for the CTEPH and PE groups. COPD (chronic obstructive pulmonary disease), IHD (ischaemic heart disease), DM (diabetes mellitus), PH (pulmonary hypertension), VQ (ventilation-perfusion).

ADAMTS13				
	Healthy control	**CTEPH**	**CTED**	**IPAH**
CTEPH	3.00×10^{-08}	-	-	-
CTED	9.70×10^{-07}	0.205	-	-
IPAH	0.373	0.003	0.001	-
PE	0.049	0.131	0.038	0.294
VWF				
	Healthy control	**CTEPH**	**CTED**	**IPAH**
CTEPH	4.00×10^{-12}	-	-	-
CTED	2.20×10^{-06}	0.834	-	-
IPAH	0.071	0.006	0.021	-
PE	0.433	1.90×10^{-04}	0.002	0.433

TABLE S2

ADAMTS13 and VWF antigen level pairwise diagnostic group comparisons from figure 1. Dunn's test with FDR adjustment was used to calculate *p*-values.

	β (%)	95% CI (%)	*p*	β* (%)	95% CI* (%)	*p**
Healthy Control	Reference					
CTEPH	-23.4	-30.9, -15.1	5.91×10^{-07}	-17.6	-27.9, -5.78	0.005
CTED	-25.9	-35.1, -15.4	1.18×10^{-05}	-23.1	-35.1, -8.75	0.003
IPAH	-2.18	-14.7, 12.2	0.752	1.17	-14.7, 20.0	0.894
PE	-12.0	-24.0, 1.97	0.089	-7.84	-23.3, 10.7	0.381
Female	Reference					
Male	-1.07	-7.57, 5.89	0.756	-1.98	-9.92, 6.65	0.641
Age	-0.518	-0.732, -0.303	3.30×10^{-06}	-0.541	-0.810, -0.271	9.99×10^{-5}
Batch1	Reference					
Batch2	-2.16	-10.5, 6.95	0.630	13.5	1.71, 26.8	0.024
Caucasian	Reference					
Non-Caucasian	-5.58	-15.2, 5.10	0.293	-6.7	-18.4, 6.59	0.306

TABLE S3

Multivariable linear regression model of ADAMTS13 antigen levels. Beta (β) coefficients and 95% confidence intervals are presented as percentage change with respect to healthy controls. The reference diagnostic group is healthy control, the reference sex is female, the reference batch is batch1, the reference ethnicity is Caucasian and the β coefficient for age is per year. n=343 individuals included in the models.

* Model additionally adjusted for VWF antigen levels.

	β (%)	95% CI (%)	*p*
CTEPH	75.5	44.8, 113	2.00×10^{-8}
CTED	89.5	48.0, 143	6.19×10^{-7}
IPAH	26.7	-1.93, 63.7	0.070
PE	19.4	-9.26, 57.2	0.205
Male	7.11	-5.65, 21.6	0.288
Age	0.584	0.180, 0.990	0.005
Batch	7.33	-9.10, 26.7	0.403
Non-Caucasian	-14.5	-30.0, 4.42	0.124

TABLE S4

Multivariable linear regression model of VWF antigen levels. Reference groups are the same as described in supplementary table S3. n=343 individuals included in the model.

Groups	Thresholds	Healthy	CTEPH

		Control	
low VWF & high ADAMTS13	VWF: < 165% (Q1-Q3) ADAMTS13 > 88% (Q2-Q4)	38 (56)	33 (16)
low VWF & low ADAMTS13	VWF: < 165% (Q1-Q3) ADAMTS13: <= 88% (Q1)	13 (19)	45 (22)
high VWF & high adamts13	VWF: >= 165% (Q4) ADAMTS13: > 88% (Q2-4)	12 (18)	39 (19)
high VWF & low ADAMTS13	VWF: >= 165% (Q4) ADAMTS13: <=88% (Q1)	5 (7)	91 (44)

TABLE S5

Summary table for combined ADAMTS13 and VWF groups. Threshold criteria in described in the table with the n (%) for CTEPH and healthy controls. ADAMTS13 and VWF antigen levels were converted to a percentage (the median of the healthy control group) and the quartile thresholds were then determined (healthy control group). The Odds ratios for each group are shown in supplementary figure S2.

	β **(%)**	**95% CI (%)**	***p***
ABO **group - O**	Reference		
ABO **group - A**	19.8	-1.75, 46.1	0.074
ABO **group - B**	51.3	5.30, 117	0.025
ABO **group - AB**	4.41	-26.9, 49.1	0.811
CTEPH	Reference		
CTED	7.43	-18.5, 41.6	0.609
Male	4.40	-11.9, 23.7	0.617
Age	0.921	0.341, 1.50	0.002
Batch	11.8	-5.88, 32.8	0.203
Non-Caucasian	-1.71	-51.7, 100	0.962

TABLE S6

Multivariable linear regression model of VWF antigen levels and genetic *ABO* groups in CTEPH / CTED. The reference ABO group is O and the reference diagnostic group is CTEPH. Otherwise, reference groups are the same as described in supplementary table S3. n=196 included in the model.

rsID	Chr	Position	β (%)	95% CI (%)	*p*
rs3739893	9	136243324	-37.1	-48.1, -23.8	3.78×10^{-6}
rs28407036	9	136252654	-39.0	-51.3, -23.5	2.42×10^{-5}
rs8181039	9	136253927	-37.4	-49.5, -22.5	2.42×10^{-5}
rs78883179	9	136241818	-41.0	-54.1, -24.1	5.05×10^{-5}
rs77533110	9	136286789	-43.0	-56.4, -25.5	5.20×10^{-5}

TABLE S7

Protein quantitative trait loci for ADAMTS13 antigen levels in CTEPH. Associations were assessed using multivariable linear regression and the SNPs included were those in the ADAMTS13 gene ± 40Kilobases (n=396). The model was adjusted for age, sex and batch. A Bonferroni *p*-value threshold $<1.26 \times 10^{-4}$ (0.05/396 variants) was used to denote statistical significance. rsID (reference SNP identification), Chr (chromosome), position (base position). GRCh37 was used for the genomic positions of the SNPs. n=207 individuals included in the model.

	β (%)	95% CI (%)	*p*	Variance (%)
rs3739893	-37.1	-48.1, -23.8	3.78×10^{-6}	7.70
Age	-0.935	-1.21, -0.660	2.23×10^{-10}	16.3
Male	-4.55	-11.7, 3.34	0.253	0.651
Batch	-2.40	-9.67, 5.47	0.537	0.0748

TABLE S8

Multivariable linear regression with the percentage of variance of ADAMTS13 antigen levels explained by SNPs and other characteristics. Reference groups are the same as described in supplementary table S3. Partitioning of the variance explained by each variable within the models was performed by averaging over orders using the R package 'relaimpo' [20]. n=207 individuals included in the model.

	β	95% CI	*p*
(Intercept)	0.827	-0.194, 0.359	0.557
CTEPH	0.192	-0.135, 0.519	0.249
CTED	-0.325	-0.785, 0.136	0.167
IPAH	0.218	-0.236, 0.672	0.345
PE	-0.0728	-0.499, 0.353	0.737
Male	0.286	0.0305, 0.541	0.028
Age	0.00101	-0.00386, 0.00587	0.684
Batch	-0.0237	-0.111, 0.0638	0.594
Non-Caucasian	-0.0535	-0.16, 0.0527	0.322
CTEPH:Age	-0.00768	-0.0133, -0.00207	0.007
CTED:Age	0.000386	-0.00766, 0.00844	0.925
IPAH:Age	-0.00420	-0.012, 0.00358	0.289
PE:Age	-0.000410	-0.00825, 0.00743	0.918
Age:Sex	-0.00500	-0.00917, -0.000817	0.019

TABLE S9

Multivariable linear regression model of log transformed ADAMTS13 antigen levels (dependent variable) with interaction terms. The reference groups for diagnostic group, batch and ethnicity are the same as described in supplementary table S3. The interaction terms included in the model are those that were significant ($p<0.05$) and informative from the combination of variables in supplementary table S3. The beta coefficients should be interpreted with consideration of the interaction effects. For

example, the predicted ADAMTS13 antigen level in an 80-year-old male Caucasian from experimental batch1 would be: exp(0.0827 + 0.192 + (80 x 0.00100) + (0.286) + (80 x -0.00768) + (80 x -0.00500)) = 0.688 µg/mL. This is 34% lower than an 80-year-old male Caucasian healthy control from experimental batch1: exp(0.827 + (80 x 0.00100) + 0.286 + (80 x -0.00500)) = 1.04 µg/mL. n=343 included in the model.

	CTEPH (n=208)	Healthy control (n=68)
Q1 (<88%)	136 (65)	16 (24)
Q2 (88-100%)	24 (12)	18 (26)
Q3 (100-114%)	12 (6)	17 (25)
Q4 (>114%)	36 (17)	17 (25)

TABLE S10

ADAMTS13 antigen level quartiles for CTEPH and healthy controls. ADAMTS13 levels were divided by the median of the healthy control group and expressed as a percentage. The CTEPH group was then divided into quartiles (Q1-Q4) of the ADAMTS13 distribution of the healthy control group.

Supplementary References

1. Taboada D, Pepke-Zaba J, Jenkins DP, et al. Outcome of pulmonary endarterectomy in symptomatic chronic thromboembolic disease. *Eur Respir J* 2014; **44**: 1635-1645.
2. Chion CK, Doggen CJ, Crawley JT, et al. ADAMTS13 and von Willebrand factor and the risk of myocardial infarction in men. *Blood* 2007;**109**:1998-2000.
3. Andersson HM, Siegerink B, Luken BM, et al. High VWF, low ADAMTS13, and oral contraceptives increase the risk of ischemic stroke and myocardial infarction in young women. *Blood* 2012; **119**: 1555-1560.
4. Kokame K, Nobe Y, Kokubo Y, et al. FRETS-VWF73, a first fluorogenic substrate for ADAMTS13 assay. *Br J Haematol* 2005; **129**: 93-100.
5. Pos W, Luken BM, Kremer Hovinga JA, et al. VH1-69 germline encoded antibodies directed towards ADAMTS13 in patients with acquired thrombotic thrombocytopenic purpura. *J Thromb Haemost* 2009; **7**: 421-428.
6. Galie N, Humbert M, Vachiery JL, et al. 2015 ESC/ERS Guidelines for the diagnosis and treatment of pulmonary hypertension: The Joint Task Force for the Diagnosis and Treatment of Pulmonary Hypertension of the European Society of Cardiology (ESC) and the European Respiratory Society (ERS): Endorsed by: Association for European Paediatric and Congenital Cardiology (AEPC), International Society for Heart and Lung Transplantation (ISHLT). *Eur Heart J* 2016; **37**: 67-119.
7. Jenkins D, Madani M, Fadel E, et al. Pulmonary endarterectomy in the management of chronic thromboembolic pulmonary hypertension. *Eur Respir Rev* 2017; **26**: 160111.
8. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature* 2007; **447**: 661-678.
9. GenomeStudio Data Analysis Sofware. Genotyping module v2.0. https://emea.illumina.com/techniques/microarrays/array-data-analysis-experimental-design/genomestudio.html. Accessed Feb 2017.
10. Guo Y, He J, Zhao S, et al. Illumina human exome genotyping array clustering and quality control. *Nat Protoc* 2014; **9**: 2643-2662.
11. Auton A, Brooks LD, Durbin RM, et al. A global reference for human genetic variation. *Nature* 2015; **526**: 68-74.
12. Price AL, Weale ME, Patterson N, et al. Long-range LD can confound genome scans in admixed populations. *Am J Hum Genet* 2008; **83**: 132-135.
13. Loh PR, Danecek P, Palamara PF, et al. Reference-based phasing using the Haplotype Reference Consortium panel. *Nat Genet* 2016; **48**: 1443-1448.
14. Durbin R. Efficient haplotype matching and storage using the positional Burrows-Wheeler transform (PBWT). *Bioinformatics* 2014; **30**: 1266-1272.
15. McCarthy S, Das S, Kretzschmar W, et al. A reference panel of 64,976 haplotypes for genotype imputation. *Nat Genet* 2016; **48**: 1279-1283.
16. Pare G, Chasman DI, Kellogg M, et al. Novel association of ABO histo-blood group antigen with soluble ICAM-1: results of a genome-wide association study of 6,578 women. *PLoS Genet* 2008; **4**: e1000118.

17. Ma Q, Jacobi PM, Emmer BT, et al. Genetic variants in ADAMTS13 as well as smoking are major determinants of plasma ADAMTS13 levels. *Blood Adv* 2017; **1**: 1037-1046.
18. de Vries PS, Boender J, Sonneveld MA, et al. Genetic variants in the ADAMTS13 and SUPT3H genes are associated with ADAMTS13 activity. *Blood* 2015; **125**: 3949-3955.
19. Suhre K, Arnold M, Bhagwat AM, et al. Connecting genetic risk to disease end points through the human blood plasma proteome. *Nat Commun* 2017; **8**: 14357.
20. Groemping U. Relative Importance for Linear Regression in R: The Package relaimpo. *Stat Softw* 2006; **17**: 27.
21. Machiela MJ, Chanock SJ. LDlink: a web-based application for exploring population-specific haplotype structure and linking correlated alleles of possible functional variants. *Bioinformatics* 2015; **31**: 3555-3557.
22. Pena EA, Slate EH. gvlma: Global Validation of Linear Models Assumptions. R package version 1.0.0.2. 2014.
23. Chang CC, Chow CC, et al. Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience* 2015; **4**: 7.
24. Li H, Handsaker B, Wysoker A, et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 2009; **25**: 2078-2079.
25. R Core Team. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. 2017.
26. RStudio Team. RStudio: Integrated Development Environment for R. RStudio, Inc., Boston, MA. 2016.
27. Venables WN, Ripley BD. *Modern Applied Statistics with S.* 4 ed: Springer, New York. 2002.
28. Hothorn T, Hornik K, van de Wiel MA, et al. Implementing a Class of Permutation Tests: The coin Package. *J Stat Softw* 2008; **28**: 23.
29. Pohlert T. The Pairwise Multiple Comparison of Mean Ranks Package (PMCMR). R package. 2014.
30. Zheng X, Levine D, Shen J, et al. A high-performance computing toolset for relatedness and principal component analysis of SNP data. *Bioinformatics* 2012; **28**: 3326-3328.
31. Long JA. jtools: Analysis and Presentation of Social Scientific Data. R package version 0.9.3. 2017.
32. Kennedy N. forestmodel: Forest Plots from Regression Models. R package version 0.4.3. 2017.
33. Wickham H. tidyverse: Easily Install and Load the 'Tidyverse'. R package version 1.2.1. 2017.
34. Suntharalingam J, Machado RD, Sharples LD, et al. Demographic features, BMPR2 status and outcomes in distal chronic thromboembolic pulmonary hypertension. *Thorax* 2007; **62**: 617-622.
35. Cannon JE, Su L, Kiely DG, et al. Dynamic Risk Stratification of Patient Long-Term Outcome After Pulmonary Endarterectomy: Results From the United Kingdom National Cohort. *Circulation* 2016; **133**: 1761-1771.

36. Klok FA, van der Hulle T, den Exter PL, et al. The post-PE syndrome: a new concept for chronic complications of pulmonary embolism. *Blood Rev.* 2014; **28**: 221-226.
37. Schleef M, Strobel E, Dick A, et al. Relationship between ABO and Secretor genotype with plasma levels of factor VIII and von Willebrand factor in thrombosis patients and control individuals. *Br J Haematol* 2005; **128**: 100-107.