Interactions between rheumatologists and cardio-/ pulmonologists in the assessment and use of outcome measures in pulmonary arterial hypertension related to systemic sclerosis

D. Huscher[1*], D. Pittrow[2*], O. Distler[3], C.P. Denton[4], I. Foeldvari[5], M. Humbert[6], M. Matucci-Cerinic[7], O. Kowal-Bielecka[8], J. Avouac[9], F. Behrens[10], P. Nash[11], C.F. Opitz[12], L.J. Rubin[13], J.R. Seibold[14], V. Strand[15], D.E. Furst[16], for all the co-authors of the EPOSS-OMERACT group

[1]*German Rheumatism Research Centre, Epidemiology, Berlin, Germany;*
[2]*Institute for Clinical Pharmacology, Medical Faculty, Technical University, Dresden, Germany.*

[*]*both authors contributed equally*

For a complete list of the co-authors' affiliations, see page 52.

Doerte Huscher
David Pittrow, MD, PhD
Oliver Distler, MD
Christopher P. Denton, MD
Ivan Foeldvari, MD
Marc Humbert, MD
Marco Matucci-Cerinic, MD
Otylia Kowal-Bielecka, MD
Jerome Avouac, MD
Frank Behrens, MD
Peter Nash, MD
Christian F. Opitz, MD
Lewis J. Rubin, MD
James R. Seibold, MD
Vibeke Strand, MD
Daniel E. Furst, MD
Please address correspondence to:
Daniel E. Furst,
1000 Veterans Ave, Rm 32-59,
Los Angeles, CA 90025, USA.
E-mail: defurst@mednet.ucla.edu

Reprints will not be available from the author.

Received on January 27, 2010; accepted in revised form on May 12, 2010.

Clin Exp Rheumatol 2010: 28 (Suppl. 58): S47-S52.

Key words: Systemic sclerosis, pulmonary arterial hypertension, outcome measures, Delphi.

Conflict of interest: see page 52.

ABSTRACT

Objective. *Pulmonary arterial hypertension in patients with systemic sclerosis is a disease involving multiple organ systems. We investigated the differences in perceptions of how to measure PAH-SSc among cardiologists, pulmonologists and rheumatologists. We also examined how a Delphi exercise can improve agreement among these subspecialties.*

Methods. *The outcome measures derived from the recent Delphi survey were used for a detailed analysis of the contribution of the various specialties contributing to it. We compared rheumatologists and cardiologist/pulmonologists with regards to preferences and ratings of various endpoints and the actual use of tools to measure these outcomes. We also examined the effects of the Delphi process among these groups.*

Results. *We could show that the different expert groups each tended to contribute differently to the development of the core set of measures and that interactions in the Delphi process resulted in convergence of rankings. Despite agreement on the high importance of the domains in the Delphi, the use of tools within those domains was sometimes divergent and dependent on specialty.*

Conclusion. *Based on these results, use of differing tools in the diagnosis and treatment of PAH-SSc can be anticipated. Further, the convergence of results provides evidence, for the first time, for the ability of various approaches in these disciplines to reach harmonious endpoints of care for PAH-SSc patients. A collaborative, interdisciplinary approach is advantageous for PAH-SSc patients.*

Introduction

In the last decade, substantial improvements have been made in the treatment of pulmonary arterial hypertension (PAH), mainly owing to the introduction of new therapeutic agents including prostaglandin derivatives (epoprostenol, iloprost, and treprostinil), endothelin receptor antagonists (bosentan, sitaxsentan and ambrisentan, and phosphodiesterase-5-inhibitors (sildenafil and tadalafil) (1, 2). While at first glance the design and endpoints used in these randomised controlled PAH trials appear well established, they have been criticised on the basis of a recent metaanalysis that populations were too homogeneous (thus not reflecting real life populations), as was the methodology (3). This causes concern, as the surrogate endpoints used in these studies are, beyond common belief, not necessarily valid. Experts who convened in workshops on endpoints in PAH trials from the Third World Symposium on Pulmonary Hypertension in 2003 in Venice and 2008 in Dana Point concluded that none of the endpoints currently used in PAH trials is optimal (4, 5). For example, the 6-minute walk test (6-MWT), the most widely used surrogate endpoint and the only measure of exercise capacity accepted by the FDA, is not validated for PAH patients with less severe disease (NYHA/WHO functional class I/II) (4). None of the typically used surrogate parameters in PAH, including 6-MWT, correlates well with survival (3).

In PAH associated with systemic sclerosis (PAH-SSc), the validation of possible study endpoints is even less convincing than in idiopathic PAH (6). In a systematic review performed at the

OMERACT-6 workshop on Outcome Measure Development for Clinical Trials in SSc, a variety of endpoints used in clinical trials were assessed according to the criteria of the "OMERACT Filter" of truth, discrimination, and feasibility (7). The only PAH endpoint that was judged to be "ready for use in clinical trials in SSc" patients was right heart catheterisation (7). However, this tool is invasive and therefore usually not appropriate for repeated measures or routine follow-up, and obviously cannot provide information beyond haemodynamics. Echo Doppler has not been fully validated for use in PAH-SSc (8). A combination of noninvasive testing methods (echo Doppler, MRI and pulmonary function tests) was recently shown to be inferior to RHC in the initial evaluation of PAH-SSc (9). Further, Health Assessment Questionnaire-Disability Index (HAQ-DI), a self-assessment measure of function, is not an adequate measure of PAH status in PAH-SSc patients (10).

Against this background, we performed a multidisciplinary Delphi survey to define a core set of outcome measures for clinical trials in PAH-SSc on a statistical basis modified by logical and medical rationale (11). A final set of seven domains and 10 tools were rated as important by the group composed of rheumatologists, cardiologists and pulmonologists. They were recommended for formal validation in prospective trials (Table I).

It is evident and has often been stressed that the clinical manifestations of PAH in SSc require a close interdisciplinary approach between various specialists to optimise diagnosis and treatment in this vulnerable patient group (12, 13). As opposed to other settings such as cancer care (14), the extent to which interdisciplinary collaboration benefits a PAH-SSc patient has not been examined. The available core set of outcome measures derived from the recent Delphi survey can be used for a detailed analysis of the contribution of cardiology, pulmonology and rheumatology to a consensus approach to PAH-SSc patient care. Therefore, we aimed to compare rheumatologists and cardiologist/pulmonologists with regards to (a) preferences and ratings of various endpoints (domains and tools), and (b) the self-reported actual use of outcome measures.

Table I. Final core set of domains and measurement tools defined by Delphi survey [adapted from reference (11)].

Domain	Measurement tools
Lung vascular	Right heart catheter, echocardiography
Exercise testing	6MWT, oxygen saturation at exercise
Cardiac function	Right heart catheter, echocardiography
Dyspnea	Dyspnea VAS
Discontinuation of treatment	Adverse events, serious adverse events
Quality of life	SF-36, HAQ DI
Global state by physician	Survival

6MWT: 6-minute walking test; VAS: visual analogue scale; SF-36: short form 36 score; HAQ DI: Health Assessment Questionnaire disability index.

Materials and methods

This analysis is based on data gathered in a Delphi study, conducted in 2006 (11). In brief, respondents (rheumatologists, cardiologists, pulmonologists) of this three-stage Delphi survey were asked to score each domain and tool on the survey for use as outcome measures in randomised controlled trials in PAH-SSc using a 5-point scale (1="not appropriate at all", 5="very appropriate"). In addition, participants were asked in round 1 whether they were actually using the tool. After stage 2 and stage 3, the number of domains and of tools were reduced according to a cluster analysis based on the revised ratings. This led to a final core set of domains and tools considered to be important for use as outcome measures in randomised controlled trials in PAH-SSc.

Among participants, the cardiologist group (n=15) was rather small for subgroup analysis. We compared the three subgroups cardiologists, pulmonologists (n=30) and rheumatologists (n=45) for possible merging. Based on 3 criteria – use of the 73 predefined tools, ranking of the 17 domains and ranking of the 86 domain-specific listed tools in round 1 – merging of the pulmonologists and cardiologists was appropriate (see results) and these two groups were combined in subsequent analyses.

Limited to the 17 domains, we then analysed how ratings developed over the course of the 3 Delphi stages for rheumatologists and the combined group of cardio-/pulmonologists. Thereby we aimed to detect the individual contribution of the two expert groups to the selection of the domains during the Delphi process and to identify possible interactions regarding the rating behavior of the two groups.

Finally, we examined the actual use of tools from the available core set of outcome measures by rheumatologists and cardio-/pulmonologists, looking for potential differences in the application of the selected tools. We counted tools with multiple listings as used, if "I use it" was ticked in any of the domains, hence the number of tools evaluated for "use of tool" reduced to 73 from originally 86 domain-specific tools.

Mean values of all participants of a considered expert group for the use of a tool or rating of a domain or tool were the basis for all analyses.

For the initial comparison of the 3 expert groups, normal distribution of the pairwise differences was verified with the Kolmogorov-Smirnov-test. Since for all merging criteria pairwise differences were normally distributed, differences between expert groups for the 3 criteria were tested with the paired *t*-test. Hence mean and standard deviation are given. SPSS 15.0 was used for analysis.

Results

Initial differences between specialist groups

The cardiologist group with 15 raters was rather small for subgroup analysis, so we aimed at combining this group with either pulmonologists or rheumatologists for the intended analysis.

Therefore we compared initial rating behavior and the daily use of tools by the 3 specialties.

In Fig. 1, each boxplot illustrates the distribution of the absolute differences for ratings or frequency of use between any two specialist groups. Outliers are illustrated as separate circles with the respective name of the domain or tool.

First, differences in the use (0-100%) of the 73 distinct tools listed in round 1 were compared. Cardiologists and pulmonologists showed similar frequencies in use of tools in most cases (Fig. 1a), since the mean difference between these two subspecialties was -0.6% (±12.1%, $p>0.05$). The comparison of rheumatologists and cardiologists not only resulted in a higher deviation (±21.3%), but also a mean difference of 9.1%, indicating a higher use of the listed tools by cardiologists compared to rheumatologists ($p=0.0005$). Pulmonologists were most dissimilar from rheumatologists in the use of tools ($p<0.0001$) with a mean difference of 9.8% (±19.3%).

A homogenous picture was seen for ratings of domains (scale: 1 to 5 points) in round 1 (Fig. 1b). There was essentially no difference between cardiologists and pulmonologists (-0.002±0.25 points, $p>0.05$). The boxplot for the differences between rheumatologists and cardiologists looks quite similar, but with a "soft" skewing towards higher ratings by cardiologists (0.06±0.37 points, $p>0.05$). Again, the largest numerical, though not significant, differences and the greatest deviation was seen when comparing pulmonologists and rheumatologists (0.07±0.43 points, $p>0.05$). In agreement with the use of tools, rating of domains was similar for cardiologists and pulmonologists.

For ratings of tools on a scale from 1 to 5 points (Fig. 1c), cardiologists tended to rate tools slightly but not significantly higher (0.10±0.47 points, $p>0.05$) than pulmonologists. The comparison of cardiologists and rheumatologists showed the lowest mean difference (-0.04±0.53 points, $p>0.05$). Rheumatologists gave about 0.13 points (±0.47) higher ratings for tools than pulmonologists ($p=0.01$).

Taken together, these findings did not comprise any significant differences between cardiologists and pulmonologists, in contrast to the other pairwise expert comparisons. On the basis of the above analysis showing that cardiologists and pulmonologists use tools in a very similar manner and gave comparable initial ratings for domains and tools, both groups were merged and analysed jointly in later analyses.

Development of ratings for domains

To assess the development of ratings and the rater interaction, we surveyed the ratings for the 17 domains over the course of the Delphi exercise (Fig. 2).

In round 1, the combined specialist group cardio-/pulmonologists (Ca/Pu) differed most from the rheumatologists. The Ca/Pu group had higher ratings on a 1-5 points rating scale for "miscellaneous symptoms" (+1.2 points) and "biomarkers" (+0.5 points), and lower ratings for "pulmonary arterial pressure" (-0.7 points) compared to rheumatologists (Rh). A high degree of agreement between the Ca/Pu and rheumatology groups was noted on the importance of the domains "cardiac function", "exercise capacity", "dyspnea", "global state as assessed by physician", "participation" and "quality of life" (mean rating >4 out of 5, difference in ratings ≤0.1 points).

In round 2, the differences in the "pulmonary arterial pressure" domain decreased, because cardio-/pulmonologists increased their ratings. For "miscellaneous symptoms" the differences also decreased, because rheumatologists increased their ratings. Differences increased by at least 0.2 units in only three domains: (1) in the "participation" domain, the two groups diverged because both groups decreased their ratings by different amounts; (2) in the "biomarkers" domain, the differences increased because rheumatologists decreased their ratings; and (3) in the "heart imaging" domain both ratings diverged.

In round 3, the "biomarkers" domain again converged, with both groups migrating toward one another. For the now combined domain "pulmonary arterial pressure (including lung vascular)" convergence was marked, since cardio-/pulmonologists increased their ratings.

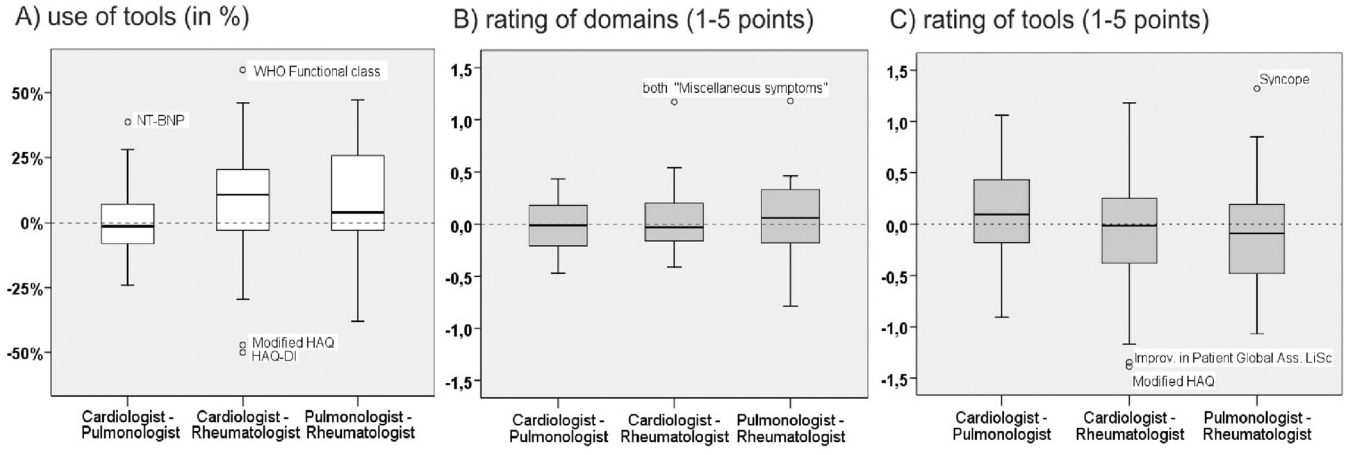

Fig. 1. Pairwise differences between the three specialists groups: boxplots illustrate the distributions of absolute differences in (**A**) mean use of 73 tools, (**B**) mean rating of 17 domains in round 1, and (**C**) mean rating of 86 domain-assigned tools in round 1.

The domain with the highest disagreement was "lung parenchymal", which remained unchanged over the course of the 3 rounds (final rating: Ca/Pu: 4.4, Rh: 4.8 points). For all the other domains the already small differences from round 2 remained unchanged.

Taken together, these data show that in this Delphi exercise, there was immediate agreement among 5 domains already in round 1. Out of the 7 domains that achieved increasing agreement during the 3 rounds, one expert group adapted their ratings to the others' in 5 cases. Interactions occurred between the two expert groups, with more frequent adaptation of one group to the other, rather than convergence on the median. An unchanged but very small difference (≤0.4 points) remained for 3 domains. An afterwards increased divergence (+0.2 units) was noted for 2 domains. By round 3, all ratings differed by less than 0.4 points (<10% of the associated ratings). In the end, agreement emerged regarding the high importance of seven out of 17 domains to be measured in patients with PAH-SSc in clinical trials (core domains).

Use of tools of the final core set by the specialist groups

A selection process similar to that used for domains was used for the tools. Here, also, a core set of highly rated tools (Table I) emerged after the three rounds of the Delphi survey (11). In the final round both expert groups gave mean ratings of at least 4.0 points (again on a rating scale from 1 to 5 points) to all tools in Table I. These were ultimately selected by the terminal cluster analysis as those with high importance.

Despite agreement on their importance, use of these tools was different between rheumatologists and cardio-/pulmonologists (Fig. 3). The greatest differences for tool use were found for "survival" (38%), "HAQ-disability index" (35%), "adverse" or "serious adverse events" (26%), "6-minute walking test" (15%), and "cardiac right ventricular function with pulmonary capillary wedge pressure" (21%). The last tool was selected by the terminal cluster analysis but not included into the core set (11). The difference for the "SF-36 questionnaire" was 6% and thus marginal. For all other tools the differences between groups were ±3%. Remarkably, the "severity of dyspnea visual analogue scale" as a part of the final core set is only used by less than a quarter in each specialist group. Despite these differences in tool use between rheumatologists and cardio-/pulmonologists, we already saw agreement on the high importance of most of these tools in round 1 (Fig. 3), when each participant gave her/his evaluation unaware of the other participants' ratings.

Discussion

The present analysis investigated possible differences between specialist groups in the choice and rankings of outcome measures in PAH-SSc, which usually involves treatment by several specialists. To our knowledge, no similar analysis on the interaction of medical disciplines in the care of patients has been published previously.

Pulmonologists and cardiologists were very similar in their use of tools and also showed agreement on the importance of the predefined domains and tools right from the start. Thus, they could be regarded as one joint expert

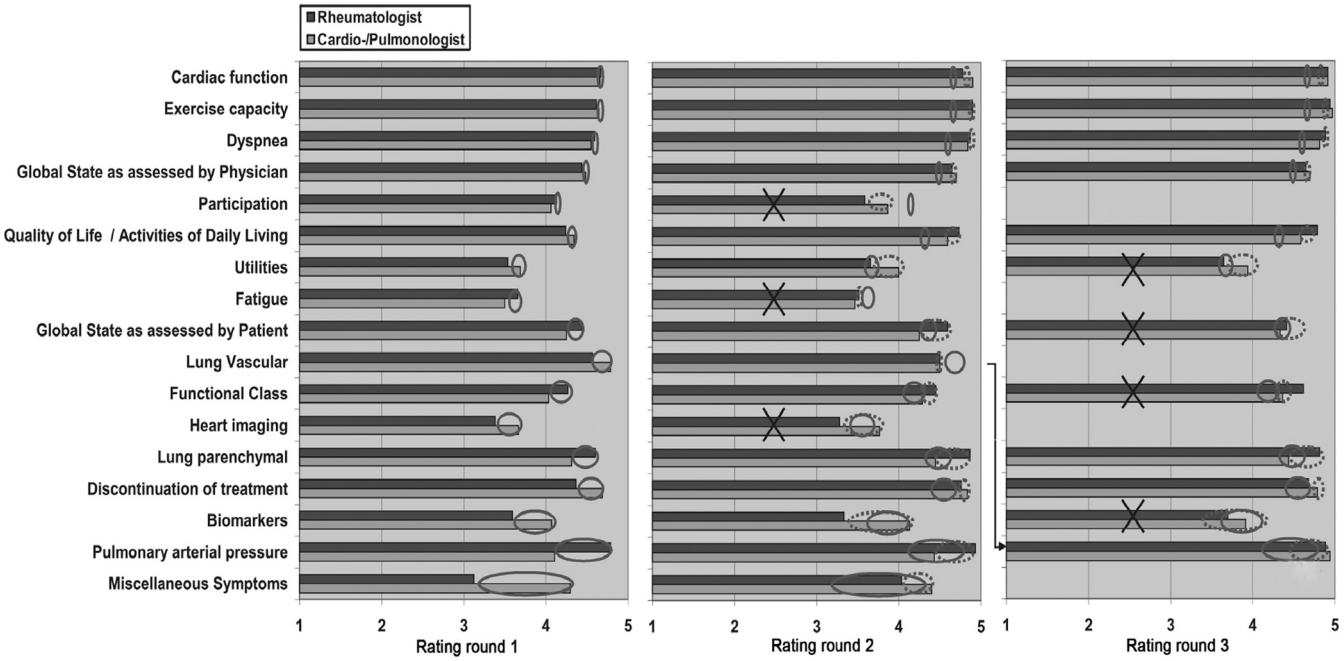

Fig. 2. Comparison of cardio-/pulmonologists with rheumatologists for ratings of domains in round 1-3; domains are sorted by ascending differences of rating in round 1.

Differences in ratings of round 1 are illustrated as solid line ellipse, differences in ratings of round 2 are illustrated as dashed line ellipse. Out-crossed domains were eliminated after that round as result of a cluster analysis, which grouped domains according to low and high ratings. To ease and support the selection process, domains "lung vascular" and "pulmonary arterial pressure" were merged after round 2, as well "miscellaneous symptoms" dropped after moving its left tools to suitable domains (11).

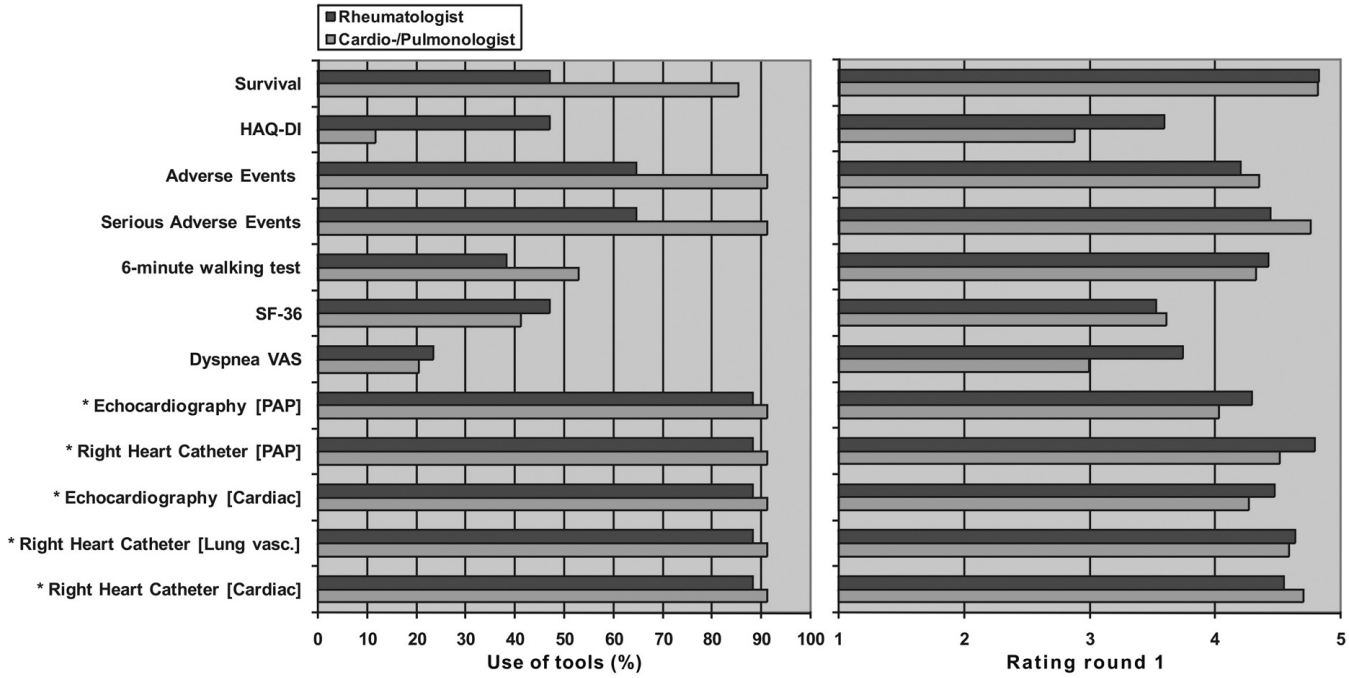

Fig. 3. Tools of the final core set: use of these tools and their rating in round 1 sorted by descending differences in use. For tools listed in several domains, the referring domain is given in square brackets: PAP: Pulmonary arterial pressure; Cardiac: Cardiac function; Lung vasc.: Lung vascular.

group for this analysis. Merging them led to balanced group sizes compared to rheumatologists.

In the comparison of cardio-/pulmonologists and rheumatologists, it is evident that the assessment of tools/domains and their actual use have to be regarded separately. The final core set consists of domains and tools with broad agreement regarding their importance, although the actual use of certain tools was different, defined by specialty. This can be illustrated by the example of a cardiologist or pulmonologist who knows about the diagnostic value of a HAQ-DI, but does not personally use the questionnaire.

The interaction of the two expert groups deserves further consideration. In the second round, agreement increased for some domains. Agreement usually resulted as one adapted to the other, rather than a "meeting in the middle". This indicates that one group influenced the other by the median rating of the previous round shown in the rating list. Thus, a "regression to the mean" for the ratings, which one could expect for this kind of internet-based, purely number-controlled Delphi exercise, did not generally occur. At the same time, disagreement increased for some domains since one group reduced their already initial lower rating or increased their already initial higher rating. Thus, we observed the effect that less convinced raters insisted on their previous rating while more convinced raters even increased their scoring. Vice versa, convinced raters kept their scoring, but less convinced raters reduced their scoring further, even in consideration of a lower or respectively higher median general rating. Such effects of reciprocal influence on decision making are an important and vital feature of the Delphi process, regardless of the topic (15). The dynamics and reasons for such changes deserve further research.

For the domains in the last round (round 3) of the Delphi process, ratings generally converged. The dynamic of the scoring underlines the deliberated harmonisation of ratings during the Delphi exercise. Particularly domains with agreement regarding their low importance were removed over the course of cluster-analysis after round 2 and round 3. Thus, individual expert opinion did play a role in the selection process and could trigger moving a certain domain into the final core set.

Most tools of the final core-set were already rated as "important" in Delphi round 1 by both expert groups. At the same time use of these tools was remarkably different for some of them. This is one of the most important findings of this study, with direct practical applications because it indicates that the employment of a combination of all these tools in patient care will depend on multidisciplinary teamwork. While rheumatologists lead in the application of the HAQ-DI, cardio-/pulmonologists dominate in undertaking the 6-minute-walking test or measuring the cardiac right ventricular function. In addition, we could identify highly scored tools that were not frequently applied by any expert group. Here questions arise as to why these tools are scored as important but not used. Ongoing research of the EPOSS group will answer whether tools of the final core set fulfill the OMERACT criteria.

These differences imply a potentially different approach to patient care between cardiologists/pulmonologists and rheumatologists, at least within the

arena of pulmonary hypertension. The rheumatologists were more involved with overall patient function and quality of life (more use of functional and quality of life tools) while the cardiologists/pulmonologists were more comfortable with the specific quantification and effects of the pulmonary status per se. Both are, of course, legitimate, but each group could benefit from the approach of the other. This exercise helped broaden perspective and undoubtedly made all the physicians more aware of the tools used by their colleagues caring for the same patient but from a different perspective. Specifically this study can, and should, lead to more uniformity of care across disciplines and, hopefully, a higher quality of care.

In conclusion, we show that involving expert groups from different specialties in the Delphi exercise is a useful approach, as it influenced the selection process of the final core set and, we believe, led to a more balanced and credible outcome. Similar interactions between disciplines can be anticipated in the daily clinical care of PAH-SSc patients. This provides strong experimental evidence for interdisciplinary care of patients with PAH-SSc, a disease involving several organ systems. In addition, the data provide, for the first time, experimental evidence that the use of a range of important diagnostic tools in the care of PAH-SSc patients depends on interdisciplinary teamwork.

Acknowledgements

The study was supported with a joint unrestricted educational grant from Actelion Pharmaceuticals, Allschwil, Switzerland; Pfizer Pharmaceuticals Ltd., London, UK; United Therapeutics; and Bayer Schering Pharma, Germany.

Co-authors' affiliations

[3]Department of Rheumatology, University Hospital Zurich, Switzerland;
[4]Centre for Rheumatology, Royal Free and University College, Medical School, London, United Kingdom;
[5]Hamburger Zentrum für Kinder- und Jugendrheumatologie, General Hospital Eilbek, Germany;
[6]Service de Pneumologie et Reanimation Respiratoire, Centre des Maladies Vasculaires Pulmonaires, Hopital Antoine-Beclere, Université Paris-Sud, Clamart, France;
[7]Department of Medicine, Division of Rheumatology, Denothe Centre, University of Florence, Italy;
[8]Department of Rheumatology and Internal Medicine, Medical University of Bialystok, Poland;
[9]Rheumatology A Dept, RDU, Paris, France;
[10]Division of Rheumatology/ ZAFES, J.W. Goethe University, Frankfurt, Germany;
[11]Department of Medicine, University of Queensland, Queensland, Australia;
[12]Medizinische Klinik II, DRK-Kliniken Berlin, Westend, Berlin, Germany;
[13]Division of Pulmonary and Critical Care Medicine, University of California, San Diego School of Medicine, La Jolla, USA;
[14]University of Michigan Scleroderma Program, Ann Arbor, Michigan, USA;
[15]Division of Rheumatology, Stanford University, Portola Valley, CA, USA;
[16]Division of Rheumatology, Department of Medicine, David Geffen School at UCLA, Los Angeles, CA, USA.

Conflict of interest: Dr Distler has a consultancy relationship and/or has received research funding from Actelion, Pfizer, Encysive, FibroGen, Ergonex, NicOx, BMS and Biovitrum in the area of potential treatments of scleroderma and its complications. He has also received lecture honoraria from Actelion, Pfizer, Encysive and Ergonex; Dr Denton has received research grants and lecture fees or acted as a consultant to Actelion, Pfizer, GSK and United Therapeutics; Dr Seibold has funded research and/or consultancy relationships with Actelion, Gilead Pfizer and United Therapeutics; Dr Strand is a consultant and serves on the advisory boards of the following companies from whom she has received only consulting fees and/or honoraria (she does not participate in Speakers' Bureaus, nor holds stock): Abbott Immunology, Alder, Amgen Corporation, AstraZeneca, BiogenIdec, BMS, CBio, CanFite, Centocor, Chelsea, Crescendo, Cypress Biosciences Inc., Eurodiagnostica, FibroGen, Forest Laboratories, Genentech, GlaxoSmithKline, Human Genome Sciences, Idera, Incyte, Jazz Pharmaceuticals, Lexicon Genetics, Lilly, Logical Therapeutics, Lux Biosciences, Medimmune, MerckSerono, Nicox, Novartis, NovoNordisk, Nuon, Ono Pharmaceuticals, Pfizer, Rigel, Roche, Sanofi-Aventis, Savient, Schering-Plough, SKK, UCB, Xdx; Dr Furst has received consultancy fees and grant support from Actelion and Gilead, and honoraria from Actelion; the other co-authors have declared no competing interests.

References

1. GALIE N, HOEPER MM, HUMBERT M *et al*.: Guidelines for the diagnosis and treatment of pulmonary hypertension: The 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 the International Society of Heart and Lung Transplantation (ISHLT). *Eur Heart J* 2009; 30: 2493-537.
2. BARST RJ, GIBBS JS, GHOFRANI HA *et al*.: Updated evidence-based treatment algorithm in pulmonary arterial hypertension. *J Am Coll Cardiol* 2009; 54 (Suppl.): S78-84.
3. MACCHIA A, MARCHIOLI R, MARFISI R *et al*.: A meta-analysis of trials of pulmonary hypertension: a clinical condition looking for drugs and research methodology. *Am Heart J* 2007; 153: 1037-47.
4. HOEPER MM, OUDIZ RJ, PEACOCK A *et al*.: End points and clinical trial designs in pulmonary arterial hypertension: clinical and regulatory perspectives. *J Am Coll Cardiol* 2004; 43 (Suppl. S): 48S-55S.
5. MCLAUGHLIN VV, BADESCH DB, DELCROIX M *et al*.: End Points and Clinical Trial Design in Pulmonary Arterial Hypertension. *J Am Coll Cardiol* 2009; 54 (Suppl. S):S97-107.
6. FURST D, KHANNA D, MATUCCI-CERINIC M, PITTROW D, STRAND V: Systemic sclerosis - continuing progress in developing clinical measures of response. *J Rheumatol* 2007; 34: 1194-200.
7. MERKEL PA, CLEMENTS PJ, REVEILLE JD, SUAREZ-ALMAZOR ME, VALENTINI G, FURST DE: Current status of outcome measure development for clinical trials in systemic sclerosis. Report from OMERACT 6. *J Rheumatol* 2003; 30: 1630-47.
8. KOWAL-BIELECKA O, AVOUAC J, PITTROW D *et al*.: Echocardiography as an outcome measure in scleroderma-related pulmonary arterial hypertension: a systematic literature analysis by the EPOSS group. *J Rheumatol* 2010; 37: 105-15.
9. HSU VM, MOREYRA AE, WILSON AC *et al*.: Assessment of pulmonary arterial hypertension in patients with systemic sclerosis: comparison of noninvasive tests with results of right-heart catheterization. *J Rheumatol* 2008; 35: 458-65.
10. CHOW S, POPE JE, MEHTA S: Lack of correlation of the health assessment questionnaire disability index with lung parameters in systemic sclerosis associated pulmonary arterial hypertension. *Clin Exp Rheumatol* 2008; 26: 1012-7.
11. DISTLER O, BEHRENS F, PITTROW D *et al*.: Defining appropriate outcome measures in pulmonary arterial hypertension related to systemic sclerosis: A Delphi consensus study with cluster analysis. *Arthritis Rheum* 2008; 59: 867-75.
12. GIBBS J, HIGENBOTTAM T: Recommendations on the management of pulmonary hypertension in clinical practice. *Heart* 2001; 86: i1-i13.
13. BLACK C: Pulmonary arterial hypertension: are we doing enough to identify systemic sclerosis patients at high risk of this rare condition? *Rheumatology* 2005; 44: 141-2.
14. LUTTERBACH J, PAGENSTECHER A, SPREER J *et al*.: The brain tumor board: lessons to be learned from an interdisciplinary conference. *Onkologie* 2005; 28: 22-6.
15. LINSTONE H, TUROFF M (Eds.), *The Delphi Method. Techniques and Applications*: New Jersey Institute of Technology, California, USA; 2002.