Time for precision medicine in systemic sclerosis-associated pulmonary arterial hypertension

David Launay1,2,3, Sébastien Sanges1,2,3 and Vincent Sobanski1,2,3,4

Affiliations: 1Univ. Lille, U1286 – INFINITE – Institute for Translational Research in Inflammation, Lille, France. 2Inserm, Lille, France. 3CHU Lille, Service de Médecine Interne et Immunologie Clinique, Centre de référence des maladies autoimmunes systémiques rares du Nord et Nord-Ouest de France (CeRAINO), Lille, France. 4Institut Universitaire de France (IUF), Lille, France.

Correspondence: David Launay, CHU Lille, Service de Médecine Interne et Immunologie Clinique, F-59000 Lille, France. E-mail: david.launay@univ-lille.fr

Integration of omics with clinical data through machine learning is the future of personalised and precision medicine in systemic sclerosis-associated pulmonary arterial hypertension

https://bit.ly/3jsyBoV

Cite this article as: Launay D, Sanges S, Sobanski V. Time for precision medicine in systemic sclerosis-associated pulmonary arterial hypertension. Eur Respir J 2021; 57: 2100205 [https://doi.org/10.1183/13993003.00205-2021].

Pulmonary arterial hypertension (PAH) is both a frequent and dismal complication of systemic sclerosis (SSc) [1]. A recent meta-analysis established that the pooled prevalence of PAH in SSc was 6.4% (95% CI 5–8.3%) and overall incidence 18.2 cases per 1000 person-years (95% CI 12–27.4) [2]. PAH is also one of the first causes of death in SSc, with a standardised mortality ratio of 5.27 (95% CI 2.98–9.31) [3] and a median survival of 3 years [1], which has not improved meaningfully for many years. Very consistently, prognosis of SSc-associated PAH (SSc-PAH) is worse than idiopathic PAH [4]. There are many explanations for this poor prognosis [5], including the severity and multisystemic condition of SSc itself [6, 7], heart involvement, frequent pulmonary veno-occlusive disease and late diagnosis (72–85% of patients have New York Heart Association (NYHA) functional class III or IV symptoms at diagnosis [8], leading to a worse response to treatment and poor outcome [9].

Improving SSc-PAH prognosis is a major challenge and undoubtedly requires diagnosis and treatment of patients as soon as possible. Indeed, Humbert et al. [10] have shown that an early diagnosis was associated with a better outcome and some clinical trials have shown that treating patients with less severe PAH was also associated with a better survival [11]. The main question remains of how to improve diagnosis of patients at an earlier stage of PAH. Systematic screening is most often cited as the best tool to capture patients with less severe PAH, opening hope for better prognosis with early intervention. Thus, screening for PAH in SSc is uniformly advised in guidelines [12]. A major issue is how to screen, and many efforts have been recently made to find the best way to screen SSc patients. A recent review by Weatherald et al. [8] has described the current approaches for screening of PAH in SSc patients. Transthoracic echocardiography, pulmonary functional tests and cardiac biomarkers (N-terminal pro-brain natriuretic peptide, NT-proBNP) are the most used and cited tools in guidelines. Recently, composite screening algorithms have been assessed associating clinical parameters as well as echocardiography, pulmonary functional tests and biomarker data. The DETECT algorithm has been validated in a study including SSc patients with a disease duration >3 years and a diffusing capacity of the lung for carbon monoxide (Dlco)
<60% predicted, which enriched the study population with patients at higher likelihood of having PAH [13]. Thanks to this algorithm, which was a major breakthrough, the sensitivity was high (96%) and most patients had mild or early disease as evidenced by 64% being in NYHA functional class I or II with an average mean pulmonary artery pressure (mPAP) of 32.5±8.3 mmHg and mean pulmonary vascular resistance (PVR) of 4.6±2.8 Wood units. Yet, DETECT has three caveats: 1) it is validated and recommended in selected SSc patients (disease duration >3 years and \( D_{LCO} <60\% \)); 2) the number of patients sent to right heart catheterisation (RHC) is high (between 40 and 60%); and 3) it has not been validated for the new definition of PAH with mPAP >20 mmHg. Besides DETECT, other composite algorithms exist, including the Australian Scleroderma Interest Group (ASIG), based on \( D_{LCO} \), forced vital capacity %/\( D_{LCO} \) % and NT-pro BNP values [14], and European Society of Cardiology/European Respiratory Society 2015 recommendations for echocardiography [12]. Thus, recent 2018 guidelines recommend that for SSc spectrum patients with uncorrected \( D_{LCO} <80\% \) of predicted, annual screening should be considered with these tools. For those with \( D_{LCO} \geq80\% \) pred, screening with echocardiography may be considered.

Yet, there is considerable room for improvement in PAH screening. Indeed, the current situation is not satisfactory because of the lack of a universal and validated screening method for the whole population of SSc patients, and the desperate need of an algorithm which would ideally be cost-effective, especially by limiting the screening tools and the number of patients unnecessarily sent to RHC. Among promising future tools, serum biomarkers are of special interest, as they are not invasive and could reflect endothelial dysfunction, inflammation, autoimmunity, epigenetic changes, extracellular matrix deposition and metabolite changes, which have recently been extensively reviewed by Odler et al. [15]. Usually, biomarkers are chosen \textit{a priori} based on pathophysiological considerations. For example, recently CCL21 has been assessed as a potential marker of PAH and proved to be interesting [16]. Recent advances in biomarker discovery are 1) to use without any \textit{a priori} tools like omics or multiplexed approaches; and 2) to integrate them in non-supervised machine learning methods combining all data from clinical characteristics to biomarkers to offer a true personalised decision to perform RHC, with the double aim of limiting the risks of not diagnosing PAH and/or performing unnecessary RHC.

In this context, the study performed by Bauer et al. [17] and published in the current issue of the \textit{European Respiratory Journal} is timely and important. They used serum samples from SSc patients included in the DETECT study as a discovery cohort and serum samples from the University of Sheffield, UK as a confirmatory cohort. 313 analytes were assessed using a multiplexed immunoassay aiming to define a proteomic signature to classify patients with SSc into those with and without PAH. Eight proteins were common in the top 20 variables of importance in both cohorts. Indeed, collagen IV, endostatin, IGFBP-2, IGFBP-7, MMP-2, neuropilin-1, NT-proBNP and RAGE discriminated SSc patients with PAH from those without pulmonary hypertension in the DETECT discovery cohort (average area under the receiver operating characteristic curve values of 0.741, 65.1% sensitivity/69.0% specificity) and was reproduced in the Sheffield cohort (81.1% accuracy, 77.3% sensitivity/86.5% specificity). Interestingly, correlations between these proteins and clinical variables related to PAH were generally weak. The best correlation was observed between NT-proBNP and pulmonary vascular resistances. Most proteins identified in this study as possible biomarkers of PAH in SSc have been previously found to play significant roles in the pathophysiology of PAH. Notably, it was not mandatory in the selection process that proteins should have a pathophysiological role. Thus, without any \textit{a priori} choice, interesting biomarkers were spontaneously associated with a significant pathogenic role in PAH.

These interesting results lead to some observations and discussion. First, their performance to classify patients with or without PAH was different between the discovery cohort and the confirmatory cohort. This confirms that PAH is indeed a heterogeneous complication with patients of different severity [5]. More validation studies including a larger number of patients are thus mandatory. Second, the performance of this eight-protein panel was superior to NT-proBNP alone, suggesting that it is adding a value to this important biomarker and that it could capture patients with less severe PAH before cardiac stress becomes dominant. When compared to the performance of the DETECT algorithm (sensitivity: 96% and specificity: 48%), we observe that the eight-protein panel is less sensitive but more specific. This is interesting as the strategy to use this panel has yet to be defined. One could hypothesise that the first step of DETECT (determining whose patients should undergo an echocardiography) could remain the same, while adding the proteomic signature to the second step could enhance the specificity and thus lower the number of unnecessary RHC procedures carried out. Third, previous studies have also performed without an \textit{a priori} approach using a different technique (SOMAscan) to analyse the proteomic signature. Rice et al. [18] identified midkine and follistatin-like 3 as SSc-PAH biomarkers among >1000 candidate proteins. Unfortunately, we cannot compare with the present study as these proteins were not analysed in the study by Bauer et al. [17]. Yet, this shows that future studies should use a high-throughput omics

\[ \text{https://doi.org/10.1183/13993003.00205-2021} \]
approach allowing assessment of the largest number of analytes. Fourth, this study used the previous definition of PAH with a cut-off of 25 mmHg for mPAP. This definition has changed and the cut-off is now 20 mmHg. The authors have re-run their analysis with this new definition and the results were similar, with only two proteins dropping from the panel. Yet, this shows that future studies should take the new definition of PAH, probably also analysing the results with PVR above 3 Wood units and between 2 and 3 as recently suggested [19]. Fifth, some recent studies have also assessed transcriptomic approaches in the biomarker discovery journey of PAH in SSc. ZHENG et al. [20] found that IFIT2, IFIT3, RSAD2 and PARP14 gene expressions in peripheral blood mononuclear cells may serve as potential biomarkers in SSc-PAH. This opens the field of integrating not only proteomic signature, such as in the study by BAUER et al. [17], but broader omics in the future to pave the path to precision medicine in SSc-PAH. One question remains over whether these new biomarkers tools could be integrated in the routine care. In conclusion, the study of BAUER et al. [17] is opening new avenues in PAH screening in SSc, suggesting that our good old algorithms could soon be outdated by computer-assisted algorithms. These could not only include clinical and echocardiography data, but also broad omics signatures allowing a very precise decision of which patients should undergo invasive RHC, which remains the gold standard for the diagnosis and classification of pulmonary hypertension. Future studies should also include SSC patients with different mechanisms of pulmonary hypertension like diastolic dysfunction, interstitial lung disease and pulmonary veno-occlusive disease.

Conflict of interest: D. Launay reports grants and personal fees from Takeda and Octapharma, personal fees from Biocrypt, grants from CSL Behring, outside the submitted work. S. Sanges reports grants from Shire, Sanofi Genzyme and SOBI, outside the submitted work. V. Sobanski reports personal fees from Boehringer Ingelheim, grants and personal fees from Grifols, grants from GSK, Octapharma, Pfizer and Shire, outside the submitted work.

References