Mendelian randomization analysis of red cell distribution width in pulmonary arterial hypertension

Authors

Anna Ulrich, MSc [1], John Wharton, Ph.D. [1], Timothy E. Thayer, M.D. [2], Emilia M. Swietlik, M.D. [3,4], Tufik R Assad, M.D. [5], Ankit A. Desai, M.D. [6], Stefan Gräf, Ph.D. [3,7,8], Lars Harbaum, M.D. [1], Marc Humbert, M.D., Ph.D. [9], Nicholas W. Morrell, M.D. [3,7], William C. Nichols, Ph.D. [10], Florent Soubrier, M.D., Ph.D. [11], Laura Southgate, Ph.D. [12], David-Alexandre Trégouët, Ph.D. [13], Richard C. Trembath, F.R.C.P. [14], Evan L. Brittain, M.D. [2,15], Martin R. Wilkins, M.D. [1], Inga Prokopenko*, Ph.D [16,17], Christopher J. Rhodes*, Ph.D. [1]

on behalf of The NIHR BioResource – Rare Diseases Consortium‡, UK PAH Cohort Study Consortium‡ and the US PAH Biobank Consortium‡

‡Members listed in the appendix

*These authors contributed equally to this study

Affiliations

[1] National Heart and Lung Institute, Hammersmith Campus, Imperial College London, London, United Kingdom;

[2] Vanderbilt University Medical Center, Division of Cardiovascular Medicine, Nashville TN, USA;

[3] Department of Medicine, University of Cambridge, Cambridge, United Kingdom;

[4] Pulmonary Vascular Disease Unit, Royal Papworth Hospital NHS Foundation Trust, Cambridge, United Kingdom;

[5] Williamson Country Medical Center, Franklin TN, USA;

[6] Indiana University, Indianapolis IN, USA

[7] NIHR BioResource - Rare Diseases, Cambridge, United Kingdom;

[8] Department of Haematology, University of Cambridge, Cambridge, United Kingdom;

[9] Université Paris-Sud, Faculté de Médecine, Université Paris-Saclay; AP-HP, Service de Pneumologie, Centre de référence de l'hypertension pulmonaire, Hôpital Bicêtre, Le Kremlin-Bicêtre; INSERM UMR_S 999, Hôpital Marie Lannelongue, Le Plessis Robinson, Paris, France;

[10] Division of Human Genetics, Cincinnati Children's Hospital Medical Center, Department of Pediatrics, University of Cincinnati College of Medicine, Cincinnati OH, USA;

[11] Sorbonne Universités, UPMC Univ. Paris 06, Institut National pour la Santé et la Recherche Médicale (INSERM), Unité Mixte de Recherche en Santé (UMR_S) 1166, Paris, France;

[12] Molecular and Clinical Sciences Research Institute, St George's University of London, London, United Kingdom;

[13] INSERM UMR_S 1219, Bordeaux Population Health research center, University of Bordeaux, Bordeaux, France;

[14] Division of Genetics and Molecular Medicine, King's College London, London, United Kingdom;

[15] Vanderbilt Translational and Clinical Cardiovascular Research Center, Nashville TN, USA;

[16] Department of Clinical and Experimental Medicine, University of Surrey, Guildford, United Kingdom;

[17] Department of Metabolism, Digestion and Reproduction, Imperial College London, London, United Kingdom;

Corresponding author's contact details:

Christopher J. Rhodes*, National Heart and Lung Institute, Medicine, Imperial College London, London, W12 0NN, United Kingdom. email: crhodes@imperial.ac.uk*

Pulmonary arterial hypertension (PAH) is a rare disease that leads to premature death from right heart failure. It is strongly associated with elevated red cell distribution width (RDW), a correlate of several iron status biomarkers. High RDW values can signal early stage iron deficiency or iron deficiency anaemia. This study investigated if elevated RDW is causally associated with PAH.

A two-sample Mendelian randomization (MR) approach was applied to investigate whether genetic predisposition to higher levels of RDW increases the odds of developing PAH. Primary and secondary MR analyses were performed using all available genome-wide significant RDW variants (n = 179) and five genome-wide significant RDW variants that act via systemic iron status, respectively.

We confirmed the observed association between RDW and PAH (OR = 1.90, 95% CI = 1.80 - 2.01) in a multi-centre case-control study (*N* cases = 642, *N* disease controls = 15,889). The primary MR analysis was adequately powered to detect a causal effect (OR) from between 1.25-1.52 or greater based on estimates reported in the RDW GWAS or from our own data. There was no evidence for a causal association between RDW and PAH in either the primary (OR_{causal} = 1.07, 95% CI = 0.92 – 1.24) or the secondary (OR_{causal} = 1.09, 95% CI = 0.77 – 1.54) MR analysis.

The results suggest that at least some of the observed association of RDW with PAH is secondary to disease progression. Results of iron therapeutic trials in PAH should be interpreted with caution as any improvements observed may not be mechanistically linked to the development of PAH.

Take home message – Mendelian randomization using genetic data from the largest-to-date pulmonary arterial hypertension (PAH) cohort do not support RDW or iron deficiency as a cause of PAH, which is important when interpreting iron replacement trials in this condition.

Introduction

Pulmonary arterial hypertension (PAH) is a rare disease with an estimated prevalence of 7-26 cases/million in the developed world [1]. It is characterized by increased pulmonary vascular resistance due to vasoconstriction and structural remodeling of pulmonary arterioles, leading to right ventricular hypertrophy and end-stage right heart failure [2]. Despite increased awareness and new therapeutic options, annual mortality remains around 10% [1]. Approximately 70–80% of heritable PAH and 10–20% of idiopathic PAH patients are known to harbour mutations in the bone morphogenetic protein type II receptor (*BMPR2*) gene [3]. A recent large study of over 1,000 PAH patients confirmed the prevalence of causal mutations in *BMPR2*, as well as in five other established genes (*TBX4, ACVRL1, ENG, SMAD9*, and *KCNK3*), and identified PAH-associated mutations in four new genes (*ATP13A3, SOX17, AQP1, and GDF2*), all together accounting for 23.5% of the cases studied [4]. The rare mutations in all these genes are inherited in an autosomal dominant manner and exhibit reduced penetrance, indicating that other genetic, epigenetic and/or environmental factors influence the development of PAH.

We and others have demonstrated that one factor strongly correlated with survival in PAH is red cell distribution width (RDW) [5, 6]. A recent hypothesis-free phenome-wide analysis indicated PAH, among several disease descriptors, as the most strongly associated with RDW (odds ratio, OR = 2.0, 95% CI = 1.75-2.4 per % increase in RDW) in a hospital population [7]. RDW - a measure of red blood cell (RBC) size variability in an individual – is part of the full blood count in standard hospital practice and readily available as a biomarker. RDW correlates with iron status biomarkers and high values can signal early stage iron deficiency or iron deficiency anemia [8]. RDW increases with decreasing iron as available body iron stores fail to meet the iron demand of RBC synthesis resulting in RBCs of varied size.

Iron deficiency is commonly observed in PAH patients and is under investigation as a therapeutic target [9-14]. An observed correlation between two traits does not necessarily imply that interventions on one trait will change the other and there are numerous examples where false positive associations have led to unsuccessful randomized controlled trials [15, 16]. Clinical trials are expensive and time-consuming, and recruitment can be challenging, especially in rare diseases. With the growing availability of genetic data in large disease-focused and population-based studies, testing causal relationships between traits of interest has become possible by harnessing the naturally occurring genetic variation in the population. The collection of methods used to test causal relationships using genetic variants is called Mendelian randomization (MR) [17, 18]. MR has been successfully used to help prioritize intervention and drug targets and to identify causal factors for several diseases [19-23]. In general, candidate drugs with genetic evidence for effectiveness are more successful in drug trials compared to those without such genetic support [24].

It remains to be investigated whether elevated RDW is largely a consequence of PAH or plays a causal role in the condition. The common genetic variation determining RDW levels has been defined in very large population studies [25], with more power than equivalent studies of iron status. This makes RDW a strong candidate for MR analysis and our primary aim was to dissect the epidemiological association between RDW and PAH by assessing causality using MR. We refined the estimate of association between RDW and PAH using biomarker data in 642 PAH cases and over 15,000 controls with common diseases from the *Imperial College PH Biobank*, the *UK PAH Cohort*

and the *Vanderbilt University Medical Centre (VUMC)* and applied MR to investigate whether being genetically predisposed to higher levels of RDW increases the odds of developing PAH.

Methods

Definition of pulmonary arterial hypertension (PAH)

PAH was defined by internationally agreed criteria [26], specifically mean pulmonary artery pressure >25mmHg, pulmonary vascular resistance >3 Woods and pulmonary capillary wedge pressure <15 mmHg. Patients with concurrent diseases known to cause pulmonary hypertension were excluded such that all were considered to have idiopathic, heritable or anorexigen-induced PAH.

Data: genetic and phenotype data in contributing studies

For our analyses we used both individual-level and summary-level data. Individual-level data including clinical laboratory RDW was available for a total of 524 PAH patients from the Imperial College PH Biobank and the multicentre *UK PAH Cohort,* a study that facilitates collaboration and the sharing of data and samples between the specialist pulmonary hypertension centres in the United Kingdom [27].

In addition, the hospital population-based *VUMC* study provided longitudinal clinical laboratory RDW measurements, detailed clinical diagnoses and genome-wide genotype array data (genotyping platform: Illumina MEGAex) for an additional 118 PAH patients and 15,889 common disease controls (*Supplementary Table 1*). *VUMC* hosts a collection of electronic medical records linked to genetic data derived from blood collected during routine clinical assessment in outpatient clinics where all patients are shown the consent form during check-in [28, 29]. The exclusion criteria (*Supplementary Figure 1*) and the imputation of genotype array data are described in the *Supplementary Methods*.

Summary-level data for both RDW and PAH susceptibility were used in the MR analyses to maximize the sample size and therefore the statistical power. Genetic instruments serving as a proxy for RDW were obtained from the largest-to-date (over 170,000 individuals) GWAS on hematological traits (hereon referred to as the 'RDW GWAS') in a population-based study by Astle et al. (25).

A large proportion of the idiopathic and familial PAH cases from the *UK PAH Cohort* were whole genome sequenced as part of the *UK National Institute for Health Research BioResource (NIHRBR) [4, 30]* study. For genetic effects on PAH susceptibility, we used a recent study published by our group and others investigating in the largest-to-date genome-wide association study the effects of common genetic variation on PAH risk (PAH GWAS) involving a total of 11,744 individuals of which 2,085 were PAH cases. The results of the PAH GWAS were combined through the meta-analysis of four independent studies, one of which is the *NIHRBR* with 847 PAH cases and 5,048 rare disease controls. The other major contributing study, the *US PAH Biobank (PAHB)* used a control population with mixed common diseases recruited from outpatient clinics (694 PAH cases and 1560 controls), whilst the two smaller studies used population-based controls, with 269/275 PAH cases and 1068/1983 controls in the PHAAR and BHFPAH studies, respectively [4].

Statistical analyses

To confirm and refine the estimate for the association between RDW and PAH, we combined PAH patients from the *Imperial College PH Biobank*, the *UK PAH Cohort* and the *VUMC* and compared them to common disease controls from the *VUMC* recruited in outpatient clinics (*Supplementary Methods, Supplementary Table 1*).

To test for causality between RDW and PAH, we applied a two-sample MR strategy that requires effect estimates for the genetic instrument on the risk factor (here RDW) and the outcome (here PAH) from two non-overlapping datasets (*Supplementary Methods and Supplementary Figure 2*). The genetic instrument for RDW comprised genetic variants associated with RDW levels in the RDW GWAS at a study-specific genome-wide level of significance ($P< 8.31\times10^{-9}$). In the PAH GWAS, 179 variants (inclusive of 12 proxy variants with a minimum r^2 of 0.8) out of the 212 independent ($r^2 <$ 0.01) RDW quantitative trait loci (QTL) were available after excluding 13 palindromic variants (A/T or C/G) with intermediate allele frequencies (minor allele frequency > 45%) to ensure that the effects of the variants for the two traits can be harmonized to the same allele. To obtain the causal estimate, we applied the inverse variance weighted (IVW) [31] - and weighted median [32] methods. We assessed heterogeneity between the causal estimates from each QTL using the Cochran's Q test (*Supplementary Methods*).

In the primary MR analysis, we included all available genome-wide significant RDW QTL (n = 179). The secondary analysis explicitly tested the role of iron deficiency in the RDW-PAH association using five RDW QTL mapped to genes involved in iron homeostasis (*HFE, TMPRSS6, TFRC* and *TFR2*) from the full set of genome-wide significant RDW QTL (*Supplementary Figure 2*). All of these five RDW QTL concomitantly increase serum iron, ferritin, and transferrin saturation and decrease transferrin, reflecting systemic iron status (*Supplementary Table 2*) and were first reported by an independent GWAS - the Genetics of Iron Status GWAS - as genome-wide significant signals for at least two of the above-mentioned iron status biomarkers [33]. These five RDW QTL are among the signals which explain the highest proportion of variance in the RDW GWAS, highlighting the importance of iron availability in RDW levels.

Furthermore, we validated the RDW genetic instrument as a proxy for RDW levels in our common disease controls from VUMC. To do so, we regressed the first available RDW measurement on the RDW genetic risk score (GRS) constructed for each individual (*Supplementary Methods*) and obtained the proportion of variance explained (R^2). The R^2 for the same RDW GRS in the RDW GWAS study populations (UK Biobank – UKB and INTERVAL) was calculated from the summary statistics of the RDW-QTL associations (*Supplementary Methods*).

Since genetic variants typically explain a small proportion of the variability in the associated trait, MR studies often require large sample sizes to detect a non-zero causal effect. Our power to detect a causal association in the current MR analyses was calculated [34] using the R^2 values estimated in VUMC and also those estimated in the RDW GWAS populations (*Supplementary Methods*).

Results

Defining the association of RDW and PAH

Within this observational analysis, each standard unit (1.4%) increase in RDW was associated with 90% higher odds of prevalent PAH after adjusting for the effects of age and sex (odds ratio, OR = 1.90, 95% CI = 1.80 - 2.01). There were no marked between-cohort differences in RDW levels (*Figure 1* and *Supplementary Table 1*).

Genetic risk score using RDW QTL

We estimated that the 179 RDW QTL would explain over 12% of the variability (R^2 = 12.7%, 95% CI = 12.32% -12.99%) in RDW levels in the RDW GWAS population (UK Biobank and INTERVAL) in which they were discovered. The RDW GRS constructed for the VUMC controls explained 2.6% (95% CI = 2.17% - 3.19%) of the variability in the first available RDW measurement *(Supplementary Table 3)*. The set of five variants specific to iron status explain an estimated 1.7% (95% CI = 1.62% – 1.87%) of the RDW variability in the UKB and INTERVAL populations whilst they explain 0.7% (95% CI = 0.43% - 0.92%) of the total variability in RDW in the VUMC controls *(Supplementary Table 3)*. The RDW GWAS study populations had a lower mean RDW level than our common disease controls (UKB and INTERVAL = 13.45; VUMC = 13.60) and lower standard deviation (UKB and INTERVAL = 0.82; VUMC = 1.40) which could in part explain differences between the R^2 estimates in these studies.

Statistical power to detect causal effect

With a genetic instrument that explains a relatively high proportion (R^2 = 12%) of the total variation in RDW levels (Figure 2, red curve) we have 80% power to detect a causal effect as small as 1.25 (OR). If the true variance explained lies closer to that estimated in the VUMC controls (R2 = 2.6%, Figure 2, green curve), this changes to 1.52. When the variance explained by the genetic instrument is small (R^2=1.7%, Figure 2, blue curve), we are limited, with our sample size, to an OR of 1.7 or above. However, if the effect of RDW calculated in our observational analysis (OR=1.80-2.01) was causal in nature, either of the two MR analyses, based on R^2 estimates from the RDW GWAS, would have over 80% power to detect an effect of that magnitude. Using the estimates based on the VUMC data, the analysis using all 179 RDW QTL (Figure 2, green line) would have sufficient (>80%) power in our sample, while the iron-specific model (Figure 2, purple line) would be underpowered.

RDW-PAH causal relationship

We tested for a causal effect of RDW on development of PAH in our primary MR analysis using 179 RDW QTL and found no significant relationship (IVW OR_{causal} = 1.07, 95% CI = 0.93 – 1.23, Q p-value = 0.57, Figure 3). A secondary MR analysis based on five RDW QTL provided no evidence for a causal effect of iron status on PAH (IVW OR_{causal} = 1.09, 95% CI = 0.77 – 1.54, Q p-value = 0.91, Figure 3). The weighted median method, which is more robust to violations of MR instrument assumptions, yielded similar estimates for both the primary (WM OR_{causal} = 1.11, 95% CI = 0.89 – 1.38) and the secondary (WM OR_{causal} = 1.04, 95% CI = 0.68 – 1.59) MR analyses.

If the OR estimated in the primary MR analysis was indicative of the magnitude of a real causal effect, the number of PAH cases needed to detect such a causal effect with a genetic instrument that explains 10% of the variance in RDW would be ~ 20,600 (with the same 1:4.6 ratio of cases:controls as in this study) to achieve 80% power at a false positive rate of 5% (P = 0.05). We tested for heterogeneity between causal effects estimated in the four studies contributing to the PAH GWAS separately (Supplementary Figures 1-2) to assess if differences in the nature of their control populations yielded heterogenous effect estimates for the instrumental variants. The two heterogeneity tests on the IVW estimates (main MR: Q = 0.83, df = 3, p = 0.84; secondary MR: Q = 0.55, df = 3, p = 0.91) did not detect considerable variability between the causal effects in the four contributing studies based on a random-effects model (*Supplementary Figures 3 and 4*).

Discussion

We applied a two-sample MR approach to test whether the epidemiological relationship between elevated RDW levels, which are associated with iron deficiency, and PAH is causal in nature. We estimated the effect of RDW on PAH in a large sample of cases and common disease controls. By using genetic variants as instruments for RDW, we found no evidence for a causal effect of RDW on PAH of the magnitude suggested by observational studies.

Previous work has shown that iron deficiency is common in PAH and associated with a poor prognosis, reduced exercise capacity and worsening hemodynamics [9, 11, 13]. A physiological link has been described in healthy volunteers where iron infusion attenuated the rise in pulmonary artery pressure induced by acute hypoxia [35, 36] and in rats where chronic iron deficiency results in pulmonary hypertension [37]. This relationship could be driven by the role of iron in de-stabilizing the hypoxia-inducible factor, thereby deficiency can mimic the hypoxic state [38]. Our study confirmed the association of raised RDW with PAH using controls from a hospital population and cases from multiple centres (with an effect size 1.90 for one standard unit RDW, 1.4%), supporting a recent hypothesis-free phenome-wide analysis which indicated, among several disease descriptors, PAH as the most strongly associated with RDW with a similar effect size (OR 2.0, 95% CI = 1.75-2.4 per % increase in RDW) [7]. Our MR analysis was adequately powered to detect a causal role for RDW with an effect of this magnitude. The fact that we did not detect a causal effect at this level suggests that at least some of the observed association is secondary to the disease.

The results of this study may appear to be at odds with previous clinical studies of the efficacy of iron supplementation in PAH patients, which have focused on functional capacity rather disease pathology. An open-label study of twenty patients with idiopathic PAH with iron deficiency reported improved iron status, 6-minute walk distance (6MWD) and quality of life (QoL) two months after a single infusion of 1000 mg ferric carboxymaltose [14]. Another open-label study in fifteen iron deficient idiopathic PAH patients reported improvement of iron status, QoL and exercise endurance capacity on cardiopulmonary exercise testing after receiving 1000 mg of intravenous iron [10]. Neither of the clinical studies were placebo controlled, although Viethen *et al.* compared their intervention group to a group of matched iron-replete patients who did not receive iron infusion. It remains possible that iron supplementation in PAH could have benefits through mechanisms distinct from those driving the cardiovascular pathology, for example on muscle function [39].

MR studies using data from large consortia support a causal effect of iron status in other diseases. The genetic instruments (two variants in *HFE* and one variant in *TMPRSS6*) used in these studies were also used in our secondary MR analysis. Gill and colleagues found iron to have a protective effect against coronary artery disease (IVW causal OR = 0.94 per standard deviation (SD) change in serum iron) [40] but increase the risk of cardioembolic stroke (IVW causal OR = 1.16 per SD iron) [41]. The authors suggest the opposing effects of iron status on CAD and stroke might be due to different underlying mechanisms. Pichler et al. [19] have reported a that iron protects against the risk of developing Parkinson`s disease (IVW causal OR = 0.88 per SD iron). Iron deficiency is a common risk factor and these causal effect estimates in common diseases are modest; this study was not powered to detect an effect if this is also true of PAH.

In the light of this MR analysis, alternative explanations for the association of RDW and PAH have to be considered; namely, that PAH causes raised RDW (reverse causation, for example reduced oxygen delivery and/or haemolysis related to PAH may stimulate reticulocytosis which would increase RDW) or that PAH and elevated RDW are caused by an independent common mechanism. One such mechanism is chronic inflammation, which is a common feature of PAH [42] and leads to intracellular sequestration of iron. Other mechanisms which may modify RDW such as folate or vitamin B12 could be studied for their association with PAH. To directly test whether PAH is causal for raised RDW levels a stronger genetic instrument for PAH is required than currently known common variation identified by PAH GWAS.

An important strength of our study lies in the sample size available with phenotype and genetic data achieved through extensive collaboration. Given the rarity of PAH, data had to be pooled from several centres to allow the investigation of common genetic variation in PAH and to test causal relationships. Our study also has some limitations. The control population for our observational study was not specifically selected to represent a population at risk of developing PAH. An example of a population at risk of PAH are relatives of patients with pathogenic *BMPR2* mutations (and other rare pathogenic variants). Given the reduced penetrance of familial/heritable PAH (~42% in women and ~14% in men carrying known mutations) [43], following up the families of affected individuals, especially relatives harbouring mutations associated with PAH, would be an invaluable source to identify environmental triggers of PAH.

Despite our efforts to exclude controls with conditions that likely affect RDW and to use the first available RDW measurement, we expect genetic effects of RDW levels to differ between individuals with common diseases and a cohort of healthy volunteers. The R^2 of a genetic variant describes the variance explained in the phenotype in a given population at a given time. Therefore, there can be no single population parameter that applies to multiple populations or the same population at multiple timepoints. Although the R^2 estimated in the RDW GWAS is likely upwardly biased (since it contains the discovery as well as the replication samples), it might reflect better the extent to which genetic variation influences long-term RDW levels in disease-free populations than the R^2 estimate from the VUMC disease controls, who may be expected to have more variable RDW levels due to comorbidities. This highlights a potential challenge in estimating power for MR studies; we now estimate that upwards of 20,000 PAH patients would be required to detect any likely causal effect of iron deficiency.

It is important to note that in our MR analyses, causal effects were estimated using the results of the PAH GWAS [4]. Four independent studies contributed to the overall results of the PAH GWAS comparing allele frequencies of their PAH cases to control cohorts selected according to different criteria. The NIHRBR study used a control population with a mixture of rare diseases while the other major contributing study from the United States used mixed common disease controls. The two smaller contributing studies used population-based controls; a preferable design for estimating the effects of surrogate genetic variants for common risk factors. Selection bias can affect the estimates of the instrumental genetic variants on disease susceptibility. This is especially true if the control cohort is enriched for conditions also affected by the risk factor of interest.

Conclusions

There is strong observational evidence for an association between elevated RDW, a surrogate for iron deficiency, and PAH. However, this Mendelian randomization analysis does not indicate that RDW is causally linked to disease development. Our study was powered to detect a causal effect similar in size to that observed. A more modest causal effect remains possible but would require a significantly larger study population to detect. Extending international collaborations and careful follow up of populations at risk will allow increasingly sophisticated study designs to investigate causal relationships, shared underlying mechanisms with other conditions and overall genetic susceptibility in PAH.

Acknowledgements

We gratefully acknowledge the participation of patients recruited to the UK National Institute of Health Research BioResource - Rare Diseases (NIHR BR-RD) Study, the UK National Cohort of Idiopathic and Heritable PAH, and the National Institutes of Health/National Heart, Lung, and Blood Institute-sponsored National Biological Sample and Data Repository for PAH (aka PAH Biobank). We thank the physicians, research nurses and coordinators in the UK, Europe and at the 38 pulmonary hypertension centers across the US involved in the PAH Biobank (www.pahbiobank.org). The work cited here is supported by funding from the NIHR BR-RD, the British Heart Foundation (SP/12/12/29836), the BHF Cambridge and Imperial Centres of Cardiovascular Research Excellence (RE/18/4/34215), UK Medical Research Council (MR/K020919/1), the Dinosaur Trust, and BHF Programme grants to RCT (RG/08/006/25302), NWM (RG/13/4/30107) and MRW (RG/10/16/28575). Funding for the PAH Biobank is provided by NIH/NHLBI HL105333. Vanderbilt University Medical Center's BioVU is supported by numerous sources: institutional funding, private agencies, and federal grants that include the NIH funded Shared Instrumentation Grant S10RR025141; and CTSA grants UL1TR002243, UL1TR000445, and UL1RR024975. Genomic data are also supported by investigator-led projects that include U01HG004798, R01NS032830, RC2GM092618, P50GM115305, U01HG006378, U19HL065962, R01HD074711; and additional funding sources listed at https://victr.vanderbilt.edu/pub/biovu/. ELB receives funding from the NIH R01 HL146588, American Heart Association Fellow to Faculty Grant (13FTF16070002) and the Gilead PAH Scholars Award Program. The genotyping of the VESPA samples was supported by RC2GM092618. The authors acknowledge use of BRC Core Facilities provided by the financial support from the Department of Health via the National Institute for Health Research (NIHR) comprehensive Biomedical Research Centre award to Imperial College NHS Trust, Cambridge University Hospitals and Guy's and St Thomas' NHS Foundation Trust in partnership with King's College London and King's College Hospital NHS Foundation Trust and by NIHR funding to the Imperial NIHR Clinical Research Facility. CJR is supported by a British Heart Foundation Intermediate Basic Science Research Fellowship (FS/15/59/31839). LS is supported by the Wellcome Trust Institutional Strategic Support Fund (204809/Z/16/Z) awarded to St. George's, University of London. IP is supported by the Wellcome Trust (WT205915), and the EU H2020 (DYNAhealth, project number 633595). NWM is a British Heart Foundation Professor and National Institute of Health Research (NIHR) Senior Investigator. WCN is supported by NIH NHLBI HL105333. AAD receives support from NIH NHLBI R01HL136603.

References

1. Hoeper MM, Humbert M, Souza R, Idrees M, Kawut SM, Sliwa-Hahnle K, Jing Z-C, Gibbs JSR. A global view of pulmonary hypertension. *The Lancet Respiratory Medicine* 2016: 4(4): 306-322.
2. Stacher E, Graham BB, Hunt JM, Gandjeva A, Groshong SD, McLaughlin VV, Jessup M, Grizzle WE, Aldred MA, Cool CD, Tuder RM. Modern age pathology of pulmonary arterial hypertension. *American journal of respiratory and critical care medicine* 2012: 186(3): 261-272.
3. Evans JDW, Girerd B, Montani D, Wang X-J, Galiè N, Austin ED, Elliott G, Asano K, Grünig E, Yan Y, Jing Z-C, Manes A, Palazzini M, Wheeler LA, Nakayama I, Satoh T, Eichstaedt C, Hinderhofer K, Wolf M, Rosenzweig EB, Chung WK, Soubrier F, Simonneau G, Sitbon O, Gräf S, Kaptoge S, Di Angelantonio E, Humbert M, Morrell NW. BMPR2 mutations and survival in pulmonary arterial hypertension: an individual participant data meta-analysis. *The Lancet Respiratory Medicine* 2016: 4(2): 129-137.
4. Rhodes CJ, Batai K, Bleda M, Haimel M, Southgate L, Germain M, Pauciulo MW, Hadinnapola C, Aman J, Girerd B, Arora A, Knight J, Hanscombe KB, Karnes JH, Kaakinen M, Gall H, Ulrich A, Harbaum L, Cebola I, Ferrer J, Lutz K, Swietlik EM, Ahmad F, Amouyel P, Archer SL, Argula R, Austin ED, Badesch D, Bakshi S, Barnett C, Benza R, Bhatt N, Bogaard HJ, Burger CD, Chakinala M, Church C, Coghlan JG, Condliffe R, Corris PA, Danesino C, Debette S, Elliott CG, Elwing J, Eyries M, Fortin T, Franke A, Frantz RP, Frost A, Garcia JGN, Ghio S, Ghofrani H-A, Gibbs JSR, Harley J, He H, Hill NS, Hirsch R, Houweling AC, Howard LS, Ivy D, Kiely DG, Klinger J, Kovacs G, Lahm T, Laudes M, Machado RD, Ross RVM, Marsolo K, Martin LJ, Moledina S, Montani D, Nathan SD, Newnham M, Olschewski A, Olschewski H, Oudiz RJ, Ouwehand WH, Peacock AJ, Pepke-Zaba J, Rehman Z, Robbins I, Roden DM, Rosenzweig EB, Saydain G, Scelsi L, Schilz R, Seeger W, Shaffer CM, Simms RW, Simon M, Sitbon O, Suntharalingam J, Tang H, Tchourbanov AY, Thenappan T, Torres F, Toshner MR, Treacy CM, Noordegraaf AV, Waisfisz Q, Walsworth AK, Walter RE, Wharton J, White RJ, Wilt J, Wort SJ, Yung D, Lawrie A, Humbert M, Soubrier F, Trégouët D-A, Prokopenko I, Kittles R, Gräf S, Nichols WC, Trembath RC, Desai AA, Morrell NW, Wilkins MR. Genetic determinants of risk in pulmonary arterial hypertension: international genome-wide association studies and meta-analysis. *The Lancet Respiratory Medicine* 2018.
5. Rhodes CJ, Wharton J, Howard LS, Gibbs JS, Wilkins MR. Red cell distribution width outperforms other potential circulating biomarkers in predicting survival in idiopathic pulmonary arterial hypertension. *Heart* 2011: 97(13): 1054-1060.
6. Hampole CV, Mehrotra AK, Thenappan T, Gomberg-Maitland M, Shah SJ. Usefulness of red cell distribution width as a prognostic marker in pulmonary hypertension. *Am J Cardiol* 2009: 104(6): 868-872.
7. Thayer TE, Huang S, Levinson RT, Farber-Eger E, Assad TR, Huston JH, Mosley JD, Wells QS, Brittain EL. Unbiased Phenome-wide Association Studies of Red Cell Distribution Width Identifies Key Associations with Pulmonary Hypertension. *Annals of the American Thoracic Society* 2019.
8. Emans ME, van der Putten K, van Rooijen KL, Kraaijenhagen RJ, Swinkels D, van Solinge WW, Cramer MJ, Doevendans PAFM, Braam B, Gaillard CAJM. Determinants of Red Cell Distribution Width (RDW) in Cardiorenal Patients: RDW is Not Related to Erythropoietin Resistance. *Journal of Cardiac Failure* 2011: 17(8): 626-633.
9. Ruiter G, Lankhorst S, Boonstra A, Postmus PE, Zweegman S, Westerhof N, van der Laarse WJ, Vonk-Noordegraaf A. Iron deficiency is common in idiopathic pulmonary arterial hypertension. *Eur Respir J* 2011: 37(6): 1386-1391.
10. Ruiter G, Manders E, Happé CM, Schalij I, Groepenhoff H, Howard LS, Wilkins MR, Bogaard HJ, Westerhof N, van der Laarse WJ, de Man FS, Vonk-Noordegraaf A. Intravenous iron therapy in patients with idiopathic pulmonary arterial hypertension and iron deficiency. *Pulmonary Circulation* 2015: 5(3): 466-472.

11. Soon E, Treacy CM, Toshner MR, MacKenzie-Ross R, Manglam V, Busbridge M, Sinclair-McGarvie M, Arnold J, Sheares KK, Morrell NW, Pepke-Zaba J. Unexplained iron deficiency in idiopathic and heritable pulmonary arterial hypertension. *Thorax* 2011: 66(4): 326-332.
12. Rhodes CJ, Wharton J, Howard L, Gibbs JS, Vonk-Noordegraaf A, Wilkins MR. Iron deficiency in pulmonary arterial hypertension: a potential therapeutic target. *Eur Respir J* 2011: 38(6): 1453-1460.
13. Rhodes CJ, Howard LS, Busbridge M, Ashby D, Kondili E, Gibbs JS, Wharton J, Wilkins MR. Iron deficiency and raised hepcidin in idiopathic pulmonary arterial hypertension: clinical prevalence, outcomes, and mechanistic insights. *J Am Coll Cardiol* 2011: 58(3): 300-309.
14. Viethen T, Gerhardt F, Dumitrescu D, Knoop-Busch S, ten Freyhaus H, Rudolph TK, Baldus S, Rosenkranz S. Ferric carboxymaltose improves exercise capacity and quality of life in patients with pulmonary arterial hypertension and iron deficiency: A pilot study. *International Journal of Cardiology* 2014: 175(2): 233-239.
15. Hennekens CH, Buring JE, Manson JE, Stampfer M, Rosner B, Cook NR, Belanger C, LaMotte F, Gaziano JM, Ridker PM, Willett W, Peto R. Lack of Effect of Long-Term Supplementation with Beta Carotene on the Incidence of Malignant Neoplasms and Cardiovascular Disease. *New England Journal of Medicine* 1996: 334(18): 1145-1149.
16. MRC/BHF Heart Protection Study of antioxidant vitamin supplementation in 20 536 high-risk individuals: a randomised placebo-controlled trial. *The Lancet* 2002: 360(9326): 23-33.
17. Davey Smith G, Hemani G. Mendelian randomization: genetic anchors for causal inference in epidemiological studies. *Hum Mol Genet* 2014: 23(R1): R89-98.
18. Burgess S, Timpson NJ, Ebrahim S, Davey Smith G. Mendelian randomization: where are we now and where are we going? *International Journal of Epidemiology* 2015: 44(2): 379-388.
19. Pichler I, Del Greco M F, Gögele M, Lill CM, Bertram L, Do CB, Eriksson N, Foroud T, Myers RH, Consortium PG, Nalls M, Keller MF, International Parkinson's Disease Genomics C, Wellcome Trust Case Control C, Benyamin B, Whitfield JB, Genetics of Iron Status C, Pramstaller PP, Hicks AA, Thompson JR, Minelli C. Serum Iron Levels and the Risk of Parkinson Disease: A Mendelian Randomization Study. *PLOS Medicine* 2013: 10(6): e1001462.
20. Collaboration CRPCHDG, Wensley F, Gao P, Burgess S, Kaptoge S, Di Angelantonio E, Shah T, Engert JC, Clarke R, Davey-Smith G, Nordestgaard BG, Saleheen D, Samani NJ, Sandhu M, Anand S, Pepys MB, Smeeth L, Whittaker J, Casas JP, Thompson SG, Hingorani AD, Danesh J. Association between C reactive protein and coronary heart disease: mendelian randomisation analysis based on individual participant data. *BMJ (Clinical research ed)* 2011: 342: d548-d548.
21. Interleukin-6 Receptor Mendelian Randomisation Analysis C, Swerdlow DI, Holmes MV, Kuchenbaecker KB, Engmann JEL, Shah T, Sofat R, Guo Y, Chung C, Peasey A, Pfister R, Mooijaart SP, Ireland HA, Leusink M, Langenberg C, Li KW, Palmen J, Howard P, Cooper JA, Drenos F, Hardy J, Nalls MA, Li YR, Lowe G, Stewart M, Bielinski SJ, Peto J, Timpson NJ, Gallacher J, Dunlop M, Houlston R, Tomlinson I, Tzoulaki I, Luan Ja, Boer JMA, Forouhi NG, Onland-Moret NC, van der Schouw YT, Schnabel RB, Hubacek JA, Kubinova R, Baceviciene M, Tamosiunas A, Pajak A, Topor-Madry R, Malyutina S, Baldassarre D, Sennblad B, Tremoli E, de Faire U, Ferrucci L, Bandenelli S, Tanaka T, Meschia JF, Singleton A, Navis G, Mateo Leach I, Bakker SJL, Gansevoort RT, Ford I, Epstein SE, Burnett MS, Devaney JM, Jukema JW, Westendorp RGJ, Jan de Borst G, van der Graaf Y, de Jong PA, Mailand-van der Zee A-H, Klungel OH, de Boer A, Doevendans PA, Stephens JW, Eaton CB, Robinson JG, Manson JE, Fowkes FG, Frayling TM, Price JF, Whincup PH, Morris RW, Lawlor DA, Smith GD, Ben-Shlomo Y, Redline S, Lange LA, Kumari M, Wareham NJ, Verschuren WMM, Benjamin EJ, Whittaker JC, Hamsten A, Dudbridge F, Delaney JAC, Wong A, Kuh D, Hardy R, Castillo BA, Connolly JJ, van der Harst P, Brunner EJ, Marmot MG, Wassel CL, Humphries SE, Talmud PJ, Kivimaki M, Asselbergs FW, Voevoda M, Bobak M, Pikhart H, Wilson JG, Hakonarson H, Reiner AP, Keating BJ, Sattar N, Hingorani AD, Casas JP. The interleukin-6 receptor as a target for prevention of coronary heart disease: a mendelian randomisation analysis. *Lancet (London, England)* 2012: 379(9822): 1214-1224.

22. Lewis SJ, Araya R, Smith GD, Freathy R, Gunnell D, Palmer T, Munafò M. Smoking is associated with, but does not cause, depressed mood in pregnancy--a mendelian randomization study. *PloS one* 2011: 6(7): e21689-e21689.
23. Palmer TM, Nordestgaard BG, Benn M, Tybjærg-Hansen A, Davey Smith G, Lawlor DA, Timpson NJ. Association of plasma uric acid with ischaemic heart disease and blood pressure: mendelian randomisation analysis of two large cohorts. *BMJ (Clinical research ed)* 2013: 347: f4262-f4262.
24. King EA, Davis JW, Degner JF. Are drug targets with genetic support twice as likely to be approved? Revised estimates of the impact of genetic support for drug mechanisms on the probability of drug approval. *BioRxiv* 2019.
25. Astle WJ, Elding H, Jiang T, Allen D, Ruklisa D, Mann AL, Mead D, Bouman H, Riveros-Mckay F, Kostadima MA, Lambourne JJ, Sivapalaratnam S, Downes K, Kundu K, Bomba L, Berentsen K, Bradley JR, Daugherty LC, Delaneau O, Freson K, Garner SF, Grassi L, Guerrero J, Haimel M, Janssen-Megens EM, Kaan A, Kamat M, Kim B, Mandoli A, Marchini J, Martens JH, Meacham S, Megy K, O'Connell J, Petersen R, Sharifi N, Sheard SM, Staley JR, Tuna S, van der Ent M, Walter K, Wang SY, Wheeler E, Wilder SP, Iotchkova V, Moore C, Sambrook J, Stunnenberg HG, Di Angelantonio E, Kaptoge S, Kuijpers TW, Carrillo-de-Santa-Pau E, Juan D, Rico D, Valencia A, Chen L, Ge B, Vasquez L, Kwan T, Garrido-Martin D, Watt S, Yang Y, Guigo R, Beck S, Paul DS, Pastinen T, Bujold D, Bourque G, Frontini M, Danesh J, Roberts DJ, Ouwehand WH, Butterworth AS, Soranzo N. The Allelic Landscape of Human Blood Cell Trait Variation and Links to Common Complex Disease. *Cell* 2016: 167(5): 1415-1429 e1419.
26. Galie N, Humbert M, Vachiery JL, Gibbs S, Lang I, Torbicki A, Simonneau G, Peacock A, Vonk Noordegraaf A, Beghetti M, Ghofrani A, Gomez Sanchez MA, Hansmann G, Klepetko W, Lancellotti P, Matucci M, McDonagh T, Pierard LA, Trindade PT, Zompatori M, Hoeper M, Group ESCSD. 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(1): 67-119.
27. National Cohort Study of Idiopathic and Heritable PAH. [cited; Available from: http://www.ipahcohort.com
28. Roden DM, Pulley JM, Basford MA, Bernard GR, Clayton EW, Balser JR, Masys DR. Development of a large-scale de-identified DNA biobank to enable personalized medicine. *Clinical pharmacology and therapeutics* 2008: 84(3): 362-369.
29. Vanderbilt University Medical Center: Institute for Clinical and Translational Research. [cited 2019 16/01]; Available from: https://victr.vanderbilt.edu/pub/biovu/
30. Graf S, Haimel M, Bleda M, Hadinnapola C, Southgate L, Li W, Hodgson J, Liu B, Salmon RM, Southwood M, Machado RD, Martin JM, Treacy CM, Yates K, Daugherty LC, Shamardina O, Whitehorn D, Holden S, Aldred M, Bogaard HJ, Church C, Coghlan G, Condliffe R, Corris PA, Danesino C, Eyries M, Gall H, Ghio S, Ghofrani HA, Gibbs JSR, Girerd B, Houweling AC, Howard L, Humbert M, Kiely DG, Kovacs G, MacKenzie Ross RV, Moledina S, Montani D, Newnham M, Olschewski A, Olschewski H, Peacock AJ, Pepke-Zaba J, Prokopenko I, Rhodes CJ, Scelsi L, Seeger W, Soubrier F, Stein DF, Suntharalingam J, Swietlik EM, Toshner MR, van Heel DA, Vonk Noordegraaf A, Waisfisz Q, Wharton J, Wort SJ, Ouwehand WH, Soranzo N, Lawrie A, Upton PD, Wilkins MR, Trembath RC, Morrell NW. Identification of rare sequence variation underlying heritable pulmonary arterial hypertension. *Nat Commun* 2018: 9(1): 1416.
31. Lawlor DA, Harbord RM, Sterne JA, Timpson N, Davey Smith G. Mendelian randomization: using genes as instruments for making causal inferences in epidemiology. *Stat Med* 2008: 27(8): 1133-1163.

32. Bowden J, Davey Smith G, Haycock PC, Burgess S. Consistent Estimation in Mendelian Randomization with Some Invalid Instruments Using a Weighted Median Estimator. *Genetic epidemiology* 2016: 40(4): 304-314.
33. Benyamin B, Esko T, Ried JS, Radhakrishnan A, Vermeulen SH, Traglia M, Gogele M, Anderson D, Broer L, Podmore C, Luan J, Kutalik Z, Sanna S, van der Meer P, Tanaka T, Wang F, Westra HJ, Franke L, Mihailov E, Milani L, Halldin J, Winkelmann J, Meitinger T, Thiery J, Peters A, Waldenberger M, Rendon A, Jolley J, Sambrook J, Kiemeney LA, Sweep FC, Sala CF, Schwienbacher C, Pichler I, Hui J, Demirkan A, Isaacs A, Amin N, Steri M, Waeber G, Verweij N, Powell JE, Nyholt DR, Heath AC, Madden PA, Visscher PM, Wright MJ, Montgomery GW, Martin NG, Hernandez D, Bandinelli S, van der Harst P, Uda M, Vollenweider P, Scott RA, Langenberg C, Wareham NJ, InterAct C, van Duijn C, Beilby J, Pramstaller PP, Hicks AA, Ouwehand WH, Oexle K, Gieger C, Metspalu A, Camaschella C, Toniolo D, Swinkels DW, Whitfield JB. Novel loci affecting iron homeostasis and their effects in individuals at risk for hemochromatosis. *Nat Commun* 2014: 5: 4926.
34. Burgess S. Sample size and power calculations in Mendelian randomization with a single instrumental variable and a binary outcome. *International Journal of Epidemiology* 2014: 43(3): 922-929.
35. Balanos G, Dorrington K, Robbins P. Desferrioxamine elevates pulmonary vascular resistance in humans: potential for involvement of HIF-1. *J Appl Physiol* 2002(92): 2501–2507.
36. Smith TG, Balanos GM, Croft QPP, Talbot NP, Dorrington KL, Ratcliffe PJ, Robbins PA. The increase in pulmonary arterial pressure caused by hypoxia depends on iron status. *The Journal of Physiology* 2008: 586(24): 5999-6005.
37. Cotroneo E, Ashek A, Wang L, Wharton J, Dubois O, Bozorgi S, Busbridge M, Alavian KN, Wilkins MR, Zhao L. Iron homeostasis and pulmonary hypertension: iron deficiency leads to pulmonary vascular remodeling in the rat. *Circ Res* 2015: 116(10): 1680-1690.
38. Ivan M, Kondo K, Yang H. HIFa targeted for VHL-mediated destruction by proline hydroxylation: implications for O2 sensing. *Science* 2001(292): 464–468.
39. Charles-Edwards G, Amaral N, Sleigh A, Ayis S, Catibog N, McDonagh T, Monaghan M, Amin-Youssef G, Kemp GJ, Shah AM, Okonko DO. Effect of Iron Isomaltoside on Skeletal Muscle Energetics in Patients With Chronic Heart Failure and Iron Deficiency. *Circulation* 2019: 139(21): 2386-2398.
40. Gill D, Del Greco MF, Walker AP, Srai SKS, Laffan MA, Minelli C. The Effect of Iron Status on Risk of Coronary Artery Disease: A Mendelian Randomization Study-Brief Report. *Arterioscler Thromb Vasc Biol* 2017: 37(9): 1788-1792.
41. **Gill D**, **Monori G**, **Tzoulaki I**, **Dehgan A**. Iron Status and Risk of Stroke- A Mendelian Randomization Study. *Stroke* 2018.
42. Hassoun PM. Inflammation in pulmonary arterial hypertension: is it time to quell the fire? *European Respiratory Journal* 2014: 43(3): 685.
43. Machado RD, Pauciulo MW, Thomson JR, Lane KB, Morgan NV, Wheeler L, Phillips Iii JA, Newman J, Williams D, Galiè N, Manes A, McNeil K, Yacoub M, Mikhail G, Rogers P, Corris P, Humbert M, Donnai D, Martensson G, Tranebjaerg L, Loyd JE, Trembath RC, Nichols WC. BMPR2 Haploinsufficiency as the Inherited Molecular Mechanism for Primary Pulmonary Hypertension. *The American Journal of Human Genetics* 2001: 68(1): 92-102.

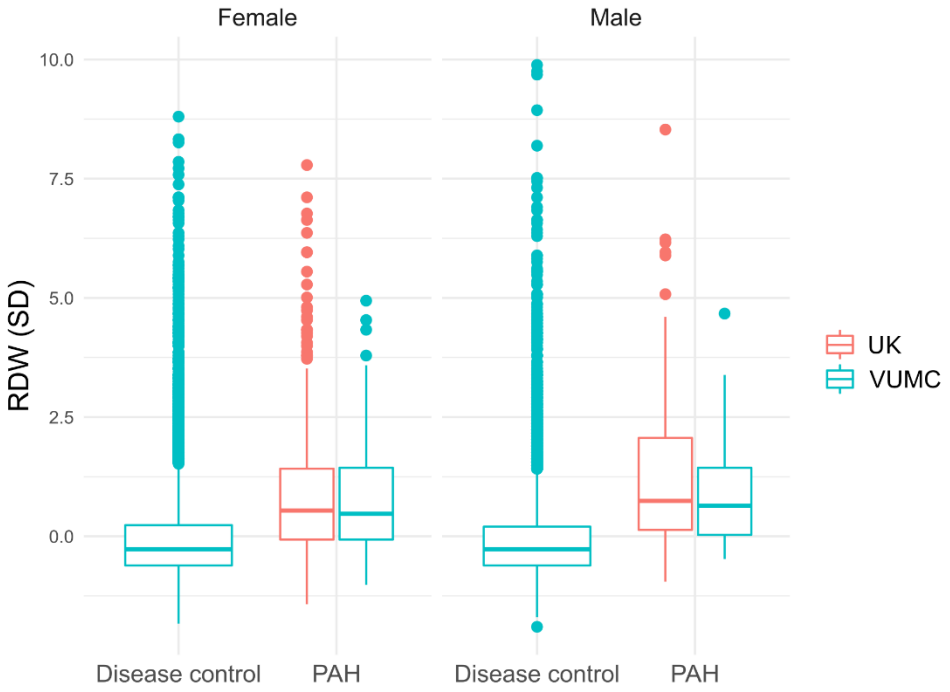


Figure 1 Boxplot of RDW levels in the merged cohort of PAH cases (N=642) and the disease control cohort (N=15,889). The bottom and the top lines of the box indicate the 25th and 75th percentiles, while the centerline indicates the median value. The whiskers extend to 1.5 times the interquartile range from both ends of the box with individual points being more extreme observations. SD = standard deviation.

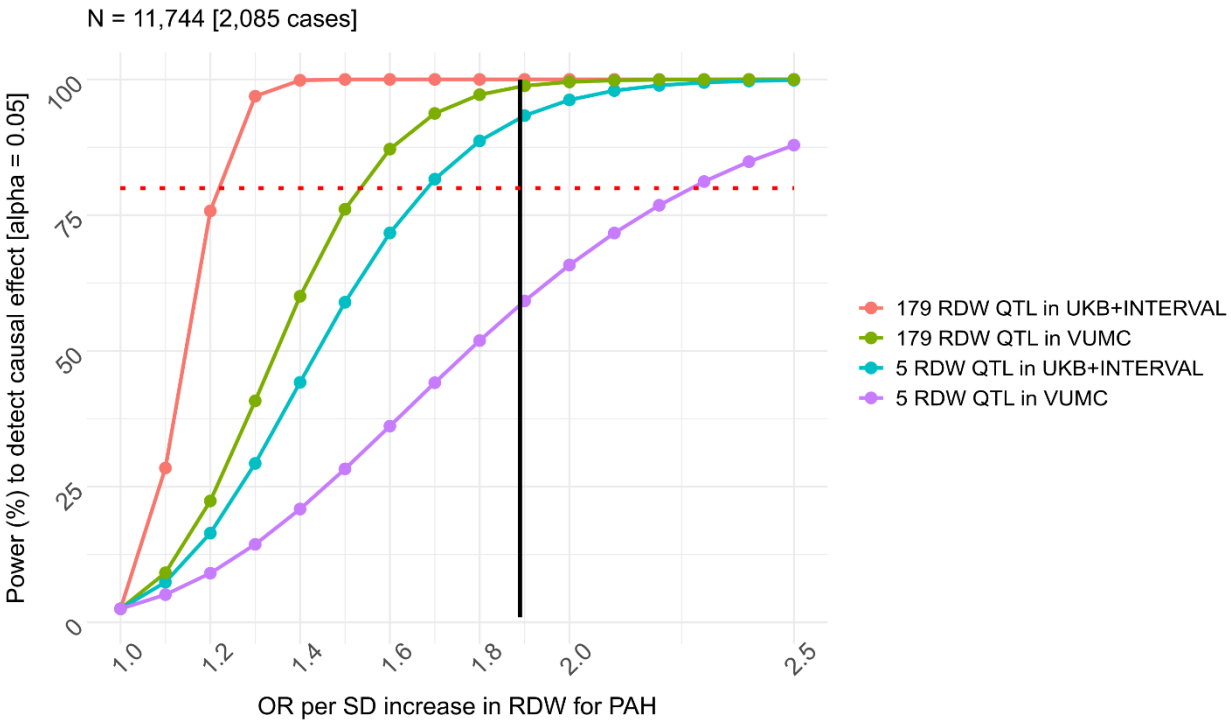


Figure 2 Power (%) to detect a causal association (y-axis) given the size of the true underlying causal effect of one standard unit increase in RDW on PAH risk (x-axis). The effect estimate obtained from the observational study is indicated with the vertical black line at OR = 1.90 whilst the red dotted line marks the desired power of 80%. Red curve: MR using all overlapping genome-wide significant variants from the RDW GWAS, given the true R^2 = 12% as per estimated in the UKB and INTERVAL cohorts. Green curve: MR using all genome-wide significant QTL from the RDW GWAS, given the true R^2 = 2.6% as per estimated in our VUMC control cohort. Blue line: MR using five genome-wide significant variants from the RDW GWAS reflecting systemic iron status, given the true R^2 = 1.7% as per estimated in the UKB and INTERVAL cohorts. Purple line: MR using five genome-wide significant QTL from the RDW GWAS reflecting systemic iron status, given the true R^2 = 0.7% as per estimated in our VUMC control cohort.

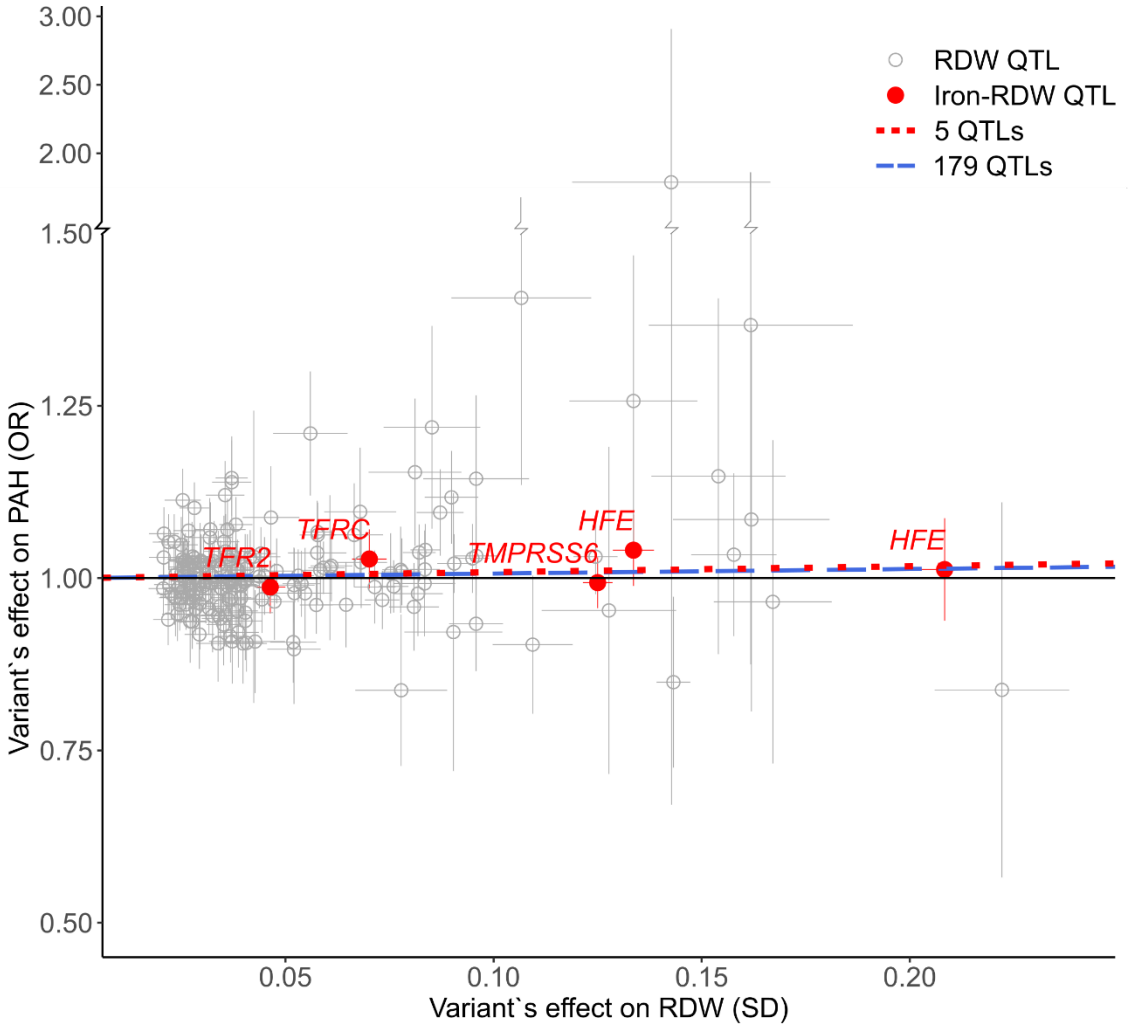


Figure 3 Scatterplot of variant – RDW associations (x-axis) plotted against variant – PAH associations (y-axis) where each dot represents a single RDW QTL. The effect estimates and their standard errors (grey bars) are given in standard units for RDW and in odds ratios for PAH. The solid black line denotes an OR of 1 (no effect) whilst the dashed blue line is the overall causal effect from the IVW regression using all 179 RDW QTL. The five iron-specific RDW QTL (red dots) used as instruments in the secondary MR analysis are labelled with their corresponding gene names and the red dotted line denotes the corresponding causal effect.

Supplementary Appendix
for

Mendelian randomization analysis of red cell distribution width in pulmonary arterial hypertension

Anna Ulrich *et al.*

Table of Contents

Supplementary Methods

RDW and PAH association

RDW values were natural log-transformed and for ease of interpretation z-score normalized to have a mean of zero and a standard deviation of one. The association between RDW levels and PAH was tested in a logistic regression framework adjusted for age and sex (Supplementary Table 1). For PAH cases we used closest to diagnosis RDW measurements. For controls we retrieved the first available RDW measurement.

We excluded individuals from non-white ethnic backgrounds to avoid potential bias from ethnicity effects. Since RDW is known to be elevated in a number of diseases, we excluded all individuals with either of the following in their medical history: polycystic kidney disease, chronic kidney disease, liver disease and transfusion therapy received. Furthermore, children and adolescents under the age of 18 as well as individuals with extreme RDW values (below 10%, N=20 or above 30%, N=15) were not included in the analysis. Out of the 35 excluded for extreme RDW, 14 were cases from VUMC in the range of 43-55 and 20 were all below 10 from one of the NIHRBR centers. Both centers confirmed reporting errors for these samples (*Supplementary Figure 1*).

Mendelian randomization; causal effect estimation using MR

Alleles were aligned to correspond to an increase in RDW followed by the harmonization of the effects. The causal effect was estimated with the inverse variance weighted (IVW) and weighted median estimator (WM) methods as implemented in the MR-Base software (1).

Inverse-variance weighted method

We used the conventional inverse variance weighted (IVW) method for estimating the causal effect. The IVW method is efficient when all variants in the genetic instrument are valid instruments. Briefly, each variant in the genetic instrument provided a causal estimate calculated by simply dividing the variant's effect on PAH by the variant's effect on RDW (ratio of coefficients or Wald ratio). These individual causal estimates were then meta-analyzed in a fixed-effects model weighted by the reciprocal of the standard error of the variant association with PAH.

This is equivalent to regressing the variant-RDW estimates on the variant-PAH estimates with the above-mentioned weighting whilst forcing the regression line to pass through the origin.

The causal estimate from the IVW method (β_{IVW}) is:

$$\beta_{IVW} = \frac{\sum_{k=1}^{K} X_k Y_k \sigma_{Yk}^{-2}}{\sum_{k=1}^{K} X_k^2 \sigma_{Yk}^{-2}}$$

Where:
k is an index for each of the variants used in the two-sample MR analysis
X is the effect estimate on the exposure (RDW) as reported in the RDW GWAS summary statistics
Y is the effect estimate on the outcome (PAH) as reported in the PAH GWAS summary statistics
The standard error of the causal estimate is:

$$se(\beta_{IVW}) = \sqrt{\frac{1}{\sum_{k=1}^{K} X_k^2 \sigma_{Yk}^{-2}}}$$

Weighted median estimator

We used the WM estimator to allow for up to (but not including 50%) of the variants in our genetic instrument to be invalid instruments.

Analogously to the IVW method, a Wald ratio (see above) is calculated for each variant. These Wald ratios are then ordered and weighted by the same weights used in the IVW method (see above). Let wj be the weight of the jth ordered Wald ratio estimate.

$$s_j = \sum_{k=1}^{j} w_k$$

Where:

k is an index for each of the variants used in the two-sample MR analysis

w is the weight of the variant

s_j is the sum of weights up to and including the jth Wald ratio estimate

The weights are standardized, so that the sum of weights is 1. The WM estimator is the median of the empirical distribution of weighted Wald ratios. Each Wald ratio is the $100(s_j - \frac{w_j}{2})$th percentile of this distribution.

Quality control and imputation procedures of VUMC genotype data

VUMC participants were genotyped in 6 batches (30,886 in total) using the Infinium Expanded Multi-Ethnic Global Array-8 (MEGA-ex) array (Illumina, San Diego, California, US).

Variant QC (pre-imputation): Variants were excluded if they had a low call rate (< 95%), deviated from the Hardy-Weinberg equilibrium ($P \leq 0.00005$), were rare (minor allele frequency ≤ 1%) or had more than two alleles.

Sample QC: Individuals with high proportions (> 5%) of missing genotype data, unresolved sex discrepancies (discordant phenotype-genotype sex information), heterozygosity outliers, self-reported and/or principal component-based ethnic outliers, intentional duplicates and related individuals (PI-HAT > 0.2) were excluded.

Imputation of non-genotyped variants: The filtered genotype array data was imputed to the Haplotype Reference Consortium panel using the free Sanger Imputation Service provided by the Wellcome Sanger Institute (2).

Imputed variants were further filtered for deviations from the Hardy-Weinberg equilibrium ($P \leq 0.00005$), rare variants (minor allele frequency ≤ 1%) which are often poorly imputed and other low-quality variants with an INFO score lower than 0.9.

Genetic risk score (GRS) derivation and calculation of variance explained (R^2) from individual-level data

Weighted genetic risk scores (GRS) comprising the single nucleotide polymorphisms (SNPs) from the RDW genetic instrument were regressed onto the first RDW values which provided the coefficient of determination (R^2) as an estimate for the correlation between RDW and the RDW GRS in our population. The GRSs can be derived by summing the effect alleles multiplied by the effect size at each of the variants (3).

GRS were calculated using the software PRSice-2 (3). Genotypes for the 179 SNPs in the RDW GRS were extracted from the imputed VUMC controls dataset. The same inclusion criteria as for our observational study were applied. Out of the 15,889 VUMC controls included in the observational study, 14,964 had genetic data that passed standard variant and sample quality control.

The GRS for an individual is the summation of the effect (trait-increasing) alleles (0, 1 or 2) weighted by the effect size of the variant taken from the genome-wide significant summary statistics of the RDW GWAS (4). We used an additive model meaning that homozygotes for the effect allele had twice the increase in RDW levels as the heterozygotes. This was in line with the model used in the RDW GWAS.

The following models were used to assess the validity of the RDW GRS as a proxy for RDW levels:

Full model: *RDW ~ GRS + sex + age + principal components (1st and 2nd) + batch (study specific)*

Null model: *sex + age + principal components (1st and 2nd) + batch (study specific)*

The R^2 for the GRS alone is calculated by subtracting the R^2 of the model not containing the GRS (null model) from the R^2 of the full model.

The confidence intervals for the GRS R^2 were computed using the adjusted bootstrap percentile method as implemented in the R software package 'boot' (5) (number of replicates = 20,000).

R^2 calculation from GWAS summary-level data

R^2 was calculated for each independent variant based from the publicly available summary statistics of the RDW GWAS in the discovery and replication populations (4). These individual estimates were then summed to give the overall variance explained by the RDW instrument.

$$R^2 = \sum_{k=1}^{K} \frac{Nsample_k + 1}{Nsample_k} \times \frac{Z_k^2}{Z_k^2 + Nsample_k} - \frac{1}{Nsample_k}$$

Where:

k is an index for each of the variants used in the two-sample MR analysis

Z is the Z-statistic as reported in the RDW GWAS summary statistics

$Nsample$ is the sample size as reported in the RDW GWAS summary statistics

The standard error for the R2 estimate was calculated as shown below:

$$SE_{R^2} = \sum_{k=1}^{K} \sqrt{\left(\frac{2}{Nsample_k}\right) \times \left(2 \times R_k^2 + \frac{1}{Nsample_k}\right)}$$

Supplementary Table and Figure Legends

Supplementary Table 1 Characteristics of the study population used for estimating/assessing the association between RDW and PAH, stratified by sex. PAH – Pulmonary Arterial Hypertension, VUMC - Vanderbilt Institute for Clinical and Translational Research, NIHRBR - UK National Institute for Health Research BioResource, RDW – red cell distribution width.

Supplementary Table 2 Logistic regression model predicting PAH disease status. We report the results of the adjusted model in the paper.

Supplementary Table 3 Five variants selected from the RDW GWAS based on their effects on systemic iron status. This table presents the effect estimates of these variants on RDW as reported by Astle et al. on RDW (Effect estimate per RDW SD; Effect estimate p-value). *The effects of these variants for the same allele go in the opposite direction on serum iron as reported by the Genetics of Iron Status GWAS (6). Elevated RDW can reflect iron deficiency which presents with decreased serum iron levels. In the Genetics of Iron Status GWAS, the two HFE variants reached genome-wide significance ($P< 5\times10^{-8}$) for all four (serum iron, transferrin, transferrin saturation, ferritin) iron status biomarkers, TMPRSS6 reached it for all but transferrin, TFRC reached it for transferrin and transferrin saturation, while TFR2 reached it for iron and transferrin saturation.

Supplementary Table 4 Summary of the RDW GRS models. The null model corresponds to the linear regression model specified above without the GRS. The R2 of the null models are identical since the sample and the covariates are the same. P.value is the significance value of the model fit (F-test). Empirical p-values that account for multiple testing and overfitting were obtained through permutation tests (n=20,000).

Supplementary Figure 1 Flow diagram of the exclusion steps in the UK centers and the Vanderbilt University Medical Centre (VUMC) of the two cohorts (UK PAH Cohort and VUMC) participating in the RDW and PAH association analysis as described in the Supplementary methods.

Supplementary Figure 2 Mendelian Randomization (MR) analyses of red cell distribution width (RDW) and pulmonary arterial hypertension (PAH). A two-sample design was used where effect estimates for the instrumental genetic variants were taken from two non-overlapping populations. RDW QTL and their effect estimates were taken from the largest-to-date population based RDW genetic association study (GWAS) (4). Effect estimates for the RDW QTL on PAH susceptibility were obtained from the largest-to-date PAH GWAS (6). The primary MR analysis included all 179 RDW QTL while the secondary MR analysis was restricted to five out of the 179 RDW QTL acting via iron status (Supplementary Table 3).

Supplementary Figure 3 Individual MR causal estimates (IVW) for the main MR analysis of association between RDW and development of PAH - using all available RDW SNPs - from the four contributing studies in PAH GWAS. PAH ORs per one standard unit increase in RDW (dot) with the corresponding lower and upper 95% confidence intervals (horizontal line). The result of the BHFPAH (IVW causal OR = 1.54, 95% CI = 1.06 – 2.23) did not survive the correction for multiple testing and was driven by the causal estimate of one variant (rs6883412). PHAAR: Pulmonary Hypertension Allele-Associated Risk (269 PAH cases, 1068 controls). PAHB: US National Biological Sample and Data Repository for Pulmonary Arterial Hypertension (694 PAH cases, 1560 controls). NIHRBR: UK National Institute for Health Research BioResource (847 PAH cases, 5048 controls). BHFPAH: British Heart Foundation Pulmonary Arterial Hypertension (275 PAH cases, 1983 controls). Meta-analyzed: overall results of PAH GWAS including all four studies.

Supplementary Figure 4 Individual MR causal estimates (IVW) for the secondary MR analysis of association of RDW to development of PAH – using 5 SNPs related to systemic iron status - from the four contributing studies in PAH GWAS. PAH ORs per one standard unit increase in RDW (dot) with the corresponding lower and upper 95% Cis (horizontal line). PHAAR: Pulmonary Hypertension Allele-Associated Risk (269 PAH cases, 1068 controls). PAHB: US National Biological ample and Data Repository for Pulmonary Arterial Hypertension (694 PAH cases, 1560 controls). NIHRBR: UK National Institute for Health Research BioResource (847 PAH cases, 5048 controls). BHFPAH: British Heart Foundation Pulmonary Arterial Hypertension (275 PAH cases, 1983 controls). Meta-analyzed: overall results of PAH GWAS including all four studies.

Supplementary Tables and Figures

	PAH (118 VUMC; 524 NIHRBR)		Controls (15,889 VUMC)	
	Female (%)	Male (%)	Female (%)	Male (%)
N	445 (69)	197 (31)	8,539 (54)	7,350 (46)
RDW (mean/SD)	15.1/2.20	15.5/2.33	13.6/1.41	13.6/1.39
Age (mean/SD)	52.4/17.6	58.2/16.2	54.3/16.1	58.1/14.7

Supplementary Table 2 Characteristics of the study population used for estimating/assessing the association between RDW and PAH, stratified by sex. PAH – Pulmonary Arterial Hypertension, VUMC - Vanderbilt University Medical Centre, NIHRBR - UK National Institute for Health Research BioResource, RDW – red cell distribution width.

Variable	Unadjusted		Adjusted	
	OR	95% CI	OR	95% CI
RDW (SD)	1.85	1.75 – 1.94	1.90	1.80 – 2.01
Age	-	-	0.98	0.98 – 0.99
Sex *(base=female)*	-	-	0.54	0.45 – 0.64

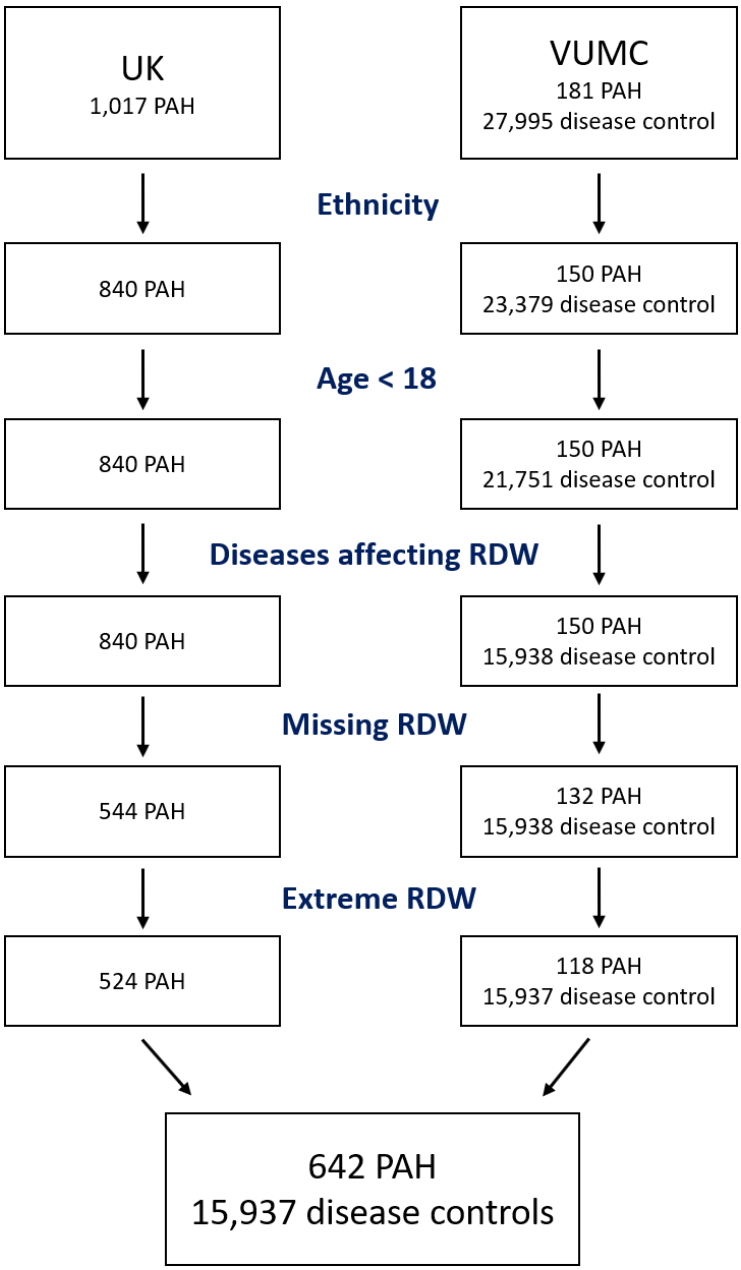
Supplementary Table 2 Logistic regression model predicting PAH disease status. We report the results of the adjusted model in the paper.

Variant information				RDW GWAS Astle et al.		Genetics of Iron Status GWAS Benyamin et al.								
Gene	Lead variant ID	Effect Allele	Allele Frequency in UKB and INTERVAL	beta RDW	*P* - value	Proxy variant ID*	beta iron	*P* – value iron	beta ferritin	*P* – value ferritin	beta TSAT	*P* – value TSAT	beta TF	P – value TF
HFE	rs144861591	C	0.92	0.21	6.50×10^{-216}	rs1800562	-0.37	4.0×10^{-77}	-0.21	1.4×10^{-29}	-0.58	1.5×10^{-178}	0.55	1.3×10^{-153}
TMPRSS6	rs855791	A	0.44	0.13	9.79×10^{-271}	-	-0.19	4.3×10^{-77}	-0.05	5.8×10^{-8}	-0.19	3.5×10^{-80}	0.04	1.3×10^{-4}
HFE	rs198851	G	0.85	0.13	2.55×10^{-161}	-	-0.19	1.6×10^{-40}	-0.06	3.6×10^{-6}	-0.23	4.7×10^{-59}	0.12	3.0×10^{-17}
TFRC	rs7619708	C	0.24	0.07	4.35×10^{-64}	rs6583288	0.03	1.2×10^{-2}	0.004	7.3×10^{-1}	0.05	3.8×10^{-6}	-0.06	3.8×10^{-8}
TFR2	rs9801017	G	0.37	0.05	9.40×10^{-37}	rs7385804	-0.06	7.2×10^{-8}	-0.02	2.5×10^{-2}	-0.05	1.8×10^{-7}	0.01	4.0×10^{-1}

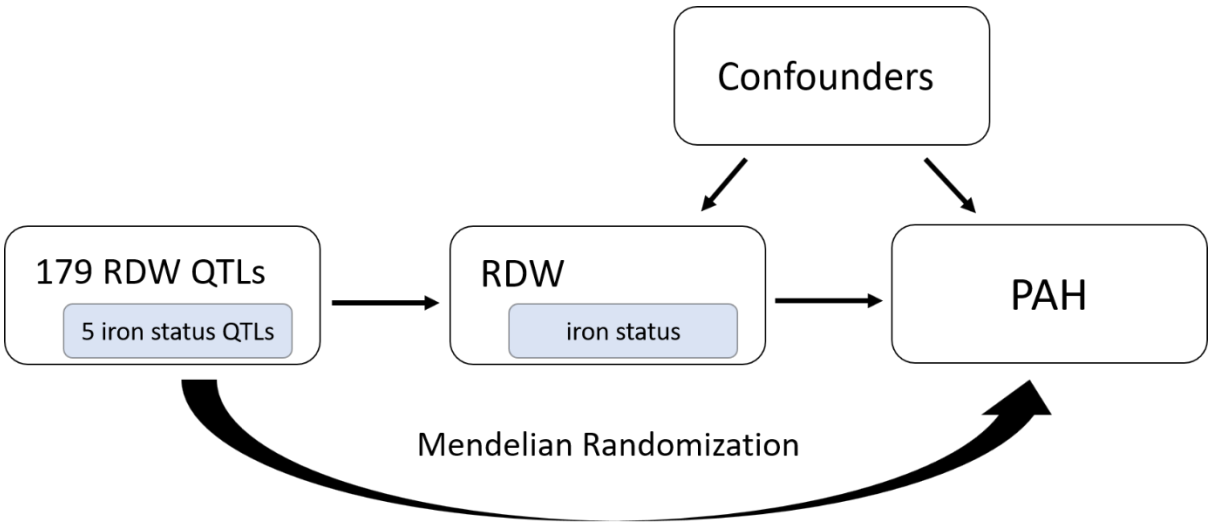
Supplementary Table 3 Five variants selected from the RDW GWAS based on their effects on systemic iron status. This table presents the effect estimates of these variants on RDW as reported by Astle et al. on RDW (Effect estimate per RDW SD; Effect estimate p-value). Elevated RDW can reflect iron deficiency which presents with decreased serum iron levels. In the Genetics of Iron Status GWAS, the two HFE variants reached genome-wide significance ($P< 5\times10^{-8}$) for all four (serum iron, transferrin, transferrin saturation, ferritin) iron status biomarkers, TMPRSS6 reached it for all but transferrin, TFRC reached it for transferrin and transferrin saturation, while TFR2 reached it for iron and transferrin saturation. The betas for RDW and the iron biomarkers from Benyamin et al. are reported in standard units. RDW = red cell distribution width; TSAT = transferrin saturation; TF = transferrin. *Where the lead RDW variant for the locus was not available in the Genetics of Iron Status GWAS we listed the results of a suitable proxy variant in strong linkage disequilibrium ($r^2 \geq 0.8$) with the RDW lead variant.

RDW GRS	GRS R^2	Null R^2	P value	Empirical P value
179 RDW QTLs	0.0264	0.054	2.4×10^{-94}	5.0×10^{-5}
5 RDW QTLs	0.0065	0.054	3.3×10^{-24}	5.0×10^{-5}

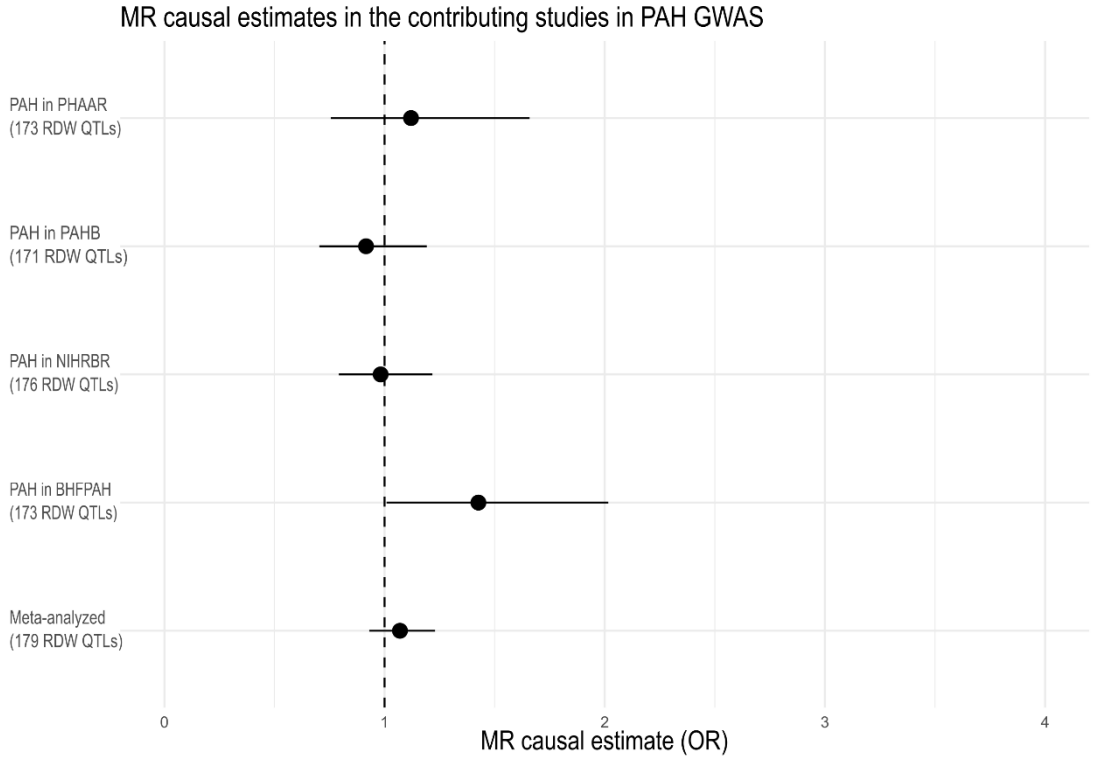
Supplementary Table 4 Summary of the RDW GRS models. The null model corresponds to the linear regression model without the GRS. The R^2 of the null models are identical since the sample and the covariates are the same. P value is the significance value of the model fit (F-test). Empirical p-values that account for multiple testing and overfitting were obtained through permutation tests (n=20,000).



Supplementary Figure 1 Flow diagram of the exclusion steps in the UK centers and the Vanderbilt University Medical Centre (VUMC) participating in the RDW and PAH association analysis as described in the Supplementary methods.

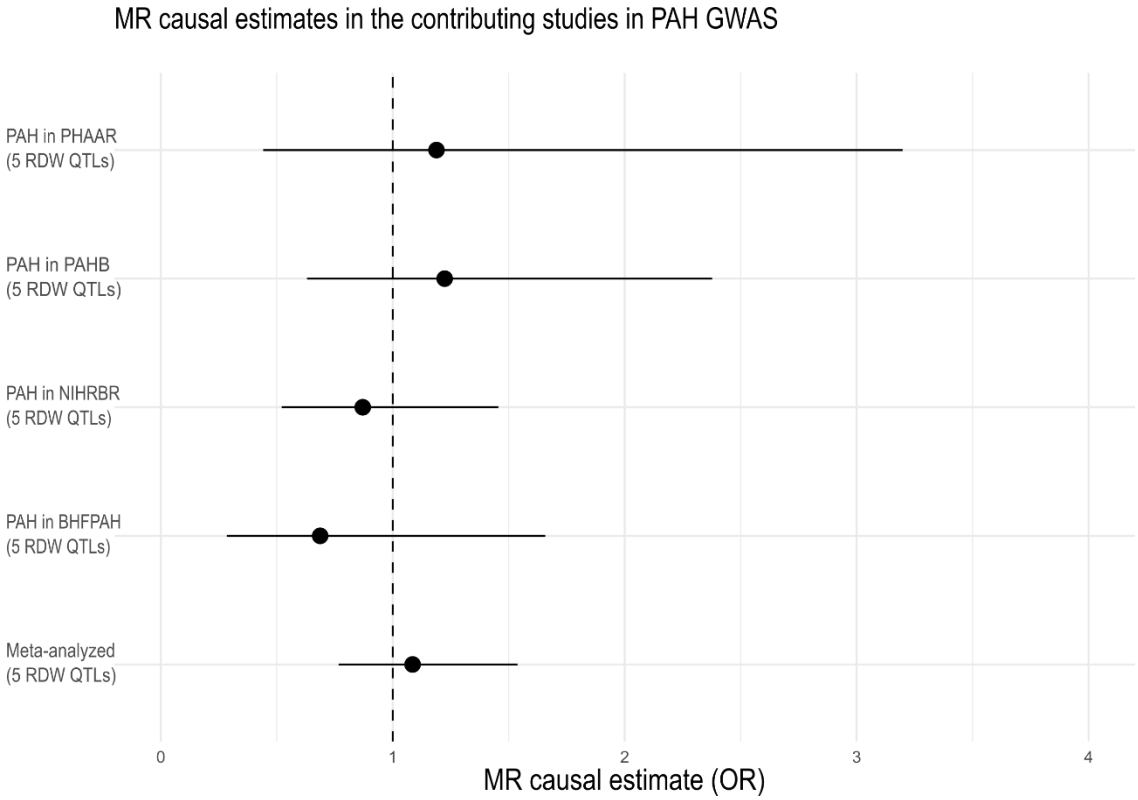


Supplementary Figure 2 Mendelian Randomization (MR) analyses of red cell distribution width (RDW) and pulmonary arterial hypertension (PAH). A two-sample design was used where effect estimates for the instrumental genetic variants were taken from two non-overlapping populations. RDW QTL and their effect estimates were taken from the largest-to-date population based RDW genetic association study (GWAS) (4). Effect estimates for the RDW QTL on PAH susceptibility were obtained from the largest-to-date PAH GWAS (7). The primary MR analysis included all 179 RDW QTL while the secondary MR analysis was restricted to five out of the 179 RDW QTL acting via iron status (Supplementary Table 3).



Supplementary Figure 3 Individual inverse variance weighted MR causal estimates for the main MR analysis of association between RDW and development of PAH - using all available RDW QTLs including suitable proxy variants with a minimum r^2 of 0.8 in each study - from the four contributing

studies in PAH GWAS. PAH ORs per one standard unit increase in RDW (dot) with the corresponding lower and upper 95% confidence intervals (horizontal line). The result of the BHFPAH (OR causal = 1.43, 95% CI = 1.01 – 2.02) did not survive the correction for multiple testing and was driven by the causal estimate of one variant (rs6883412). PHAAR: Pulmonary Hypertension Allele-Associated Risk (269 PAH cases, 1068 controls). PAHB: US National Biological Sample and Data Repository for Pulmonary Arterial Hypertension (694 PAH cases, 1560 controls). NIHRBR: UK National Institute for Health Research BioResource (847 PAH cases, 5048 controls). BHFPAH: British Heart Foundation Pulmonary Arterial Hypertension (275 PAH cases, 1983 controls). Meta-analyzed: overall results of PAH GWAS including all four studies.



Supplementary Figure 4 Individual MR causal estimates (IVW) for the secondary MR analysis of association of RDW to development of PAH – using 5 RDW QTLs related to systemic iron status - from the four contributing studies in PAH GWAS. PAH ORs per one standard unit increase in RDW (dot) with the corresponding lower and upper 95% confidence intervals (horizontal line). PHAAR: Pulmonary Hypertension Allele-Associated Risk (269 PAH cases, 1068 controls). PAHB: US National Biological Sample and Data Repository for Pulmonary Arterial Hypertension (694 PAH cases, 1560 controls). NIHRBR: UK National Institute for Health Research BioResource (847 PAH cases, 5048 controls). BHFPAH: British Heart Foundation Pulmonary Arterial Hypertension (275 PAH cases, 1983 controls). Meta-analyzed: overall results of PAH GWAS including all four studies.

References

1. Hemani G, Zheng J, Elsworth B, Wade KH, Haberland V, Baird D, et al. The MR-Base platform supports systematic causal inference across the human phenome. eLife. 2018;7:e34408.
2. the Haplotype Reference C, McCarthy S, Das S, Kretzschmar W, Delaneau O, Wood AR, et al. A reference panel of 64,976 haplotypes for genotype imputation. Nature Genetics. 2016;48:1279.
3. Euesden J, Lewis CM, O'Reilly PF. PRSice: Polygenic Risk Score software. Bioinformatics. 2015;31(9):1466-8.
4. Astle WJ, Elding H, Jiang T, Allen D, Ruklisa D, Mann AL, et al. The Allelic Landscape of Human Blood Cell Trait Variation and Links to Common Complex Disease. Cell. 2016;167(5):1415-29 e19.
5. Canty A, Ripley B. boot: Bootstrap R (S-Plus) Functions. R package version 1.3-23. ed2019.
6. Benyamin B, Esko T, Ried JS, Radhakrishnan A, Vermeulen SH, Traglia M, et al. Novel loci affecting iron homeostasis and their effects in individuals at risk for hemochromatosis. Nat Commun. 2014;5:4926.
7. Rhodes CJ, Batai K, Bleda M, Haimel M, Southgate L, Germain M, et al. Genetic determinants of risk in pulmonary arterial hypertension: international genome-wide association studies and meta-analysis. The Lancet Respiratory Medicine. 2018.

NIHR BioResource Collaborators

NIHR BioResource – Rare Disease Consortium[1], Julian Adlard[2], Munaza Ahmed[3], Tim Aitman[4,5], Hana Alachkar[6], David Allsup[7], Jeff Almeida-King[8], Philip Ancliff[9], Richard Antrobus[10], Ruth Armstrong[11,12,13], Gavin Arno[14,15], Sofie Ashford[1,16], William Astle[1,16,17], Anthony Attwood[1,16], Chris Babbs[18,19], Tamam Bakchoul[20], Tadbir Bariana[21,22], Julian Barwell[23,24], David Bennett[25], David Bentley[26], Agnieszka Bierzynska[27], Tina Biss[28], Marta Bleda[29], Harm Bogaard[30], Christian Bourne[26], Sara Boyce[31], John Bradley[1], Gerome Breen[32,33], Paul Brennan[34,35], Carole Brewer[36], Matthew Brown[1,16], Michael Browning[37], Rachel Buchan[38,39], Matthew Buckland[40], Teofila Bueser[41,42,43], Siobhan Burns[40], Oliver Burren[29], Paul Calleja[44], Gerald Carr-White[42], Keren Carss[1,16], Ruth Casey[11,12,13], Mark Caulfield[45], John Chambers[46,47], Jennifer Chambers[48,49], Floria Cheng[49], Patrick F Chinnery[1,50,51], Martin Christian[52], Colin Church[53], Naomi Clements Brod[1,16], Gerry Coghlan[40], Elizabeth Colby[27], Trevor Cole[54], Janine Collins[55], Peter Collins[56], Camilla Colombo[26], Robin Condliffe[57], Stuart Cook[38,58,59,60], Terry Cook[61], Nichola Cooper[62], Paul Corris[63,64], Abigail Crisp-Hihn[1,16], Nicola Curry[65], Cesare Danesino[66], Matthew Daniels[67,68], Louise Daugherty[1,16], John Davis[1,16], Sri V V Deevi[1,16], Timothy Dent[68], Eleanor Dewhurst[1,16], Peter Dixon[48], Kate Downes[1,16], Anna Drazyk[69], Elizabeth Drewe[70], Tina Dutt[71], David Edgar[72], Karen Edwards[1,16], William Egner[73], Wendy Erber[74], Marie Erwood[1,16], Maria C Estiu[75], Gillian Evans[76], Dafydd Gareth Evans[77], Tamara Everington[78], Mélanie Eyries[79], Remi Favier[80,81,82], Debra Fletcher[1,16], James Fox[1,16], Amy Frary[1,16], Courtney French[83], Kathleen Freson[84], Mattia Frontini[1,16], Daniel Gale[85], Henning Gall[86], Claire Geoghegan[26], Terry Gerighty[26], Stefano Ghio[87], Hossein-Ardeschir Ghofrani[62,86], Simon Gibbs[38], Kimberley Gilmour[88], Barbara Girerd[89,90,91], Sarah Goddard[92], Keith Gomez[21,22], Pavels Gordins[93], David Gosal[6], Stefan Gräf[1,16,29], Luigi Grassi[1,16], Daniel Greene[1,16,17], Lynn Greenhalgh[94], Andreas Greinacher[95], Paolo Gresele[96], Philip Griffiths[97,98], Sofia Grigoriadou[99], Russell Grocock[26], Detelina Grozeva[11], Scott Hackett[100], Charaka Hadinnapola[29], William Hague[101], Matthias Haimel[1,16,29], Matthew Hall[70], Helen Hanson[94], Kirsty Harkness[102], Andrew Harper[38,67,103], Claire Harris[64], Daniel Hart[55], Ahamad Hassan[104], Grant Hayman[105], Alex Henderson[106], Jonathan Hoffmann[54], Rita Horvath[107,108], Arjan Houweling[30], Luke Howard[38], Fengyuan Hu[1,16], Gavin Hudson[107], Joseph Hughes[26], Aarnoud Huissoon[100], Marc Humbert[89,90,91], Sean Humphray[26], Sarah Hunter[26], Matthew Hurles[109], Louise Izatt[110], Roger James[1,16], Sally Johnson[111], Stephen Jolles[112,113], Jennifer Jolley[1,16], Neringa Jurkute[14,22], Mary Kasanicki[114], Hanadi Kazkaz[115], Rashid Kazmi[31], Peter Kelleher[39], David Kiely[57], Nathalie Kingston[1], Robert Klima[44], Myrto Kostadima[1,16], Gabor Kovacs[116,117], Ania Koziell[118,119], Roman Kreuzhuber[1,16], Taco Kuijpers[120,121], Ajith Kumar[3], Dinakantha Kumararatne[114], Manju Kurian[122,123], Michael Laffan[124,125], Fiona Lalloo[77], Michele Lambert[126,127], Hana Lango Allen[1,16], Allan Lawrie[128], Mark Layton[124], Claire Lentaigne[124,125], Adam Levine[85], Rachel Linger[1,16], Hilary Longhurst[99], Eleni Louka[18,19], Robert MacKenzie Ross[129], Bella Madan[130], Eamonn Maher[11,131], Jesmeen Maimaris[88], Sarah Mangles[132], Rutendo Mapeta[1,16], Kevin Marchbank[64], Stephen Marks[9], Hugh S Markus[69], Hanns-Ulrich Marschall[133], Andrew Marshall[134,135,136], Jennifer Martin[1,16,29], Mary Mathias[137], Emma Matthews[22,138], Heather Maxwell[139], Paul McAlinden[64], Mark McCarthy[19,103,140], Stuart Meacham[1,16], Adam Mead[141], Karyn Megy[1,16], Sarju Mehta[142], Michel Michaelides[14], Carolyn Millar[124,125], Shahin Moledina[9], David Montani[89,90,91], Tony Moore[14,15], Nicholas Morrell[1,29], Monika Mozere[85], MPGN/C3 Glomerulopathy Rare Renal Disease group[143], Keith Muir[144], Andrew Mumford[145,146], Michael Newnham[29], Jennifer O'Sullivan[130], Samya Obaji[56], Steven Okoli[18,19], Andrea Olschewski[116], Horst Olschewski[116,117], Kai Ren Ong[54], Elizabeth Ormondroyd[67,68], Willem Ouwehand[1,16], Sofia Papadia[1,16], Soo-Mi Park[12,13,147], David Parry[5], Joan Paterson[11,12,13], Andrew Peacock[53], John Peden[26], Kathelijne Peerlinck[84], Christopher Penkett[1,16], Joanna Pepke-Zaba[148],

Romina Petersen[1,16], Angela Pyle[107], Stuart Rankin[44], Anupama Rao[9], F Lucy Raymond[1,11], Paula Rayner-Matthews[1,16], Christine Rees[26], Augusto Rendon[45], Tara Renton[43], Andrew Rice[149,150], Sylvia Richardson[17], Alex Richter[10], Irene Roberts[18,19,151], Catherine Roughley[76], Noemi Roy[18,19,151], Omid Sadeghi-Alavijeh[85], Moin Saleem[27], Nilesh Samani[152], Alba Sanchis-Juan[1,16], Ravishankar Sargur[73], Simon Satchell[27], Sinisa Savic[153], Laura Scelsi[87], Sol Schulman[154], Marie Scully[115], Claire Searle[155], Werner Seeger[86], Carrock Sewell[156], Denis Seyres[1,16], Susie Shapiro[65], Olga Sharmardina[1,16], Rakefet Shtoyerman[157], Keith Sibson[137], Lucy Side[3], Ilenia Simeoni[1,16], Michael Simpson[158], Suthesh Sivapalaratnam[55], Anne-Bine Skytte[159], Katherine Smith[45], Kenneth G C Smith[29,160], Katie Snape[161], Florent Soubrier[79], Simon Staines[1,16], Emily Staples[29], Hannah Stark[1,16], Jonathan Stephens[1,16], Kathleen Stirrups[1,16], Sophie Stock[1,16], Jay Suntharalingam[129], Emilia Swietlik[29], R Campbell Tait[162], Kate Talks[28], Rhea Tan[69], James Thaventhiran[29], Andreas Themistocleous[25], Moira Thomas[163], Kate Thomson[67,68], Adrian Thrasher[9], Chantal Thys[84], Marc Tischkowitz[164], Catherine Titterton[1,16], Cheng-Hock Toh[71], Mark Toshner[29], Matthew Traylor[69], Carmen Treacy[29,148], Richard Trembath[41], Salih Tuna[1,16], Wojciech Turek[44], Ernest Turro[1,16,17], Tom Vale[25], Chris Van Geet[84], Natalie Van Zuydam[25], Marta Vazquez-Lopez[49], Julie von Ziegenweidt[1,16], Anton Vonk Noordegraaf[30], Quintin Waisfisz[30], Suellen Walker[9], James Ware[38,39,58], Hugh Watkins[67,68,103], Christopher Watt[1,16], Andrew Webster[14,15], Wei Wei[50], Steven Welch[100], Julie Wessels[92], Sarah Westbury[145,146], John-Paul Westwood[115], John Wharton[62], Deborah Whitehorn[1,16], James Whitworth[11,12,13], Martin R Wilkins[62], Catherine Williamson[48,165], Edwin Wong[98], Nicholas Wood[166,167], Yvette Wood[1,16], Geoff Woods[11,114], Emma Woodward[77], Stephen Wort[39,41], Austen Worth[9], Katherine Yates[1,16,29], Patrick Yong[168], Tim Young[1,16], Ping Yu[1,16], Patrick Yu-Wai-Man[50]

Affiliations

[1]NIHR BioResource, Cambridge University Hospitals NHS Foundation, Cambridge Biomedical Campus, Cambridge, UK. [2]Chapel Allerton Hospital, Leeds Teaching Hospitals NHS Trust, Leeds, UK. [3]North East Thames Regional Genetics Service, Great Ormond Street Hospital for Children NHS Foundation Trust, London, UK. [4]MRC Clinical Sciences Centre, Faculty of Medicine, Imperial College London, London, UK. [5]Institute of Genetics and Molecular Medicine, University of Edinburgh, Edinburgh, UK. [6]Salford Royal NHS Foundation Trust, Salford, UK. [7]Queens Centre for Haematology and Oncology, Castle Hill Hospital, Hull and East Yorkshire NHS Trust, Cottingham, UK. [8]European Molecular Biology Laboratory, European Bioinformatics Institute (EMBL-EBI), Wellcome Genome Campus, Hinxton, Cambridge, UK. [9]Great Ormond Street Hospital for Children NHS Foundation Trust, London, UK. [10]University Hospitals Birmingham NHS Foundation Trust, Birmingham, UK. [11]Department of Medical Genetics, Cambridge Institute for Medical Research, University of Cambridge, Cambridge Biomedical Campus, Cambridge, UK. [12]Cancer Research UK Cambridge Centre, Cambridge Biomedical Campus, Cambridge, UK. [13]NIHR Cambridge Biomedical Research Centre, Cambridge Biomedical Campus, Cambridge, UK. [14]Moorfields Eye Hospital NHS Foundation Trust, London, UK. [15]UCL Institute of Ophthalmology, University College London, London, UK. [16]Department of Haematology, University of Cambridge, Cambridge Biomedical Campus, Cambridge, UK. [17]MRC Biostatistics Unit, Cambridge Institute of Public Health, University of Cambridge, Cambridge, UK. [18]MRC Molecular Haematology Unit, Weatherall Institute of Molecular Medicine, University of Oxford, Oxford, UK. [19]NIHR Oxford Biomedical Research Centre, Oxford University Hospitals Trust, Oxford, UK. [20]Center for Clinical Transfusion Medicine, University Hospital of Tübingen, Tübingen, Germany. [21]The Katharine Dormandy Haemophilia Centre and Thrombosis Unit, Royal Free London NHS Foundation Trust, London, UK. [22]University College London, London, UK. [23]Department of Clinical Genetics, Leicester

Royal Infirmary, University Hospitals of Leicester, Leicester, UK. [24]University of Leicester, Leicester, UK. [25]The Nuffield Department of Clinical Neurosciences, University of Oxford, John Radcliffe Hospital, Oxford, UK. [26]Illumina Limited, Chesterford Research Park, Little Chesterford, Nr Saffron Walden, UK. [27]Bristol Renal, University of Bristol, Bristol, UK. [28]Haematology Department, Royal Victoria Infirmary, The Newcastle upon Tyne Hospitals NHS Foundation Trust, Newcastle upon Tyne, UK. [29]Department of Medicine, School of Clinical Medicine, University of Cambridge, Cambridge Biomedical Campus, Cambridge, UK. [30]Department of Pulmonary Medicine, VU University Medical Centre, Amsterdam, The Netherlands. [31]Southampton General Hospital, University Hospital Southampton NHS Foundation Trust, Southampton, UK. [32]MRC Social, Genetic & Developmental Psychiatry Centre, Institute of Psychiatry, Psychology & Neuroscience, King's College London, London, UK. [33]NIHR Biomedical Research Centre for Mental Health, Maudsley Hospital, London, UK. [34]Newcastle University, Newcastle upon Tyne, UK. [35]Newcastle upon Tyne Hospitals NHS Foundation Trust, Newcastle upon Tyne, UK. [36]Department of Clinical Genetics, Royal Devon & Exeter Hospital, Royal Devon and Exeter NHS Foundation Trust, Exeter, UK. [37]Department of Immunology, Leicester Royal Infirmary, Leicester, UK. [38]National Heart and Lung Institute, Imperial College London, Royal Brompton Hospital, London, UK. [39]Royal Brompton Hospital, Royal Brompton and Harefield NHS Foundation Trust, London, UK. [40]Royal Free London NHS Foundation Trust, London, UK. [41]King's College London, London, UK. [42]Guy's and St Thomas' Hospital, Guy's and St Thomas' NHS Foundation Trust, London, UK. [43]King's College Hospital NHS Foundation Trust, London, UK. [44]High Performance Computing Service, University of Cambridge, Cambridge, UK. [45]Genomics England Ltd, London, UK. [46]Epidemiology and Biostatistics, Imperial College London, London, UK. [47]Imperial College Healthcare NHS Trust, London, UK. [48]Division of Women's Health, King's College London, London, UK. [49]Women's Health Research Centre, Surgery and Cancer, Faculty of Medicine, Hammersmith Hospital, Imperial College Healthcare NHS Trust, London, UK. [50]Department of Clinical Neurosciences, School of Clinical Medicine, University of Cambridge, Cambridge Biomedical Campus, Cambridge, UK. [51]Medical Research Council Mitochondrial Biology Unit, Cambridge Biomedical Campus, Cambridge, UK. [52]Children's Renal and Urology Unit, Nottingham Children's Hospital, QMC, Nottingham University Hospitals NHS Trust, Nottingham, UK. [53]Golden Jubilee National Hospital, Glasgow, UK. [54]West Midlands Regional Genetics Service, Birmingham Women's and Children's NHS Foundation Trust, Birmingham, UK. [55]The Royal London Hospital, Barts Health NHS Foundation Trust, London, UK. [56]The Arthur Bloom Haemophilia Centre, University Hospital of Wales, Cardiff, UK. [57]Sheffield Pulmonary Vascular Disease Unit, Royal Hallamshire Hospital NHS Foundation Trust, Sheffield, UK. [58]MRC London Institute of Medical Sciences, Imperial College London, London, UK. [59]National Heart Research Institute Singapore, National Heart Centre Singapore, Singapore, Singapore. [60]Division of Cardiovascular & Metabolic Disorders, Duke-National University of Singapore, Singapore, Singapore. [61]Imperial College Renal and Transplant Centre, Hammersmith Hospital, Imperial College Healthcare NHS Trust, London, UK. [62]Department of Medicine, Imperial College London, London, UK. [63]National Pulmonary Hypertension Service (Newcastle), The Newcastle upon Tyne Hospitals NHS Foundation Trust, Newcastle upon Tyne, UK. [64]Institute of Cellular Medicine, Faculty of Medical Sciences, Newcastle University, Newcastle upon Tyne, UK. [65]Oxford Haemophilia and Thrombosis Centre, The Churchill Hospital, Oxford University Hospitals NHS Trust, Oxford, UK. [66]Department of Molecular Medicine, General Biology, and Medical Genetics Unit, University of Pavia, Pavia, Italy. [67]Department of Cardiovascular Medicine, Radcliffe Department of Medicine, University of Oxford, Oxford, UK. [68]Oxford University Hospitals NHS Foundation Trust, Oxford, UK. [69]Stroke Research Group, Department of Clinical Neurosciences, University of

Cambridge, Cambridge Biomedical Campus, Cambridge, UK. [70]Nottingham University Hospitals NHS Trust, Nottingham, UK. [71]The Roald Dahl Haemostasis and Thrombosis Centre, The Royal Liverpool Hospital, Liverpool, UK. [72]Regional Immunology Service, Kelvin Building, Royal Victoria Hospital, Belfast, UK. [73]Sheffield Teaching Hospitals NHS Foundation Trust, Sheffield, UK. [74]Pathology and Laboratory Medicine, University of Western Australia, Crawley, Australia. [75]Ramón Sardá Mother's and Children's Hospital, Buenos Aires, Argentina. [76]Haemophilia Centre, Kent & Canterbury Hospital, East Kent Hospitals University Foundation Trust, Canterbury, UK. [77]Manchester Centre for Genomic Medicine, Saint Mary's Hospital, Manchester, UK. [78]Salisbury District Hospital, Salisbury NHS Foundation Trust, Salisbury, UK. [79]Departement de genetique, Hopital Pitie-Salpetriere, Paris, France. [80]Service d'Hématologie Biologique, Hôpital d'enfants Armand Trousseau, Paris, France, Paris, France. [81]Inserm U1170, Villejuif, France. [82]Assistance Publique-Hôpitaux de Paris, Département d'Hématologie, Hôpital Armand Trousseau, Paris, France. [83]Department of Paediatrics, School of Clinical Medicine, University of Cambridge , Cambridge Biomedical Campus, Cambridge, UK. [84]Department of Cardiovascular Sciences, Center for Molecular and Vascular Biology, University of Leuven, Leuven, Belgium. [85]UCL Centre for Nephrology, University College London, London, UK. [86]University of Giessen and Marburg Lung Center (UGMLC), Giessen, Germany. [87]Division of Cardiology, Fondazione IRCCS Policlinico S. Matteo, Pavia, Italy. [88]UCL Great Ormond Street Institute of Child Health, London, UK. [89]Universite Paris-Sud, Le Kremlin-Bicêtre, France. [90]Service de Pneumologie, DHU Thorax Innovation, Hôpital Bicêtre, Le Kremlin-Bicêtre, France. [91]INSERM U999, LabEx LERMIT, Centre Chirurgical Marie Lannelongue, Le Plessis Robinson, France. [92]University Hospitals of North Midlands NHS Trust, Stoke-on-Trent, UK. [93]East Yorkshire Regional Adult Immunology and Allergy Unit, Hull Royal Infirmary, Hull & East Yorkshire Hospitals NHS Trust, Hull, UK. [94]Department of Clinical Genetics, Liverpool Women's NHS Foundation, Liverpool, UK. [95]Institute for Immunology and Transfusion Medicine, University of Greifswald, Greifswald, Germany. [96]Section of Internal and Cardiovascular Medicine, University of Perugia, Perugia, Italy. [97]Mitochondrial Research Group, Institute of Genetic Medicine, Newcastle University, Newcastle upon Tyne, UK. [98]Institute of Genetic Medicine, Newcastle University, Newcastle upon Tyne, UK. [99]Barts Health NHS Foundation Trust, London, UK. [100]Birmingham Heartlands Hospital, Heart of England NHS Foundation Trust, Birmingham, UK. [101]ARCH: Australian Research Centre for Health of Women and Babies, Robinson Research Institute, Discipline of Obstetrics and Gynaecology, The University of Adelaide, Women's and Children's Hospital, Adelaide, Australia. [102]Department of Neurology, Sheffield Teaching Hospitals NHS Foundation Trust, Sheffield, UK. [103]Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, UK. [104]Department of Neurology, Leeds Teaching Hospital NHS Trust, Leeds, UK. [105]Epsom & St Helier University Hospitals NHS Trust, London, UK. [106]Northern Genetics Service, The Newcastle upon Tyne Hospitals NHS Foundation Trust, International Centre for Life, Newcastle upon Tyne, UK. [107]Wellcome Centre for Mitochondrial Research, Institute of Genetic Medicine, Newcastle University, Newcastle upon Tyne, UK. [108]John Walton Muscular Dystrophy Research Centre, Institute of Genetic Medicine, Newcastle University, Newcastle upon Tyne, UK. [109]Wellcome Trust Sanger Institute, Hinxton, Cambridge, UK. [110]Department of Clinical Genetics, Guy's and St Thomas' NHS Foundation Trust, London, UK. [111]Department of Paediatric Nephrology, Great North Children's Hospital, Newcastle upon Tyne Hospitals NHS Foundation Trust, Newcastle upon Tyne, UK. [112]University Hospital of Wales, Cardiff, UK. [113]Cardiff & Vale University LHB, Cardiff, UK. [114]Addenbrookes Hospital, Cambridge University Hospitals NHS Foundation Trust, Cambridge, UK. [115]University College London Hospitals NHS Foundation Trust, London, UK. [116]Ludwig Boltzmann Institute for Lung Vascular Research, Graz, Austria. [117]Dept of Internal Medicine, Division of

Pulmonology, Medical University of Graz, Graz, Austria. [118]Division of Transplantation Immunology and Mucosal Biology, Department of Experimental Immunobiology, Faculty of Life Sciences and Medicine, King's College London, London, UK. [119]Department of Paediatric Nephrology, Evelina London Children's Hospital, Guy's & St Thomas' NHS Foundation Trust, London, UK. [120]Department of Pediatric Hematology, Immunology, Rheumatology and Infectious Diseases, Emma Children's Hospital, Academic Medical Center (AMC), University of Amsterdam, Amsterdam, The Netherlands. [121]Department of Clinical Genetics, Academic Medical Center (AMC), University of Amsterdam, Amsterdam, The Netherlands. [122]Molecular Neurosciences, Developmental Neurosciences, UCL Great Ormond Street Institute of Child Health, London, UK. [123]Department of Neurology, Great Ormond Street Hospital for Children NHS Foundation Trust, London, UK. [124]Department of Haematology, Hammersmith Hospital, Imperial College Healthcare NHS Trust, London, UK. [125]Department of Haematology, Imperial College London, London, UK. [126]Division of Hematology, The Children's Hospital of Philadelphia, Philadelphia, USA. [127]Department of Pediatrics, Perelman School of Medicine at the University of Pennsylvania, Philadelphia, USA. [128]Department of Infection, Immunity & Cardiovascular Disease, University of Sheffield, Sheffield, UK. [129]Royal United Hospitals Bath NHS Foundation Trust, Bath, UK. [130]Department of Haematology, Guy's and St Thomas' NHS Foundation Trust, London, UK. [131]Cambridge NIHR Biomedical Research Centre, Cambridge Biomedical Campus, Cambridge, UK. [132]Haemophilia, Haemostasis and Thrombosis Centre, Hampshire Hospitals NHS Foundation Trust, Basingstoke, UK. [133]Wallenberg Laboratory, Department of Molecular and Clinical Medicine, Sahlgrenska Academy, University of Gothenburg, Gothenburg, Sweden. [134]Faculty of Medical and Human Sciences, Centre for Endocrinology and Diabetes, Institute of Human Development, University of Manchester, Manchester, UK. [135]Department of Clinical Neurophysiology, Manchester Royal Infirmary, Central Manchester University Hospitals National Health Service Foundation Trust, Manchester Academic Health Science Centre, Manchester, UK. [136]National Institute for Health Research/Wellcome Trust Clinical Research Facility, Manchester, UK. [137]Department of Haematology, Great Ormond Street Hospital for Children NHS Foundation Trust, London, UK. [138]The National Hospital for Neurology and Neurosurgery, University College London Hospitals NHS Foundation Trust, London, UK. [139]Royal Hospital for Children, NHS Greater Glasgow and Clyde, Glasgow, UK. [140]Oxford Centre for Diabetes, Endocrinology and Metabolism, University of Oxford, Churchill Hospital, Oxford, UK. [141]Centre for Haematology, Department of Medicine, Hammersmith Hospital, Imperial College Healthcare NHS Trust, London, UK. [142]Department of Clinical Genetics, Addenbrookes Hospital, Cambridge University Hospitals NHS Foundation Trust, Cambridge, UK. [143]MPGN/C3 Glomerulopathy Rare Renal Disease group, , UK. [144]Institute of Neuroscience and Psychology, University of Glasgow, Glasgow, UK. [145]School of Cellular and Molecular Medicine, University of Bristol, Bristol, UK. [146]University Hospitals Bristol NHS Foundation Trust, Bristol, UK. [147]East Anglian Regional Genetics Service, Cambridge University Hospitals NHS Foundation Trust, Cambridge, UK. [148]Royal Papworth Hospital NHS Foundation Trust, Cambridge, UK. [149]Pain Research, Department of Surgery and Cancer, Faculty of Medicine, Imperial College London, London, UK. [150]Chelsea and Westminster Hospital NHS Foundation Trust, London, UK. [151]Department of Paediatrics, Weatherall Institute of Molecular Medicine, University of Oxford, Oxford, UK. [152]Departments of Cardiovascular Sciences and NIHR Leicester Cardiovascular Biomedical Research Unit, University of Leicester, Leicester, UK. [153]The Leeds Teaching Hospitals NHS Trust, Leeds, UK. [154]Beth Israel Deaconess Medical Centre and Harvard Medical School, Boston, USA. [155]Department of Clinical Genetics, Nottingham University Hospitals NHS Trust, Nottingham, UK. [156]Scunthorpe General Hospital, Northern Lincolnshire and Goole NHS Foundation Trust, Scunthorpe, UK. [157]Clinical

Genetics Institute, Kaplan Medical Center, Rehovot, Israel. [158]Genetics and Molecular Medicine, King's College London, London, UK. [159]Aarhus University Hospital, Aarhus, Denmark. [160]Cambridge Institute for Medical Research, University of Cambridge, Cambridge Biomedical Campus, Cambridge, UK. [161]Department of Clinical Genetics, St George's University Hospitals NHS Foundation Trust, London, UK. [162]Glasgow Royal Infirmary, NHS Greater Glasgow and Clyde, Glasgow, UK. [163]Gartnavel General Hospital, NHS Greater Glasgow and Clyde, Glasgow, UK. [164]Addenbrooke's Treatment Centre, Addenbrooke's Hospital, Cambridge University Hospitals NHS Foundation Trust, Cambridge, UK. [165]Institute of Reproductive and Developmental Biology, Surgery and Cancer, Hammersmith Hospital, Imperial College Healthcare NHS Trust, London, UK. [166]Department of Molecular Neuroscience, UCL Institute of Neurology, London, UK. [167]UCL Genetics Institute, London, UK. [168]Frimley Park Hospital, NHS Frimley Health Foundation Trust, Camberley, UK.

US PAH Biobank Consortium

Alison Theuer[72], Anagha Malur[59], Ann Williams[36], Anne Dotson[73], Ashley Warden[18], Brandy Harrington[69], Brenda Vang[69], Caitlin Ziemak[22], Nancy Casanova[41], Elizabeth Caskey[71], Catherine MacDonald[41], Courtney Rowley[18], Daniel Larimore[18], Daniela Brady[63], David Tomer[35], Anne Davis[75], Debra Broach[50], Jane Devereux[63], Ellen Lovato[27], Eric Stratton[66], Erin Turk[50], Esperanza Jackson[21], Gina Paciotti[69], Gretchen Peichel[69], Hellina Birru[22], Holly del Junco[20], Rebecca Ingledue[36], Jackie Bruno[72], Jan Drake[24], Jennifer Marks[50], Jessica Pisarcik[67], Jill Spears[18], Joseph Santiago[24], Jordyn DeMartino[18], Joy Beckmann[57], Julia Palmer[18], Karen Visnaw[44], Karla Kennedy[37], Karlise Lewis[46], Kathleen Miller-Reed[46], Kelly Hannon[20], Kimberly McClain[73], Laura Dillon[69], Lekan Olanipekun[70], Lillian Mendibles[41], Lindsey Hawke[17], Linnea Brody[75], Louise Durst[39], Mary Andrews[65], Melissa Allahua[47], Melissa Stratoberdha[26], Merte Lemma[56], Molly Cope[73], Mya Franzo[73], Natalia Feliz[21], Nicholas Hawkes[41], Tracy Norwood[71], Opal Wilson[71], Amy Palmisciano[47], Peggy Gruhlke[39], Priscilla Correa[23], Randy Blake[57], Reema Karnekar[59], Robyn Do[18], Kristal Rohwer[39], Sara Ahmed[22], Kimberly Schiltz[15], Page Scovel[15], Shannon Cordell[60], Sharon Heuerman[27], Shazzra Ahmad[64], Sylwia Szuberla[20], Tammy Roads[36], Thoeun Iem[39], Traci Mcgaha[67], Tracy Urban[46], Valerie Aston[35], Waleed Ayesh[64], Allison Light[72], Abby Arkon[19], Andrea Tavlarides[26], Audrey Anderson[21], John Bindu[57], Deedre Boekwig[35], Donna Singleton[71], Inna Abrea[26], Jennifer Lee[40], Mark Ormiston[17], Renesa Whitman[21], Royanne Holy[40], Sisama Almeida-Peters[37], Tosin Igenoza[70], Hap Farber[66], Auvo Reponen[45], Mukta Barve[45], Amber Gygi[45], Clayborne Winslow[45]

[9] Division of Human Genetics, Cincinnati Children's Hospital Medical Center, Department of Pediatrics, University of Cincinnati College of Medicine, Cincinnati, OH, United States;
[15] Division of Cardiovascular Medicine, University of Iowa, Iowa City IA, United States;
[17] Queen’s University, Kingston ON, Canada;
[18] Medical University of South Carolina, Charleston SC, United States;
[19] Vanderbilt University-Peds, Nashville TN, United States;
[20] University of Colorado Denver, Aurora CO, United States;
[21] Baylor Research Institute, Plano TX, United States;
[22] Medstar Health, Washington D.C., United States;
[23] Allegheny-Singer Research Institute, Pittsburgh PA, United States;
[24] Ohio State University, Columbus OH, United States;
[26] Mayo Clinic Florida, Jacksonville FL, United States;
[27] Washington University, St. Louis MO, United States;

[35] Department of Medicine at Intermountain Medical Center and the University of Utah, Murray UT, United States;
[36] University of Cincinnati, Cincinnati OH, United States;
[37] Duke University Medical Center, Durham NC, United States;
[39] Mayo Clinic, Rochester MN, United States;
[40] Weill Cornell Medical College and The Houston Methodist Hospital, Houston TX, United States;
[41] Department of Medicine and Arizona Health Sciences Center, University of Arizona, Tucson, AZ, United States;
[44] Tufts Medical Center, Boston MA, United States;
[45] Cincinnati Children's Hospital, Cincinnati OH, United States;
[46] Children's Hospital of Colorado, University of Colorado Denver, Aurora CO, United States;
[47] Rhode Island Hospital, Providence RI, United States;
[50] Indiana University, Indianapolis IN, United States;
[56] Inova Heart and Vascular Institute, Falls Church VA, United States;
[57] LA Biomedical Research Institute at Harbor-UCLA, Torrance CA, United States;
[59] East Carolina University, Greenville NC, United States;
[60] Vanderbilt University Medical Center, Nashville TN, United States;
[63] Columbia University, New York NY, United States;
[64] Wayne State University, Detroit MI, United States;
[65] University Hospital of Cleveland, Cleveland OH, United States;
[66] Boston University School of Medicine, Boston MA, United States;
[67] University of Pittsburgh, Pittsburgh PA, United States;
[69] University of Minnesota, Minneapolis MN, United States;
[70] UT Southwestern, Dallas TX, United States;
[71] LSU Health, Shreveport LA, United States;
[72] University of Rochester Medical Center, Rochester NY, United States;
[73] Spectrum Health Hospitals, Grand Rapids MI, United States;
[75] Seattle Children's Hospital, Seattle WA, United States;