**CARDIOVASCULAR MEDICINE: TOWARDS A MOLECULAR CLASSIFICATION OF PULMONARY HYPERTENSION**

Martin R Wilkins
Imperial College London

This work is submitted for fulfilment of the requirements for the degree of Doctor of Science (DSc) from Imperial College London

January 2022

**LIST OF PUBLICATIONS**

**Proteomics**

1. PMID: 21041689

Abdul-Salam VB, Wharton J, Cupitt J, Berryman M, Edwards RJ, **Wilkins MR.** (2010) Proteomic analysis of lung tissues from patients with pulmonary arterial hypertension *CIRCULATION.* 122:2058-67. Doi: 10.1161/CIRCULATIONAHA.110.972745.

2. PMID: 24503951

Wojciak-Stothard B, Abdul-Salam VB, Lao KH, Tsang H, Irwin DC, Lisk C, Loomis Z, Stenmark KR, Edwards JC, Yuspa SH, Howard LS, Edwards RJ, Rhodes CJ, Gibbs JS, Wharton J, Zhao L, **Wilkins MR**. (2014) Aberrant chloride intracellular channel 4 expression contributes to endothelial dysfunction in pulmonary arterial hypertension. *CIRCULATION*. 129:1770-80. Doi:10.1161/CIRCULATIONAHA.113.006797

3. PMID: 28624389

Rhodes CJ, Wharton J, Ghataorhe P, Watson G, Girerd B, Howard LS, Gibbs JSR, Condliffe R, Elliott CA, Kiely DG, Simonneau G, Montani D, Sitbon O, Gall H, Schermuly RT, Ghofrani HA, Lawrie A, Humbert M, **Wilkins MR.** (2017). Plasma proteome analysis in patients with pulmonary arterial hypertension: an observational cohort study. *LANCET RESPIRATORY MEDICINE.* 5: 717-726. doi:10.1016/S2213-2600(17)30161-3.

4. PMID: 35081018

Rhodes CJ, Wharton J, Swietlik EM, Harbaum L, Girerd B, Coghlan JG, Lordan J, Church C, Pepke-Zaba J, Toshner M, Wort SJ, Kiely DG, Condliffe R, Lawrie A, Gräf S, Montani D, Boucly A, Sitbon O, Humbert M, Howard LS, Morrell NW, **Wilkins MR**, on behalf of the UK National PAH Cohort Study Consortium (2022). Using the Plasma Proteome for Risk Stratifying Patients with Pulmonary Arterial Hypertension. *AMERICAN JOURNAL OF RESPIRATORY AND CRITICAL CARE MEDICINE*. DOI: https://doi.org/10.1164/rccm.202105-1118OC

**Metabolomics**

5. PMID: 27881557

Rhodes CJ, Ghataorhe P, Wharton J, Rue-Albrecht KC, Hadinnapola C, Watson G, Bleda M, Haimel M, Coghlan G, Corris PA, Howard LS, Kiely DG, Peacock AJ, Pepke-Zaba J, Toshner MR, Wort SJ, Gibbs JSR, Lawrie A, Gräf S, Morrell NW, **Wilkins MR.** (2017). Plasma Metabolomics Implicates Modified Transfer RNAs and Altered Bioenergetics in the Outcomes of Pulmonary Arterial Hypertension. *CIRCULATION.* 135: 460-+. doi:10.1161/CIRCULATIONAHA.116.024602.

6. PMID: 30478197

Harbaum L, Ghataorhe P, Wharton J, Jimenez B, Howard LSG, Gibbs JSR, Nicholson JK, Rhodes CJ, **Wilkins MR.** (2019). Reduced plasma levels of small HDL particles transporting fibrinolytic proteins in pulmonary arterial hypertension. *THORAX.* 74: 380-389. doi:10.1136/thoraxjnl-2018-212144.

**Genetics**

7. PMID: 32581362

Turro, E., Astle, W.J., Megy, K. Gräf S, Greene D, Shamardina O, Allen HL, Sanchis-Juan A, Frontini M, Thys C, Stephens J, Mapeta R, Burren OS, Downes K, Haimel M, Tuna S, Deevi SVV, Aitman TJ, Bennett DL, Calleja P, Carss K, Caulfield MJ, Chinnery PF, Dixon PH, Gale DP, James R, Koziell A, Laffan MA, Levine AP, Maher ER, Markus HS, Morales J, Morrell NW, Mumford AD, Ormondroyd E, Rankin S, Rendon A, Richardson S, Roberts I, Roy NBA, Saleem MA, Smith KGC, Stark H, Tan RYY, Themistocleous AC, Thrasher AJ, Watkins H, Webster AR, **Wilkins MR**, Williamson C, Whitworth J, Humphray S, Bentley DR, NIHR BioResource for the 100,000 Genomes Project, Kingston N, Walker

N, Bradley JR, Ashford S, Penkett CJ, Freson K, Stirrups KE, Raymond FL, Ouwehand WH. (2020) Whole-genome sequencing of patients with rare diseases in a national health system. *NATURE.* 583: 96–102. https://doi.org/10.1038/s41586-020-2434-2

8. PMID: 29650961
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. (2018). Identification of rare sequence variation underlying heritable pulmonary arterial hypertension. *NATURE COMMUNICATIONS.* 9:1416. doi:10.1038/s41467-018-03672-4.

9. PMID: 30527956
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 HA, 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, MacKenzie Ross RV, 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, Vonk Noordegraaf A, Waisfisz Q, Walsworth AK, Walter RE, Wharton J, White RJ, Wilt J, Wort SJ, Yung D, Lawrie A, Humbert M, Soubrier F, Trégouët DA, Prokopenko I, Kittles R, Gräf S, Nichols WC, Trembath RC, Desai AA, Morrell NW, **Wilkins MR**; UK NIHR BioResource Rare Diseases Consortium; UK PAH Cohort Study Consortium; US PAH Biobank Consortium.. (2019). Genetic determinants of risk in pulmonary arterial hypertension: international genome-wide association studies and meta-analysis. *LANCET RESPIRATORY MEDICINE.* 7:227-238. doi:10.1016/S2213-2600(18)30409-0

10. PMID: 25373139
**Wilkins MR**, Aldashev A A, Wharton J, Rhodes CJ, Vandrovcova J, Kasperaviciute D, Bhosle SG, Mueller M, Geschka S, Rison S, Kojonazarov B, Morrell NW, Neidhardt I, Surmeli NB, Aitman TJ, Stasch J-P, Behrends S, Marletta MA. (2014). alpha 1-A680T Variant in GUCY1A3 as a Candidate Conferring Protection From Pulmonary Hypertension Among Kyrgyz Highlanders. *CIRCULATION-CARDIOVASCULAR GENETICS.* 7: 920-U505. doi:10.1161/CIRCGENETICS.114.000763

11. PMID: 26258299
Zhao L, Oliver E, Maratou K, Atanur SS, Dubois OD, Cotroneo E, Chen CN, Wang L, Arce C, Chabosseau PL, Ponsa-Cobas J, Frid MG, Moyon B, Webster Z, Aldashev A, Ferrer J, Rutter GA, Stenmark KR, Aitman TJ, **Wilkins MR**. (2015). The zinc transporter ZIP12 regulates the pulmonary vascular response to chronic hypoxia. *NATURE.* 524:356-60. DOI: 10.1038/nature14620

**microRNAs and transcriptomics**

12. PMID: 23220912
Rhodes C J, Wharton J, Boon RA, Roexe T, Tsang H, Wojciak-Stothard B, Chakrabarti A, Howard LS, Gibbs JSR, Lawrie A, Condliffe R, Elliott CA, Kiely DG, Huson, L, Ghofrani HA, Tiede H, Schermuly R, Zeiher AM, Dimmeler S, **Wilkins MR**. (2013). Reduced MicroRNA-150 Is Associated with Poor Survival in Pulmonary Arterial Hypertension. *AMERICAN JOURNAL OF RESPIRATORY AND CRITICAL CARE MEDICINE.* 187:294-302. doi:10.1164/rccm.201205-0839OC

13. PMID: 32352834
Rhodes CJ, Otero-Núñez P, Wharton J, Swietlik EM, Kariotis S, Harbaum L, Dunning MJ, Elinoff JM, Errington N, Thompson AAR, Iremonger J, Coghlan JG, Corris PA, Howard LS, Kiely DG, Church C, Pepke-Zaba J, Toshner M, Wort SJ, Desai AA, Humbert M, Nichols WC, Southgate L, Trégouët DA, Trembath RC, Prokopenko I, Gräf S, Morrell NW, Wang D, Lawrie A, Wilkins MR. (2020). Whole-Blood RNA Profiles Associated with Pulmonary Arterial Hypertension and Clinical Outcome. *AMERICAN JOURNAL OF RESPIRATORY AND CRITICAL CARE MEDICINE.* 202:586-594. doi:10.1164/rccm.202003-0510OC

14. PMID: 34876579
Kariotis S, Jammeh E, Swietlik EM, Pickworth JA, Rhodes CJ, Otero P, Wharton J, Iremonger J, Dunning MJ, Pandya D, Mascarenhas TS, Errington N, Thompson AAR, Romanoski CE, Rischard F, Garcia JGN, Yuan JX, An TS, Desai AA, Coghlan G, Lordan J, Corris PA, Howard LS, Condliffe R, Kiely DG, Church C, Pepke-Zaba J, Toshner M, Wort S, Gräf S, Morrell NW, **Wilkins MR**, Lawrie A, Wang D; UK National PAH Cohort Study Consortium. (2021) Biological heterogeneity in idiopathic pulmonary arterial hypertension identified through unsupervised transcriptomic profiling of whole blood. *NATURE COMMUNICATIONs.* 12:7104. DOI: 10.1038/s41467-021-27326-0

15. PMID: 21737024
Rhodes CJ, Howard LS, Busbridge M, Ashby D, Kondili E, Gibbs JSR, Wharton J, **Wilkins MR.** (2011). Iron Deficiency and Raised Hepcidin in Idiopathic Pulmonary Arterial Hypertension Clinical Prevalence, Outcomes, and Mechanistic Insights. *JOURNAL OF THE AMERICAN COLLEGE OF CARDIOLOGY.* 58:300-309. doi:10.1016/j.jacc.2011.02.057

**Imaging**

16. PMID: 23900048
Zhao L, Ashek A, Wang L, Fang W, Dabral S, Dubois O, Cupitt J, Pullamsetti SS, Cotroneo E, Jones H, Tomasi G, Nguyen QD, Aboagye EO, El-Bahrawy MA, Barnes G, Howard LS, Gibbs JS, Gsell W, He JG, **Wilkins MR**. (2013). Heterogeneity in Lung 18FDG Uptake in PAH: Potential of Dynamic 18FDG-PET with Kinetic Analysis as a Bridging Biomarker for Pulmonary Remodeling Targeted Treatments. *CIRCULATION.* 128:1214-24. DOI: 10.1161/CIRCULATIONAHA.113.004136

17. PMID: 29070699
Michelakis ED, Gurtu V, Webster L, Barnes G, Watson G, Howard L, Cupitt J, Paterson I, Thompson RB, Chow K, O'Regan DP, Zhao L, Wharton J, Kiely DG, Kinnaird A, Boukouris AE, White C, Nagendran J, Freed DH, Wort SJ, Gibbs JSR, **Wilkins MR**. (2017). Inhibition of pyruvate dehydrogenase kinase improves pulmonary arterial hypertension in genetically susceptible patients. *SCIENCE TRANSLATIONAL MEDICINE.* 9(413):eaao4583. DOI: 10.1126/scitranslmed.aao4583.

18. PMID: 15750042.
**Wilkins MR**, Paul GA, Strange JW, Tunariu N, Gin-Sing W, Banya WA, Westwood MA, Stefanidis A, Ng LL, Pennell DJ, Mohiaddin RH, Nihoyannopoulos P, Gibbs JS. Sildenafil versus Endothelin Receptor Antagonist for Pulmonary Hypertension (SERAPH) study. (2005). *AMERICAN JOURNAL OF RESPIRATORY AND CRITICAL CARE MEDICINE.* 171:1292-7. DOI: 10.1164/rccm.200410-1411OC

19. PMID: 20460548.
**Wilkins MR**, Ali O, Bradlow W, Wharton J, Taegtmeyer A, Rhodes CJ, Ghofrani HA, Howard L, Nihoyannopoulos P, Mohiaddin RH, Gibbs JS; Simvastatin Pulmonary Hypertension Trial (SiPHT) Study Group. (2010). Simvastatin as a treatment for pulmonary hypertension trial. *AMERICAN JOURNAL OF RESPIRATORY AND CRITICAL CARE MEDICINE.* 181:1106-13.
DOI: 10.1164/rccm.2009111-699oc

20. PMID: 28092203
Dawes TJW, de Marvao A, Shi W, Fletcher T, Watson GMJ, Wharton J, Rhodes CJ, Howard LSGE, Gibbs JSR, Rueckert D, Cook SA, **Wilkins MR**, O'Regan DP. (2017). Machine Learning of Three-dimensional Right Ventricular Motion Enables Outcome Prediction in Pulmonary Hypertension: A Cardiac MR Imaging Study. *RADIOLOGY.* 283:381-390. doi:10.1148/radiol.2016161315

21. PMID: 29869959
Dawes TJW, Cai J, Quinlan M, de Marvao A, Ostrowski PJ, Tokarczuk PF, Watson GMJ, Wharton J, Howard LSGE, Gibbs JSR, Cook SA, **Wilkins MR**, O'Regan, DP. (2018). Fractal Analysis of Right Ventricular Trabeculae in Pulmonary Hypertension. *RADIOLOGY.* 288:386-395.
doi:10.1148/radiol.2018172821

22. PMID: 30535300
Attard M, Dawes TJW, de Marvao A, Biffi C, Shi W, Wharton J, Rhodes, CJ, Ghataorhe P, Gibbs JSR, Howard LSGE, Rueckert D, **Wilkins MR**, O'Regan DP. (2019). Metabolic pathways associated with right ventricular adaptation to pulmonary hypertension: 3D analysis of cardiac magnetic resonance imaging. *EUROPEAN HEART JOURNAL-CARDIOVASCULAR IMAGING.* 20:668-676.
doi:10.1093/ehjci/jey175

23. PMID: 30801055
Bello GA, Dawes TJW, Duan J, Biffi C, de Marvao A, Howard LSGE, Gibbs JSR, Wilkins MR, Cook SA, Rueckert D, O'Regan DP. (2019). Deep-learning cardiac motion analysis for human survival prediction. *NATURE MACHINE INTELLIGENCE.* 1:95-+. doi:10.1038/s42256-019-0019-2

**Reviews**

24. PMID: 18591337
Wilkins MR, Wharton J, Grimminger F, Ghofrani HA. (2008). Phosphodiesterase inhibitors for the treatment of pulmonary hypertension. *EUROPEAN RESPIRATORY JOURNAL. 32*:198-209.
doi:10.1183/09031936.00124007

25. PMID: 32201940
Harbaum L, Rhodes CJ, Otero-Nunez P, Wharton J, Wilkins MR. (2020). The application of 'omics' to pulmonary arterial hypertension. BRITISH JOURNAL OF PHARMACOLOGY. 178:108-120.
DOI: 10.1111/bph.15056

26. PMID: 33888254
Oldham WM, Hemnes AR, Aldred MA, Barnard J, Brittain EL., Chan SY, Cheng F, Cho MH, Desai AA, Garcia JGN, Geraci MW, Ghiassian SD, Hall KT, Horn EM, Jain M, Kelly RS, Leopold JA, Lindstrom S, Modena BD, Nichols WC, Rhodes CJ, Sun W, Sweatt AJ, Vanderpool RR, **Wilkins MR**, Wilmot B, Zamanian RT, Fessel JP, Aggarwal NR, Loscalzo J, Xiao, L. (2021). NHLBI-CMREF Workshop Report on Pulmonary Vascular Disease Classification JACC State-of-the-Art Review. *JOURNAL OF THE AMERICAN COLLEGE OF CARDIOLOGY.* 77:2040-2052. doi:10.1016/j.jacc.2021.02.056

27. PMID: 33541614
**Wilkins, M.R.** (2021) Personalized Medicine for Pulmonary Hypertension: The Future Management of Pulmonary Hypertension Requires a New Taxonomy. *CLINICS CHEST MEDICINE.* 42:207-216. DOI: 10.1016/j.ccm.2020.10.004

**STATEMENT AS TO HOW THE PUBLICATIONS SUBMITTED HAVE CONTRIBUTED TO KNOWLEDGE IN THE FIELD OF CARDIOVASCULAR MEDICINE**

**Introduction**

The healthy adult pulmonary circulation is a low resistance vascular bed, with a mean resting pulmonary artery pressure (mPAP) of 14.0±3.3 mmHg. The first world symposium on the subject in 1973 accepted a mPAP ≥25mmHg as an arbitrary definition of pulmonary hypertension[1]. This was revised to >20mmHg in 2018, in part recognising that elevated mPAP represents a continuum of risk, with measured evidence of increased mortality beginning at 19mmHg relative to 10mmHg[2].

The 1973 symposium acknowledged that pulmonary hypertension occurs most commonly in association with lung and left heart disease, but the focus of the meeting was the pathology, clinical features and epidemiology of the more rare presentation of pulmonary hypertension of unknown cause, termed *primary pulmonary hypertension*. A clinical classification of chronic pulmonary heart disease was provided in Annex 1 of the symposium record, with 3 categories based on diseases primarily affecting airways, the thoracic cage and the pulmonary vasculature, respectively[1]. This was revisited and developed further 25 years later at what is recognised as the second world symposium[3]. Here the clinical presentations of pulmonary hypertension were sorted into 5 main categories and the term *pulmonary arterial hypertension or PAH* was introduced to replace *primary pulmonary hypertension*. At the third world symposium in 2003, the classification was widely regarded as useful for clinical and epidemiological purposes, less so for research[4].

The current clinical classification[2], visited in 2018, still holds to the 5 major category structure but is increasingly challenged by observations from molecular science. Patients do not always fit easily into one of the main categories. Not only is there overlap between them, but clinical heterogeneity is present within each category. This lack of precision arises from a lack of understanding of the molecular drivers of pulmonary hypertension and impairs novel drug development.

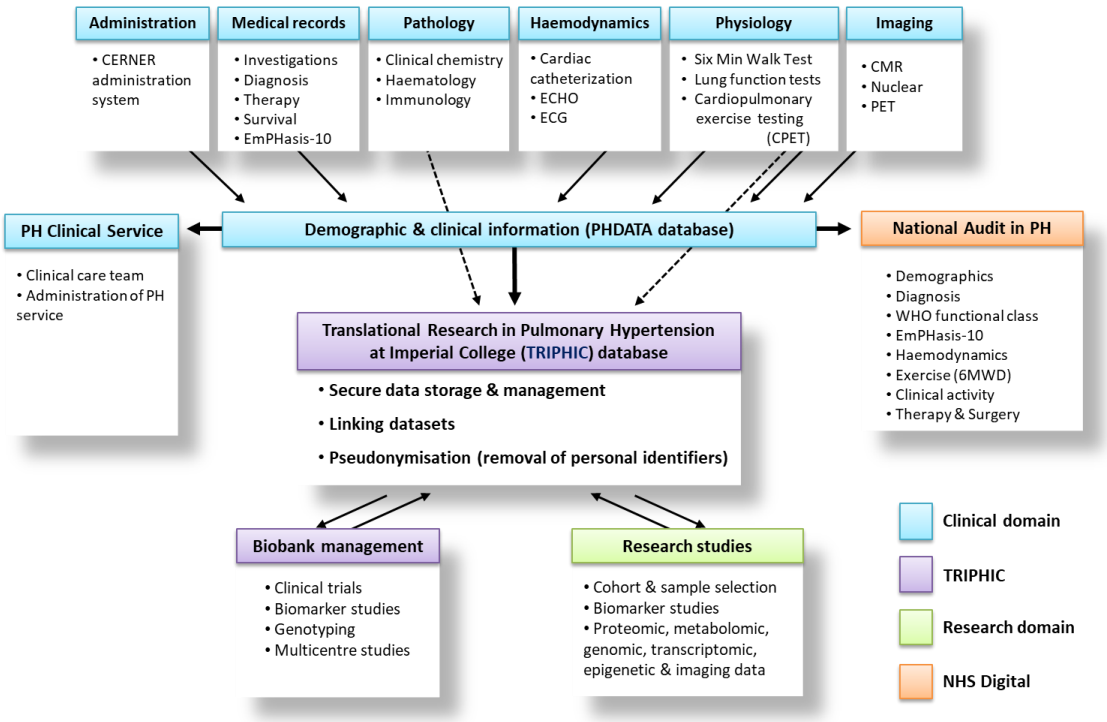
For the past 20 years, my research has been directed at discovering and exploiting the molecular drivers of pulmonary hypertension and has tested the established clinical classification, with a view to providing a better tailored, more personalised approach to the management of the condition.

**Over-arching hypothesis**

The hypothesis that underpins my research is that deep molecular phenotyping of patients presenting with a clinical diagnosis of pulmonary hypertension will identify subgroups that share one or more druggable disease-associated pathways and accessible biomarkers of response to treatment

**Translational Research in Pulmonary Hypertension at Imperial College (TRIPHIC)**

In 2002, I began to assemble data and biosamples from a cohort of patients presenting to Hammersmith Hospital with a tentative diagnosis of pulmonary hypertension. Patients were reviewed and investigated by the clinical team and assigned a diagnostic category and management pathway following international guidelines. Clinical data were collated and, in 2013, I registered a database called TRIPHIC that complied with contemporary regulatory requirements and linked to samples taken at clinical and research visits, aliquoted and stored at -80$^0$C. The database now stands at 2,121 patients and captures clinical features from 20 primary sources, including key investigations such as haemodynamics, imaging and functional measurements, co-morbidities and treatments. The biobank holds over 70,000 coded retrievable samples – DNA, mRNA, plasma, serum, urine, cells and tissue. The majority of patients have a clinical diagnosis of PAH (33%) or chronic thromboembolic disease (23%). Patients with suspected pulmonary hypertension in whom the diagnosis was excluded by cardiac catheterisation (18%) provide disease controls.




**Translational Research in Pulmonary Hypertension at Imperial College (TRIPHIC) database**

The database and biobank have provided the foundation for the molecular interrogation of pulmonary hypertension by myself and others, both locally and internationally. It preceded, and was the major contributor to, the UK National Pulmonary Arterial Hypertension Cohort Study led by Nicholas Morrell from Cambridge, which networked data and sample collection from around 8 expert hospital centres in the UK; the Hammersmith Hospital cohort constituted over 25% of the national cohort. TRIPHIC has provided data and samples for overseas investigators. And it has fuelled a steady stream of discoveries shared in high profile publications that are shaping our concepts about the pathogenesis, diagnosis and management of pulmonary hypertension and the development of new therapies.

**Proteomics**

In an early attempt to identify key molecular pathways in pulmonary vascular disease I used label-free liquid chromatography tandem mass spectrometry to analyse the protein profiles in lung samples from 8 patients with PAH and 8 control subjects (PMID: 21041689). The PAH samples were acquired from patients at lung transplantation while the control specimens came from uninvolved regions of lobectomy tissue from patients undergoing surgery for bronchial carcinoma. From 362 proteins detected, 25 were differentially expressed (either increased or decreased) in PAH lung tissue. Several had been reported in PAH lung previously, giving confidence to the approach and, in particular, credibility to the novel observation of upregulation of chloride intracellular channel protein 4 (CLIC4) in PAH. The latter was confirmed by Western blotting and immunostaining and was taken forward as a molecule of interest with my colleague, Beata Wojciak-Stothard at Imperial (PMID: 24503951). We showed that deletion of CLIC4 attenuated the development of pulmonary hypertension in mice exposed to hypoxia while over-expression compromised endothelial barrier function. We concluded that approaches that reverse increased pulmonary endothelial CLIC4 expression may be beneficial in patients with PAH.

A limitation of studying end-stage disease at lung transplantation is that repair and other compensatory mechanisms may overshadow aberrant pathways responsible for the disease. Lung biopsy is not a safe procedure in PAH and so I turned my attention to proteins that circulate in blood. The plasma proteome is an amalgamation of proteins that are secreted or leak from damaged tissue

and provides an accessible “liquid biopsy” that informs on health. The vast pulmonary vascular bed would be expected to be a major contributor to the plasma proteome and the advent of high-throughput platform technologies offered a route to conducting an unbiased screen of changes in the plasma proteome with pulmonary vascular disease.

I opted for the aptamer-based SomaScan platform operated by SomaLogic. In the first study of this kind, I assayed samples from patients (n=218) with well-phenotyped idiopathic or heritable PAH registered in the TRIPHIC biobank and 136 patients from 2 international cohorts for 1129 circulating proteins (PMID: 28624389). I argued that protein levels associated with survival might be more likely to be related to the biology of condition and, working with Chris Rhodes at Imperial College, found 14 proteins that robustly identified survivors, independent of the cardiac biomarker, N-terminal pro-brain natriuretic peptide (NT-proBNP), and confirmed 9 by targeted immunoassay. These 9 proteins described several pathological processes related to pulmonary vascular disease, such as inflammation, pulmonary vascular cellular dysfunction and structural dysregulation, iron status, and coagulation. A risk score was constructed out of these 9 proteins that improved risk estimates above established clinical risk equations and offered the first multi-marker panel for risk stratifying PAH patients. I recently revisited and refined this multimarker approach by measuring 4,152 proteins by aptamer assay in patients (n=357) with idiopathic, heritable or drug-induced-PAH from the UK National Cohort of PAH and a French cohort (n=79) (PMID:35081018). Thirty-one proteins robustly informed prognosis independent of NT-proBNP and 6 minute walk distance in the UK Cohort. The addition of a weighted combination score of 6 proteins to NT-proBNP improved prediction of 5-year outcomes.

**Metabolomics**

The plasma metabolome, the sum total of circulating metabolites, is, arguably, closer to clinical phenotype than proteins. It provides an equally rich ‘vocabulary’ for describing disease. I used ultraperformance liquid chromatography mass spectrometry to analyze 1416 metabolites in plasma from patients with idiopathic or heritable PAH (n=365) from the UK National PAH Cohort study and identified 52 that distinguished these patients from healthy controls (n=121); 20 of the 53 also separated PAH from disease controls (n=139) (PMID: 27881557). Sixty-two metabolites were prognostic in PAH, with 36 of 62 independent of established prognostic markers. Increased circulating modified nucleosides (N2,N2-dimethylguanosine, N1-methylinosine), TCA cycle intermediates (malate, fumarate), glutamate, fatty acid acylcarnitines, and polyamine metabolites and decreased levels of steroids, sphingomyelins, and phosphatidylcholines are characteristics of patients with PAH that distinguish them from symptomatic patients without pulmonary hypertension. Serial measurements in a subset of patients suggested that correction of these metabolite disturbances is linked to improved outcomes. Of specific note, patients defined as vasoresponders, who have excellent outcomes on calcium channel blocker therapies, demonstrated metabolic profiles more similar to those of healthy control subjects than to other patients.

I followed this analysis with a closer examination of lipid metabolism in PAH, using nuclear magnetic resonance spectroscopy to measure 105 discrete lipoproteins in plasma from PAH patients registered in TRIPHIC (PMID: 30478197). Contemporaneous plasma protein levels from the previous SomaScan platform study were used to identify proteins linked to the lipoprotein subclasses. Reduced levels of small Apo A-2-rich high-density lipoprotein-4 was independently linked with higher mortality. The hypothecated mechanism is that these particles transport fibrinolytic proteins, such as alpha-2-antiplasmin, and vasoactive peptides, such as prekallikrein, and suggest that increasing levels of small high-density lipoprotein and its associated proteins in PAH may have therapeutic benefit.

**Genetics**

The National Institute of Health Research BioResource Rare Diseases (BRIDGE) consortium provided funding to sequence 13 rare diseases and PAH was adopted as one of these (PMID: 32581362). Whole

genome sequencing was performed on 1038 PAH index cases, including 271 from TRIPHIC. Nick Morrell and Stefan Graf from Cambridge led the rare variant analysis and observed over representation of rare variants in 3 new genes in PAH (*ATP13A3*, *AQP1* and *SOX17*), and confirmed the association of GDF2 with the condition (PMID: 29650961).

The common variant analysis of this dataset was led by myself with Inga Prokopenko and Chris Rhodes, both at Imperial College. In a genome-wide association study we identified two common variants that reached the accepted threshold for statistical significance; namely, a locus near *SOX17* and a second locus in *HLA-DPA1* and *HLA-DPB1* (collectively referred to as *HLA-DPA1/DPB1*) within the class II MHC region (PMID: 30527956). These two loci were confirmed in a metanalysis using data from 3 additional international studies, totalling 2085 patients of European ancestry. Deeper analysis revealed that the *SOX17* locus had two independent signals associated with PAH and functional and epigenomic data support the contention that these risk variants alter gene regulation via an enhancer active in endothelial cells. The study indicates that impairment of SOX17 function might be more common in PAH than suggested by rare mutations in *SOX17*. Survival analysis showed that the HLA-DPA1/DPB1 (rs2856830) genotype was strongly associated with survival; median survival from diagnosis in PAH patients with the C/C homozygous genotype was double that of those with the T/T genotype, despite similar baseline disease severity. It makes a case for HLA typing or rs2856830 genotyping patients in clinical trials if time to clinical worsening is a declared trial endpoint.

Genetic variants associated with disease can signal drug targets. SOX17 and HLA-DPA1/DPB1 are not readily druggable. In extending my search for druggable targets with genetic evidence of association with pulmonary hypertension I have looked for protein quantitative trait loci (pQTL) associated with PAH. This involved drawing on extensive plasma proteome and whole genome sequence data from patients (n=357) with idiopathic or heritable PAH, healthy volunteers (n=103) and relatives (n=23) of PAH patients. With Lars Harbaum and Chris Rhodes, I found that triangulating plasma proteins that (i) differentiated PAH from health, (ii) carried prognostic information and (iii) showed genetic control through a pQTL provided 8 proteins that met all 3 criteria. Mendelian randomisation analysis using these pQTL and data from our genome-wide association metanalysis support 3 pQTL as causally related to PAH; namely, netrin-4, thrombospondin-2 and endoglin. The added value of employing Mendelian randomisation is that it helps interpret the significance of altered circulating plasma protein levels. The direction of effect from Mendelian randomisation argues that there is likely to be therapeutic benefit from reducing netrin-4 levels/activity but augmenting thrombospondin-2 and endoglin levels/activity, despite the elevated levels of the latter two in PAH.

In parallel with genetic studies in PAH, I have also led a long-term project to understand the genetic basis of adaptation to hypoxia-induced pulmonary hypertension. The premise is that an increase in pulmonary vascular resistance to alveolar hypoxia, leading to pulmonary hypertension, is a characteristic of most adult mammalian pulmonary circulations; as such, understanding the genetic basis for why some pulmonary vascular beds are more resistant than others will illuminate the underlying mechanisms that might be amenable to pharmacological manipulation. In one success, I compared the exomes of Kyrgyz highlanders who appeared resistant to developing pulmonary hypertension above 3,000m with those from highlanders that were susceptible and found an activating mutation in GUY1A3, that encodes the alpha-subunit of soluble guanylate cyclase (PMID: 25373139); this observation is consistent with a therapeutic strategy that augments the cyclic GMP signalling pathway as a treatment for pulmonary hypertension (PMID: 18591337).

In another successful line of study, I took advantage of a strain of rat, the F344 strain, that, as a result of an experiment of nature, is more resistant to hypoxia-induced pulmonary hypertension than the WKY strain. Through a congenic breeding and detailed phenotyping program, with Lan Zhao at Imperial, and then comparative genomics, with Tim Aitman (also at Imperial College at the time), the

resistance phenotype was narrowed to a mutation in *Slc39a12* which predicted a truncated protein, zinc transporter 12 (Zip12) (PMID: 26258299). The expression of Zip12 in vascular cells is hypoxia dependent and regulates the influx of $Zn^{2+}$ into cells. Knockout of Zip12 in the susceptible strain confers resistance to hypoxia-induced pulmonary hypertension. Zip12 is upregulated in lung tissue from patients with PAH (and hypoxia-induced pulmonary hypertension) and is now the focus of a drug development programme, targeted at inhibiting Zip12 in PAH.

**microRNAs and Transcriptomics**

Circulating microRNAs (miRs) regulate gene expression and reports have emerged of changes in candidate miRs, such as miR-204 and miR-21, in lung tissue from patients with PAH and animal models. I took an unbiased approach, and working with Stefanie Dimmeler from Frankfurt, completed a microarray screen of plasma from 8 patients with PAH and 8 healthy controls and found that miR-150 was markedly downregulated in PAH (PMID: 23220912). Further investigation in samples from the TRIPHIC database (145 patients) and a separate validation group (of 30 treatment naive patients) from Sheffield showed that plasma levels predicted survival. Subsequent studies led by Beata Wojciak-Stothard have shown that miR-150 has anti-apoptotic, anti-proliferative and anti-inflammatory properties and that endothelium-targeted delivery of miR-150 has a protective effect in a rodent (Sugen/hypoxia) model of pulmonary hypertension[5].

Whole blood transcriptome profiles offer another approach to interrogating the pathology of pulmonary vascular disease. I initiated RNA sequencing of whole blood samples from patients with PAH (n=359) from the UK National cohort and matched controls (n=72) (PMID: 32352834). With Dr Chris Rhodes, differences in expression levels of 507 genes between PAH and controls were noted and used to construct a model that score disease severity and predict long-term survival. Using two-sample Mendelian randomisation analysis in collaboration with international colleagues from the genome-wide association study, an eQTL for *SMAD5* was associated with risk of developing PAH; lower expression was associated with increased risk. *SMAD5* encodes an intracellular transcriptional modulator that is activated by ligand binding of bone morphogenetic protein receptor 2, the most common genetic risk factor in heritable PAH, giving biological plausibility to the observation. Interestingly, *SMAD5* also controls levels of the master iron regulator, hepcidin, which I had previously reported to be elevated and likely drive iron deficiency in PAH and is associated with a poor clinical outcome (PMID: 21737024). One conclusion from the study is that therapeutic strategies that restore normal SMAD5 function may have broader benefit beyond patients with BMPR2 mutations. Another consideration is whether patients with genetically lower SMAD5 levels, or simply with the variant associated with lower SMAD5 levels, might show a differential response to novel therapeutics targeting the BMPR2 signalling pathway.

The transcriptomic data lend themselves to deeper interrogation for molecular clusters to better understand the heterogenous population left behind once known causes of pulmonary hypertension have been excluded. Unsupervised machine learning of these data, led by Allan Lawrie and Dennis Wang from Sheffield, identified 3 major patient subgroups associated with poor, moderate and good prognosis (PMID: 34876579). Significantly, the endophenotype associated with poor prognosis was in part defined by the *C/C* variant of *HLA-DPA1/DPB1* identified in the genome-wide association study.

**Imaging**

Advanced imaging affords not only deep phenotyping of patients but also the potential to understand mechanisms of response to drugs. Given the difficulty of obtaining lung tissue from patients, I explored the use of positron emission tomography (PET) as a tool for following the response of pulmonary vascular disease to treatment; specifically, using 18F-fluorodeoxyglucose (18FDG) to follow proliferation and inflammation in the diseased lung and develop a bridging biomarker to follow pulmonary vascular remodelling. A proof-of-concept study in a rodent model showed that lung 18FDG

uptake was increased in the disease model and reduced with treatment (PMID: 23900048). I then used this technique in a study of a novel drug from PAH, dichloroacetate, and showed lung parenchymal 18FDG uptake was increased in PAH and reduced in patients that showed signs of improvement on the drug (PMID: 29070699). The dynamic range of the 18FDG signal was small (i.e. in many patients, 18FDG uptake was just above background). Coupled with the need for PET, this particular tracer is not likely to be adopted for experimental medicine studies, but it illustrated the principle of using PET to follow drug response.

Cardiac magnetic resonance (CMR) is more accessible and has become the 'gold standard' for investigating cardiac structure and function. I was an early adopter of CMR and used this in two investigator-led studies to follow the response of the right heart to treatment. In the Sildenafil versus Endothelial Receptor Antagonist in Pulmonary Hypertension (SERAPH) study I reported that sildenafil had a beneficial action in reducing right ventricular hypertrophy (PMID: 20460548). In the Simvastatin in Pulmonary Hypertension (SiPHT) study I reported that high-dose simvastatin, unlike in multiple rodent studies, has no long-term beneficial effect on cardiac mass or function (PMID: 20460548).

CMR data contain more information that the simple metrics usually used to define cardiac function, such as ejection fraction and mass. Collaborating with Declan O'Regan at Imperial College and two co-supervised PhD students, Tim Dawes and Mark Attard, to mine data from over 250 patients in TRIPHIC, we were able to show that 3-dimensional cardiac motion analysis and fractal analysis of right ventricular trabeculae, by taking into account how the heart as a whole behaves in response to pulmonary hypertension, improves survival prediction (PMID: 28092203, PMID: 29869959, PMID: 30535300, PMID: 30801055).

**Summary and future directions**

The application of the tools of experimental medicine, in particular, high-throughput 'omic platforms, advanced imaging and bioinformatics, to patients with clinically identified pulmonary hypertension has challenged the conventional clinical pathways for diagnosing and managing this condition. My research has been at the forefront of this, resulting in a succession of publications that are recognised internationally and demonstrate leadership in the field. The clinical utility of the research has been the identification of novel drug targets and potential biomarkers for risk stratification and monitoring response to treatment. Beyond this is the promise of a greater overhaul of the manner in which pulmonary hypertension is categorised and treated, as outlined in a recent review (PMID: 32201940) and workshop (PMID: 33888254).

Rather than attempting to fit each patient into a clinical category based on haemodynamic measurements and the presence or absence of co-existing disease, deep-molecular phenotyping together with advanced imaging offers the future prospect of a new taxonomy of pulmonary hypertension that defines patient clusters coupled to druggable targets. The approach to treating pulmonary hypertension would then become more akin to that increasingly adopted by oncologists, where treatment is mechanism-based and selected based on the most active aberrant signalling pathway, irrespective of clinical nomenclature (PMID: 33541614).

**Additional references cited**

1. Primary pulmonary hypertension: report on a WHO meeting, Geneva, 15-17 October 1973 / edited by Shuichi Hatano and Toma Strasser (https://apps.who.int/iris/handle/10665/39094)
2. Simonneau G et al Haemodynamic definitions and updated clinical classification of pulmonary hypertension European Respiratory Journal 2019 53: 1801913; https://doi.org/10.1183/13993003.01913-2018

3. Rich S et al. Executive summary from the World Symposium on Primary Pulmonary Hypertension (Evian, France, September 6–10, 1998). http://www.wsphassociation.org/wp-content/uploads/2019/04/Primary-Pulmonary-Hypertension-Evian-1998.pdf
4. Fishman AP. Primary pulmonary arterial hypertension: A look back Journal of the American College of Cardiology (2004) 43: Supplement 12 S2 to S4. https://doi.org/10.1016/j.jacc.2004.03.019
5. Russomanno G et al. miR-150-PTPMT1-cardiolipin signaling in pulmonary arterial hypertension. Molecular Therapy - Nucleic Acids (2020) 23:142-153.
DOI:https://doi.org/10.1016/j.omtn.2020.10.042

**STATEMENT ON CONJOINT WORK**

All the work submitted has been a collaborative effort. I have led or co-led all the studies included in the publication list, having provided and/or shaped the original idea, developed the collaborations, raised the funding as principal applicant or co-applicant, provided supervision, and engaged in data analysis and interpretation. I have referenced key collaborators in the narrative overview and here expand on the background.

**Clinical studies**
The TRIPHIC database at Imperial College relied on the provision of clinical data by the clinical team, led by Simon Gibbs until his retirement and then Luke Howard. The governance of the database was co-ordinated through John Wharton at Imperial College.

**Proteomics and metabolomics**
The proteomic lung mass spectrometry was supervised Robert Edwards and the lipidomic assay by Jeremy Nicholson at Imperial College. The high-throughput plasma proteomic and metabolomic assays were commercial offerings, from SomaLogic and Metabolon respectively. Chris Rhodes and latterly Lars Harbaum from Imperial, provided invaluable bioinformatic expertise.

**Genetics**
The whole genome sequencing of PAH patients and the interpretation of rare variants was led by Nick Morrell and Stefan Graf at Cambridge. Inga Prokopenko at Imperial College provided her expertise and together with Chris Rhodes conducted the common variant analysis. The Kyrgyz expeditions were co-ordinated by the late Almaz Aldashev from Bishkek. The rat genome studies were conducted with Lan Zhao and Tim Aitman at Imperial College, who led on the phenotyping and comparative genomics respectively.

**miRNA and transcriptomics**
The miRNA assays were conducted by Stefanie Dimmeler in Frankfurt and interpreted with bioinformatic expertise from Chris Rhodes. The whole blood RNAseq was conducted by Imperial Biomedical Research Genomics Facility and data interpretation was again with Chris Rhodes. The cluster analysis was overseen by Allan Lawrie and Dennis Wang from Sheffield.

**Imaging**
The PET studies were conducted at Imperial College using local expertise to derive the data. Routine clinical imaging was provided by local hospital radiology departments. The advanced cardiac magnetic resonance studies and data analysis were led by Declan O'Regan at Imperial College.