Assessing risk in pulmonary arterial hypertension: what we know, what we don't

Raymond L. Benza[1], Harrison W. Farber[2], Mona Selej[3,4] and Mardi Gomberg-Maitland[5]

Affiliations: [1]Cardiovascular Medicine, Allegheny General Hospital, Pittsburgh, PA, USA. [2]Pulmonary and Critical Care, Boston Medical Center, Boston, MA, USA. [3]University of California San Francisco, Dept of Medicine, Cardiology, Pulmonary Hypertension Clinic, San Francisco, CA, USA. [4]Actelion Pharmaceuticals US, Inc., South San Francisco, CA, USA. [5]Cardiology, Heart and Vascular Institute, Falls Church, VA, USA.

Correspondence: Raymond L. Benza, Cardiovascular Medicine, Allegheny General Hospital, 320 East North Avenue, Pittsburgh, PA 15212, USA. Email: RBENZA@wpahs.org

@ERSpublications
Risk stratification in PAH is best served by joint use of the ERS/ESC risk table and REVEAL calculator risk tools http://ow.ly/ySrB30dKq9r

Cite this article as: Benza RL, Farber HW, Selej M, *et al*. Assessing risk in pulmonary arterial hypertension: what we know, what we don't. *Eur Respir J* 2017; 50: 1701353 [https://doi.org/10.1183/13993003.01353-2017].

Despite improvement in short-term survival, pulmonary arterial hypertension (PAH) remains an incurable disease with an unacceptable median survival of 7 years [1]. In the USA, the survival rates for PAH patients awaiting transplant continue to mirror the 2.5-year survival rate in the pretreatment era of the disease [2]. Despite treatment, PAH patients continue to experience disease progression and increased rates of hospitalisations due to right heart failure. Importantly, right heart failure hospitalisations in PAH occur at rates that are comparable to patients with left heart disease, particularly in those aged greater than 65 [3] More so, morbid events in PAH, notably hospitalisations, herald disease progression and early mortality [4, 5]. Hence, along with advancing PAH treatment options, appropriate and accurate risk prediction is essential to halt disease progression and make individualised treatment decisions.

In a progressive disease like PAH, early and accurate risk prediction allows for the identification of patients who are more likely to progress rapidly, "rapid progressors". Risk stratification is especially important in settings where clinical PAH experience is not available and could facilitate early referral to a PAH centre. A risk stratification algorithm could also offer a more individualised treatment strategy for PAH patients; by identifying risk stratum, guiding clinical decision making and informing treatment options and goals. Risk prediction modelling can help physicians allocate treatment resources in settings where they are scarce. They can also be used to inform patients of their prognosis thereby allowing them to make informed decisions about treatment options. If widely adopted, appropriate risk prediction provides an opportunity to learn about various risk phenotypes in PAH, enhance consistency of treatment approaches across practitioners and assist in the timely referral for lung transplantation. Lastly, risk model-derived equations can enhance clinical study design both by selecting the appropriate study cohort and serving as a study end-point.

Statistical models are often used to predict the probability that an individual with a given set of risk factors will experience a health outcome, usually termed an "event" [6]. In PAH, it is widely agreed that a

Received: July 06 2017 | Accepted after revision: July 17 2017

Conflict of interest: None declared.

multivariate risk prediction model that is validated serially represents the best predictor of morbidity and mortality, especially compared to measures assessed individually or at a single time point. In building such model, the challenge lies in choosing the parameters that best depict risk and their corresponding risk cut-offs/categories. In acute coronary syndrome and stroke risk in atrial fibrillation [7, 8], risk prediction models have been widely implemented. These risk assessment tools were informative, statistically sound, validated and practical to use. Similarly, in PAH, the ideal risk assessment tool has to: 1) be easy to use; 2) be applicable, with good predictive ability, at any time in a patient's course of disease, whether that patient is newly or previously diagnosed; 3) be equally predictive whether used at baseline or at follow-up; 4) be applicable to all PAH sub-groups; 5) be informed by the most recent data available even if testing was not contemporaneous; 6) retain utility when some data points or risk parameters are missing; and 7) be dynamic (*i.e.* capable of capturing change over time and changes in overall risk are reflective of changes in prognosis). To be statistically sound, the risk assessment tool must: 1) have good discrimination and calibration); 2) be derived from sizable well-defined cohorts; 3) be composed of risk parameters (and corresponding cut-points) that are evidence-based rather than expert opinion-derived; 4) incorporate "weighting" of the various parameters; and 5) be validated internally and externally.

In recent investigations of risk assessment in PAH from three European (Swedish, COMPERA (80% German) and French) [9–11] registries, an abbreviated version of the risk stratification strategy outlined in the European Society of Cardiology (ESC)/European Respiratory Society (ERS) guideline was evaluated [12]. This strategy aims to stratify patients into low, intermediate and high risk of death or transplant-free survival based on expert opinion-derived risk silos. Overall, the three bodies of work utilised a somewhat similar set of readily available clinical, haemodynamic and echocardiographic variables established in the ESC/ERS guidelines and were exclusive to newly diagnosed PAH patients (majority or entirety of whom were idiopathic, hereditable, or drug-induced). More importantly, in work that collectively includes over 3000 patients from eight European countries and despite variable follow-up, all three investigations successfully and consistently highlighted the favourable prognostic outcome of "low risk" *versus* the other risk categories both at baseline and follow-up.

In the Swedish PAH registry (SPAHR), Kylhammar *et al.* [13] showed that in PAH incident patients, the "low risk" silo at baseline (n=530) and within 1-year follow-up (n=383), confers a survival advantage. Risk assessment variables in this work included functional class (FC), 6-min walk distance (6MWD), NT-proBNP, right atrial (RA) area; mean RA pressure (mRAP), pericardial effusion, cardiac index (CI), and mixed-venous oxygen saturation (S_{vO_2}). Each variable was graded from 1 to 3 where 1 is "Low risk", 2 "Intermediate risk", and 3 "High risk". A patient's mean grade was obtained by dividing the sum of all grades by the number of available variables for each patient.

In COMPERA (n=1588), Hoeper *et al.* [14] showed that the observed mortality rates 1 year after diagnosis were 2.8% in the low risk cohort (n=196), 9.9% in the intermediate risk cohort (n=1116), and 21.2% in the high risk cohort (n=276), respectively. The selected cohort from COMPERA was a treatment-naïve, newly diagnosed PAH population with at least two of the following variables available: World Health Organization (WHO) FC, 6MWD, BNP or NT-proBNP, mRAP, CI and S_{vO_2}. Importantly, Hoeper *et al.* [14] used a similar set of variables (minus mRAP and pericardial effusion) and a similar grading strategy as Kylhammar *et al.* [13] (risk silo was defined by a mean grade obtained by dividing the sum of all grades by the number of available variables).

In the French registry (n=1017), Boucly *et al.* [15] studied incident patients with idiopathic, heritable or drug-induced PAH. The number of low-risk criteria (FC I or II, 6MWD >440 m, mRAP <8 mmHg and CI $\geqslant$2.5 $L \cdot min^{-1} \cdot m^{-2}$) was assessed at diagnosis and at first re-evaluation within 1 year of follow up. Boucly *et al.* [15] concluded that the number of low-risk criteria present at diagnosis and at first re-evaluation (median 4.4 months) discriminated risk of death or lung transplantation.

In the PAH research community, much has been learned from the collective experience of registries. These bodies of work have inspired many discussions on the ideal risk prediction tool, the appropriate cohort to examine risk (incident *versus* prevalent and retrospective *versus* prospective cohorts), the sophisticated statistical concepts surrounding such analyses, and the limitations of using cohort data. Thus, our understanding of left truncated data (and delayed entry modelling), survival and immortal time biases, varying risk hazards (accounting for risk at different times of the disease) and interactions between risk and prognosis have been accelerated [16, 17]. The consensus imparted to the PAH research community from these sessions has been that although a "simple" prognostic model might seem easier to implement, this complex disease requires a sophisticated model that is able to inform both our clinical practice and trial design.

The aforementioned evaluations of the ESC/ERS risk stratification scheme [12] are meritorious in their simplicity of approach, consistency of findings, large number of patients and relative diversity of the

cohorts. However, they do share some notable shortcomings. Methodologically, in an effort to simplify risk prediction, both the expert derived ESC/ERS risk table and the current risk prediction analyses examining its "real world" applicability, share a major limitation. Common to all three analyses, the risk assessment tool employed is based on consensus of expert opinion rather than statistical modelling of validated parameters of risk [12, 13]. In addition, these analyses did not cover the full spectrum of PAH subgroups nor, in some cases, all ages [13]. None examined risk assessment in prevalent cohorts nor included non-modifiable risk factors. Other methodological limitations that are not unique to these analyses include the use of retrospective cohorts (in SPAHR and French registries), loss of patients at follow-up (in COMPERA and SPAHR registries), the unavailability of key parameters of risk (in COMPERA and SPAHR registries) and the degree to which assessed variables were contemporaneous (in COMPERA, SPAHR and French registries). Hence, although these findings reassure us that maintaining or achieving low risk treatment goals in PAH is sound, unsurprisingly, these findings fall short of enriching our knowledge beyond that or changing the way we, as clinicians, practise PAH today. Remarkably, only a minority of patients in each analysis cohort achieved such low risk status at follow-up. The reasons for this are likely myriad and complex and certainly warrant further study.

Common to all three analyses, is the "lack of weighting" of the various prognostic parameters *i.e.* parameters were assigned similar grade and assumed to be equal. In risk assessment, especially in the context of missing variables, weighting variables is paramount. Weighting is defined as assigning a risk factor a value given according to its significance in the overall risk rating. In PAH models, this is achieved by evaluating individual factors (*e.g.* FC) based on their actual hazard ratios (observed risks) and assigning them individual scores accordingly. In future work, the performance of these models as risk predictors needs further testing for calibration and discrimination. Although promising in their consistent findings, for wider use and applicability, these models require prospective validation amongst a wider variety of WHO group I patients to assess utility and require external validation to determine generalisability. Also, although these measures appropriately identified risk at ends of the spectrum (*i.e.* high and low risk), they had poorer discrimination for those "in between". Finally, it is important that a tool has accurate assessment of prognosis at time of diagnosis and within the first year of treatment, as this is the time point when crucial treatment decisions are made and when assessment of risk is most/best calibrated. It is clear that the more variables utilised (COMPERA/SPAHR *versus* French registry) the better the delineation of risk at 1 year. For example, in the French evaluation, there was little segregation of risk within 1 year for either model (with or without haemodynamics or natriuretic peptide levels), except in those patients who did or did not have all four-risk criterion. A similar drawback (clustering of risk) was noted in the original REVEAL risk calculator-based Kaplan–Meier survival estimates [18]. Whether this early clustering of risk is related to study inclusion criterion or not will likely be borne out in future prospective investigations. Indeed, a recent refinement of the REVEAL risk calculator showed much improved delineation of risk across all risk strata particularly within the first year of risk [19].

The ESC/ERS scheme appears easy to use; however, it is more likely that the assumption that patients will neatly fit neatly into one of the three risk categories is an oversimplification of the reality of PAH. Additionally, if we were to utilise the "Rubik cube" approach [20], we are faced with approximately 729 combinations of various risk factors. Thus, to use the full spectrum of the ERS/ESC table results in complex blending of risk. Additionally, because the risk assessment is derived from a multivariable model, it will underestimate the relative risk contribution of factors associated with high risk such as connective tissue disease-associated PAH. Unfortunately, other risk tools like the REVEAL risk calculator, a robust statistical and clinically validated tool, were overlooked in recent guideline statements [9]. Importantly, work from REVEAL and other large and small international center has informed the selection of those PAH prognostic factors (and their corresponding cut points, *i.e.* 6MWD, NT-ProBNP) in the ERS/ESC consensus-based risk assessment scheme. It is possible that the REVEAL calculator's absence from the recent guidelines may be due to perceived shortcomings of the tool; most of which are misconceptions on its ease of use (or lack thereof), "complexity" due to the number of variables, its derivation from a largely previously diagnosed cohort and the use of non-modifiable variables and "invasive" measurements. In regards to the later, an important observation in the French study was that invasive haemodynamics was not necessary to determine a low risk profile. Indeed, the use of a risk predictive tool that does not require routine invasive assessment would be ideal. Similar to the REVEAL calculator, in which the inclusion of invasive measures is neither necessary nor required, it is now safe to say that we have two easy means to calculate low risk profiles in the clinic. It is also worth clarifying that the REVEAL score only utilises 12 variables, not 13 or 19 as mentioned in the French and COMPERA articles, respectively. In fact, only seven are actually required at any particular time to keep the score adequately calibrated. Lastly, Boucly *et al.* [15] suggest that the use of non-modifiable variables may be a disadvantage when calculating risk. We contend, however, that the inclusion of non-modifiable risk factors (*e.g.* age/gender and PAH subtype) are inseparable from risk projections; are known to shape the natural course of the disease; and their use is

widely accepted in many disease areas in medicine [7, 8]. Alas, despite all these aforementioned attributes for both approaches, the current reality is that neither an equation nor a cube will perfectly define an individual's risk assessment.

In sum, despite the progress we have made in understanding risk in PAH and multiple attempts to devise the perfect risk prediction tool we have not yet created it. Every tool has inherent weaknesses and none have been studied prospectively. The ERS/ESC consensus-based risk assessment schemes best illustrate that our clinical acumen is quite good when used to create a low risk clinical profile. To emphasise a recent perspective in the *New England Journal of Medicine* on prediction in medicine, "the practice of medicine is constantly evolving in response to new technology, epidemiology, and social phenomena, we will always be chasing a moving target" [21]. Thus, to predict prognosis accurately, we as a community must use all the tools available to us. So, instead of working in silos developing multiple tools, we should work as a community and join forces. Collaboration to devise a clinically meaningful prognostic algorithm, which can then be investigated in a collective, collegial and prospective manner, should be our goal. Ultimately, we all seek the same end, a risk assessment tool in which change in score by treatment choices is associated with change in outcome. In such a manner, we can move together to improve outcome in this destructive disease.

References

1 Benza RL, Miller DP, Barst RJ, *et al.* An evaluation of long-term survival from time of diagnosis in pulmonary arterial hypertension from the REVEAL Registry. *Chest* 2012; 142: 448–456.
2 Gomberg-Maitland M, Glassner-Kolmin C, Watson S, *et al.* Survival in pulmonary arterial hypertension patients awaiting lung transplantation. *J Heart Lung Transplant* 2013; 32: 1179–1186.
3 George MG, Schieb LJ, Ayala C, *et al.* Pulmonary hypertension surveillance: United States, 2001 to 2010. *Chest* 2014; 146: 476–495.
4 Frantz RP, Schilz RJ, Chakinala MM, *et al.* Hospitalization and survival in patients using epoprostenol for injection in the PROSPECT observational study. *Chest* 2015; 147: 484–494.
5 Frost AE, Badesch DB, Miller DP, *et al.* Evaluation of the predictive value of a clinical worsening definition using 2-year outcomes in patients with pulmonary arterial hypertension: a REVEAL Registry analysis. *Chest* 2013; 144: 1521–1529.
6 Pavlou M, Ambler G, Seaman SR, *et al*. How to develop a more accurate risk prediction model when there are few events. *BMJ* 2015; 351: h3868.
7 Antman EM, Cohen M, Bernink PJ, *et al.* The TIMI risk score for unstable angina/non-ST elevation MI: A method for prognostication and therapeutic decision making. *JAMA* 2000; 284: 835–842.
8 Gage BF, van Walraven C, Pearce L, *et al.* Selecting patients with atrial fibrillation for anticoagulation: stroke risk stratification in patients taking aspirin. *Circulation* 2004; 110: 2287–2292.
9 Olsson KM, Delcroix M, Ghofrani HA, *et al.* Anticoagulation and survival in pulmonary arterial hypertension: results from the Comparative, Prospective Registry of Newly Initiated Therapies for Pulmonary Hypertension (COMPERA). *Circulation* 2014; 129: 57–65.
10 Rådegran G, Kjellström B, Ekmehag B, *et al.* Characteristics and survival of adult Swedish PAH and CTEPH patients 2000-2014. *Scand Cardiovasc J* 2016; 50: 243–250.
11 Humbert M, Sitbon O, Chaouat A, *et al.* Pulmonary arterial hypertension in France: results from a national registry. *Am J Respir Crit Care Med* 2006; 173: 1023–1030.
12 Galiè N, Humbert M, Vachiery J-L, *et al.* 2015 ESC/ERS Guidelines for the diagnosis and treatment of pulmonary hypertension. *Eur Respir J* 2015; 46: 903–975.
13 Kylhammar D, Kjellström B, Hjalmarsson C, *et al.* A comprehensive risk stratification at early follow-up determines prognosis in pulmonary arterial hypertension. *Eur Heart J* 2017; in press [https://doi.org/10.1093/eurheartj/ehx257].
14 Hoeper MM, Kramer T, Pan Z, *et al.* Mortality in pulmonary arterial hypertension: prediction by the 2015 European pulmonary hypertension guidelines risk stratification model. *Eur Respir J* 2017; 50: 170740.
15 Boucly A, Weatherald J, Savale L, *et al.* Risk assessment, prognosis and guideline implementation in pulmonary arterial hypertension. *Eur Respir J* 2017; 50: 170889.
16 Miller DP, Gomberg-Maitland M, Humbert M. Survivor bias and risk assessment. *Eur Respir J* 2012; 40: 530–532.
17 Humbert M, Sitbon O, Chaouat A, *et al.* Survival in patients with idiopathic, familial, and anorexigen-associated pulmonary arterial hypertension in the modern management era. *Circulation* 2010; 122: 156–163.
18 Benza RL, Gomberg-Maitland M, Miller DP, *et al.* The REVEAL Registry risk score calculator in patients newly diagnosed with pulmonary arterial hypertension. *Chest* 2012; 141: 354–362.
19 Benza RL, Elliott CG, Farber HW, *et al.* Updated risk score calculator for pulmonary arterial hypertension patients. *J Heart Lung Transplant* 2017; 36: S19.
20 Weatherald J, Sitbon O, Humbert M. Validation of a risk assessment instrument for pulmonary arterial hypertension. *Eur Heart J* 2017; in press [https://doi.org/10.1093/eurheartj/ehx301].
21 Chen JH, Asch SM. Machine learning and prediction in medicine – beyond the peak of inflated expectations. *N Engl J Med* 2017; 376: 2507–2509.