




Novel techniques for advancing our understanding of pulmonary arterial hypertension

Emily M. Beck¹, Nathan D. Hatton¹ and John J. Ryan²

Affiliations: ¹Division of Pulmonary Medicine, Dept of Medicine, University of Utah, Salt Lake City, UT, USA. ²Division of Cardiovascular Medicine, Dept of Medicine, University of Utah, Salt Lake City, UT, USA.

Correspondence: John J. Ryan, University of Utah Health Science Center, 30 North 1900 East, Room 4A100, Salt Lake City, UT 84132, USA. E-mail: john.ryan@hsc.utah.edu

 @ERSpublications
New, sophisticated tools are being used to advance our knowledge of PAH and facilitate individualised treatment based on phenotypes <http://ow.ly/qP6h30ocllp>

Cite this article as: Beck EM, Hatton ND, Ryan JJ. Novel techniques for advancing our understanding of pulmonary arterial hypertension. *Eur Respir J* 2019; 53: 1900556 [<https://doi.org/10.1183/13993003.00556-2019>].

Pulmonary arterial hypertension (PAH) is a rare disease of the pulmonary vasculature characterised by progressive remodelling and narrowing of the pulmonary arteries due to pathological abnormalities of fibroblasts, inflammatory cells, endothelial cells and smooth muscle cells [1]. Decades of clinical, translational and bench research have advanced our understanding of the underlying mechanisms contributing to the development of PAH. However, despite considerable heterogeneity of this disease, current diagnostic and treatment guidelines remain largely standardised. Aside from cohorts with long-term response to calcium channel blockers, individualised treatment options targeted at the various phenotypes of PAH have not been identified. Ongoing research efforts aim to discover therapies to target disease mechanisms, including genetic and epigenetic mutations, altered metabolism, hormone signalling, inflammatory pathways, and oxidative stress [2].

It is in this background that CORNET *et al.* [3] thoughtfully sought a novel approach to further our understanding of the pathophysiology of PAH. Nine incident cases of PAH were noted in the French Pulmonary Hypertension registry between 2006 and 2010 in patients receiving dasatinib, a potent second-generation protein kinase inhibitor (PKI), used for the treatment of chronic myelogenous leukaemia. These cases were observed 8 to 48 months after initial exposure to dasatinib and characterised by moderate-to-severe PAH with functional and haemodynamic impairments. Improvements, but not complete recovery, were observed within 4 months of dasatinib discontinuation [4]. In the past 3 years there have been numerous other reports of incident PAH or worsening of existing PAH associated with other PKIs, including the Src-family of PKIs (namely bosutinib and ponatinib) and the JAK1/2 PKI ruxolitinib [5, 6].

In their report published in this issue of the *European Respiratory Journal*, CORNET *et al.* [3] used the World Health Organization's pharmacovigilance database, Vigibase, to identify 22 PKIs associated with at least one case of PAH. Pharmacovigilance is the practice of extracting adverse drug events (ADEs) from spontaneous reporting systems to identify previously unestablished post-market ADEs. Vigibase has been collecting data from individual case safety reports since 1968 and contains ~16 million reports of suspected ADEs from more than 150 countries.

Received: March 19 2019 | Accepted: March 20 2019

Copyright ©ERS 2019

While there is no gold standard data-mining method for spontaneous reporting systems, disproportionality analysis is a commonly applied technique to calculate observed-to-expected ratios of potential ADEs [7]. In this study, five of the 22 PKIs were associated with a positive disproportionality signal: dasatinib, bosutinib, ponatinib, ruxolitinib and nilotinib. To identify pharmacodynamic properties, a systematic literature review was performed and yielded affinity data for 16 of the 22 PKIs. The affinity data was subsequently correlated with 35 protein kinases (PKs) involved in PAH pathophysiology. Five PKI targets correlated with disproportionality signals: four Src kinases (c-src, c-yes, Lck, Lyn) and TEC. Finally, cluster analysis was performed on the identified PKIs with the PK affinity data to characterise the affinity profile of each drug. Interestingly, the affinity profile of dasatinib differed from that of the other drugs. These findings are consistent with previously published mechanistic data and case reports of PKI-associated PAH and support the potential pathogenic role of Src PKs [8, 9]. Due to the rare incidence of PKI-associated PAH, it is likely that unelucidated mechanisms such as genetic, epigenetic, inflammatory or environmental factors contribute to the development of disease. Nevertheless, the findings in this paper highlight an exciting breakthrough into possible mechanisms underlying PKI-associated PAH.

Perhaps even more exciting is the novel approach these investigators took in analysing existing resources, in this case Vigibase, to advance our understanding of PAH as a heterogeneous disease. This paper complements increasingly sophisticated approaches being used by PAH researchers to phenotype and identify personalised targets in PAH (figure 1). Examples include the multicentre Pulmonary Vascular Disease through Phenomics (PVDOMICS) study, currently underway, which aims to utilise deep phenotyping with a comprehensive “omic” analysis to discover molecular-based subtypes of pulmonary vascular disease that will allow reclassification, development of targeted treatment options, and individualised management of PAH based on phenotypic characteristics [10].

Another example is the work of HEMNES *et al.* [11] who have described transcriptional and genetic differences that characterise vasodilator-responsive PAH (VR-PAH) from non-responsive PAH (VN-PAH). Transcriptome analysis was performed on peripheral blood leukocytes which identified 13 of the 25 interrogated genes to be differentially expressed in VR-PAH *versus* VN-PAH, thus serving as a proof-of-concept study that gene expression differences can be detected in the blood and used to predict vaso-responsiveness. Whole exome sequencing was subsequently performed on 36 PAH patients, which was mapped to genome-wide association study data to reveal multiple rare genetic variants that converge on key biological pathways implicated in PAH. Compared to VN-PAH, an enrichment of genes involved in vascular smooth muscle contraction was seen in VR-PAH, suggesting a genetic predisposition to vaso-responsiveness [12]. Whole exome sequencing was also performed on a paediatric population of mutation-negative PAH cases by N. Zhu and co-workers, who found that *de novo* variants in novel genes may explain ~19% of paediatric-onset idiopathic PAH cases [13, 14].

With the nearly universal digitalisation of medical records, development of electronic databases, and the expanding field of bioinformatics and genomic research, there is an increasing ability to integrate individual health metrics and molecular data with biorepositories and clinical pharmacology research [15]. Artificial intelligence (AI), which uses techniques such as machine learning (ML) and natural language processing (NLP), is being utilised in precision medicine to mimic human thought processes to solve complicated problems, extract, analyse, and interpret large amounts of data, and transform unstructured text into usable data [16]. Recently, SWEATT *et al.* [17] applied ML to analyse the proteomic immune profile of peripheral blood from PAH patients and found four distinct immune clusters demonstrating different clinical risk profiles and outcomes independent of PAH subtype, patient demographics, comorbidities or background therapies. Using ML tools and Bayesian statistics, efforts are ongoing to reanalyse REVEAL registry data to improve prediction accuracy of 1-year survival in PAH [18]. ML has also been used to create computation-derived metrics in magnetic resonance imaging (MRI) to improve the diagnostic accuracy of MRI in PH and predict survival based on patterns of right ventricular systolic motion [19, 20]. Bayer and Merck received an FDA Breakthrough Device Designation in 2018 for the development of AI software to identify pattern recognition in chronic thromboembolic pulmonary hypertension (CTEPH). This approach utilises deep learning methodology to integrate cardiac, perfusion and pulmonary vessel information with clinical history to support radiologists in identifying CTEPH [21]. Computer science principles of AI, ML and NLP are increasingly being applied to databases, including in the field of pharmacovigilance, thereby increasing the relevance of the work presented here.

Several challenges limit the use of pharmacovigilance databases in real-time research, including biased, incomplete or underreporting of events and a significant delay between reporting and confirmation of ADEs. These challenges can be overcome by incorporating NLP algorithms into electronic health record (EHR) databases [7]. For example, by extracting outcome information from unstructured text contained in the EHR and combining this with the FDA Adverse Event Reporting System (a US-based spontaneous reporting system) dataset, WANG *et al.* [22] recently demonstrated the feasibility of using NLP to capture

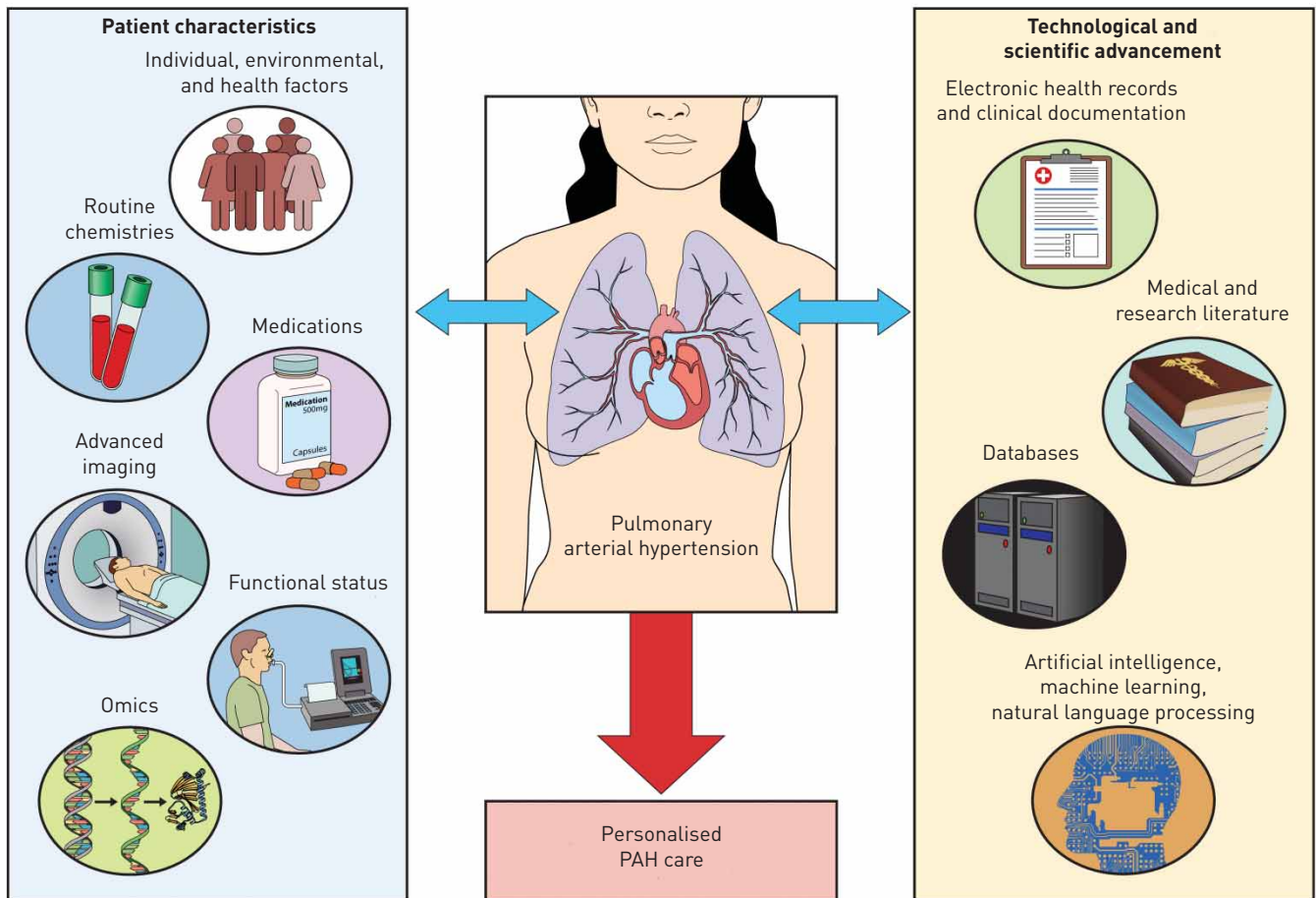


FIGURE 1 Sophisticated approaches being used to phenotype and identify personalised targets in pulmonary arterial hypertension (PAH).

ADEs not previously identified in structured EHR data. Application of NLP algorithms to other datasets, including internet-based reporting systems, and published literature, has the potential for real-time identification of ADEs that would otherwise go unanalysed [23]. The power of applying NLP and pharmacovigilance monitoring to the EHR is strengthened when genetic information is incorporated, as this may reveal individual genetic polymorphism that alter drug metabolism or predispose to ADEs [24].

Linking EHR data with genomics through AI, ML and NLP is a valuable resource with immense potential that could be expanded upon in PAH research to further elucidate disease phenotypes. Despite the obvious challenges of a patient population that does not likely reflect the entire community, inconsistencies of data collection within and between EHRs, absence of standardised phenotype ontologies, and ethical and privacy concerns, there are a number of advantages to using data extracted from EHRs coupled to biobanks. Traditional studies are time-intensive and expensive, with strict inclusion criteria and relatively short follow-up times. EHR-driven genetic research allows for analysis of large and diverse cohorts with robust clinical data gathered in real-world settings over years. Researchers can rapidly and inexpensively generate large datasets with the possibility of direct clinical applications that guide clinicians to improve patient care in real time [25, 26]. Currently, the Electronic Medical Records and Genomics (eMERGE) Network is working to “develop, disseminate, and apply approaches to research that combine biorepositories with electronic medical record systems for genomic discovery and genomic medicine implementation research” [27].

The paper by CORNET *et al.* [3] demonstrates that novel research methods and resources can be used to further our understanding of the pathophysiology of PAH. Armed with new, sophisticated tools, the future of PAH diagnosis, treatment and prognosis lies with a more precise understanding of individual disease phenotypes that will direct preventative strategies, individualise treatments, and improve prognostication.

Conflict of interest: None declared.

References

- 1 Thenappan T, Ormiston ML, Ryan JJ, *et al.* Pulmonary arterial hypertension: pathogenesis and clinical management. *BMJ* 2018; 360: j5492.
- 2 Sitbon O, Gomberg-Maitland M, Granton J, *et al.* Clinical trial design and new therapies for pulmonary arterial hypertension. *Eur Respir J* 2019; 53: 1801908.
- 3 Cornet L, Khouri C, Roustit M, *et al.* Pulmonary arterial hypertension associated with protein kinase inhibitors: a pharmacovigilance–pharmacodynamic study. *Eur Respir J* 2019; 53: 1802472.
- 4 Montani D, Bergot E, Gunther S, *et al.* Pulmonary arterial hypertension in patients treated by dasatinib. *Circulation* 2012; 125: 2128–2137.
- 5 Seegobin K, Babbar A, Ferreira J, *et al.* A case of worsening pulmonary arterial hypertension and pleural effusions by bosutinib after prior treatment with dasatinib. *Pulm Circ* 2017; 7: 808–812.
- 6 Low AT, Howard L, Harrison C, *et al.* Pulmonary arterial hypertension exacerbated by ruxolitinib. *Haematologica* 2015; 100: e244–e245.
- 7 Wang X, Hripcsak G, Markatou M, *et al.* Active computerized pharmacovigilance using natural language processing, statistics, and electronic health records: a feasibility study. *J Am Med Inform Assoc* 2009; 16: 328–337.
- 8 Weatherald J, Chaumais MC, Savale L, *et al.* Long-term outcomes of dasatinib-induced pulmonary arterial hypertension: a population-based study. *Eur Respir J* 2017; 50: 1700217.
- 9 Nagaraj C, Tang B, Balint Z, *et al.* Src tyrosine kinase is crucial for potassium channel function in human pulmonary arteries. *Eur Respir J* 2013; 41: 85–95.
- 10 Hemnes AR, Beck GJ, Newman JH, *et al.* PVDOMICS: a multi-center study to improve understanding of pulmonary vascular disease through phenomics. *Circ Res* 2017; 121: 1136–1139.
- 11 Hemnes AR, Trammell AW, Archer SL, *et al.* Peripheral blood signature of vasodilator-responsive pulmonary arterial hypertension. *Circulation* 2015; 131: 401–409.
- 12 Hemnes AR, Zhao M, West J, *et al.* Critical genomic networks and vasoreactive variants in idiopathic pulmonary arterial hypertension. *Am J Respir Crit Care Med* 2016; 194: 464–475.
- 13 Zhu N, Gonzaga-Jauregui C, Welch CL, *et al.* Exome sequencing in children with pulmonary arterial hypertension demonstrates differences compared with adults. *Circ Genom Precis Med* 2018; 11: e001887.
- 14 Cirulis MM, Ryan JJ. Who’s who of pulmonary hypertension: redefining classification to advance precision care. *Circ Genom Precis Med* 2018; 11: e002116.
- 15 Jensen PB, Jensen LJ, Brunak S. Mining electronic health records: towards better research applications and clinical care. *Nat Rev Genet* 2012; 13: 395–405.
- 16 Krittanawong C, Zhang H, Wang Z, *et al.* Artificial intelligence in precision cardiovascular medicine. *J Am Coll Cardiol* 2017; 69: 2657–2664.
- 17 Sweatt AJ, Hedlin HK, Balasubramanian V, *et al.* Discovery of distinct immune phenotypes using machine learning in pulmonary arterial hypertension. *Circ Res* 2019; 124: 904–919.
- 18 Kanwar M, Lohmueller L, Correa P, *et al.* (501) Application of Bayesian model to predict outcomes in pulmonary arterial hypertension. *J Heart Lung Transplant* 2018; 37: Suppl., S207.
- 19 Lungu A, Swift AJ, Capener D, *et al.* Diagnosis of pulmonary hypertension from magnetic resonance imaging-based computational models and decision tree analysis. *Pulm Circ* 2016; 6: 181–190.
- 20 Dawes TJW, de Marvao A, Shi W, *et al.* Machine learning of three-dimensional right ventricular motion enables outcome prediction in pulmonary hypertension: a cardiac MR imaging study. *Radiology* 2017; 283: 381–390.
- 21 May J. FDA Grants Breakthrough Device Designation for CTEPH Pattern Recognition Artificial Intelligence Software from Bayer and Merck [press release]. Bayer, 2018. Date last accessed: March 8, 2019.
- 22 Wang L, Rastegar-Mojarad M, Ji Z, *et al.* Detecting pharmacovigilance signals combining electronic medical records with spontaneous reports: a case study of conventional disease-modifying antirheumatic drugs for rheumatoid arthritis. *Front Pharmacol* 2018; 9: 875.
- 23 Wong A, Plasek JM, Montecalvo SP, *et al.* Natural language processing and its implications for the future of medication safety: a narrative review of recent advances and challenges. *Pharmacotherapy* 2018; 38: 822–841.
- 24 Clark DW, Donnelly E, Coulter DM, *et al.* Linking pharmacovigilance with pharmacogenetics. *Drug Saf* 2004; 27: 1171–1184.
- 25 Hall JL, Ryan JJ, Bray BE, *et al.* Merging electronic health record data and genomics for cardiovascular research: a science advisory from the American Heart Association. *Circ Cardiovasc Genet* 2016; 9: 193–202.
- 26 Krishnamoorthy P, Gupta D, Chatterjee S, *et al.* A review of the role of electronic health record in genomic research. *J Cardiovasc Transl Res* 2014; 7: 692–700.
- 27 Gottesman O, Kuivaniemi H, Tromp G, *et al.* The Electronic Medical Records and Genomics (eMERGE) Network: past, present, and future. *Genet Med* 2013; 15: 761–771.