Author’s response to reviews

Title: Prognostic value of cardiopulmonary exercise testing in patients with systemic sclerosis

Authors:

Ralf Ewert (ewert@uni-greifswald.de)

Till Ittermann (Itterman@uni-greifswald.de)

Dirk Habedank (d.habedank@drk-kliniken-berlin.de;db.habedank@t-online.de;db.habedank@tonline.de)

Matthias Held (matthias.held@missioklinik.de)

Tobias Lange (Tobias.Lange@klinik.uni-regensburg.de)

Michael Halank (Michael.Halank@uniklinikum-dresden.de)

Jörg Winkler (Lungewinkler@aol.com)

Sven Gläser (Sven.Glaeser@vivantes.de)

Horst Olschewsky (horst.olschewski@medunigraz.at)

Gabor Kovacz (gabor.kovacs@klinikum-graz.at)

Version: 1 Date: 21 Aug 2019

Author’s response to reviews:

We wish to thank the reviewers for their profound revision and the helpful comments. We have thoroughly revised our manuscript and re-written many aspects. Please find our answers in red below. We have also uploaded a version with the correction tracking function of MS Word, and our final version without the correction function. We hope that we could respond to all the points that have been raised during the review process.

We look forward to the assessment and decision.

To Reviewer 1 (Prof. Harrison W. Farber)
"Comments to Author:
In this submission, Ewert et al. have retrospectively assessed CPETs in 210 patients with scleroderma (SSc) from six centers. After examining multiple different metrics, the authors found that peak VO2 (<64.5% of predicted) and VE/VCO2-slope >35 were predictive of prognosis in these patients. These are interesting observations, but raise several questions.

1) What is the appropriate control group for this population? It is clear that SSc patients fare much worse than the general population - CPET metrics are not necessary to know this. One could argue that in SSc patients with ILD, this should be patients with similar degrees of ILD (but w/o scleroderma) or
in SSc-PH/PAH patients, a like group without SSc. In the group used in this submission it is not mentioned whether the groups were matched by sex (since SSc patients are predominantly female).

Answer

We think this is a justified objection. The reference values of all parameters are published and consented (ECSC, SHIP), so that a “normal” control group will not add so much information. We therefore omitted the control group, what also focused the manuscript a bit in the methods and the limitations section. Instead of this, we took up the reviewer’s proposal that differentiation in subgroups with and without ILD is more useful and calculated all parameters additionally with this splitting (table S1). As to be assumed, patients with and without ILD showed significant differences in all parameters that characterize pulmonary restriction and diffusion, had more often a RHC and a higher proportion of patients with pulmonary limitation at exercise (VE/MVV > 80%). However, there were no significant differences in co-morbidity, echocardiography, most hemodynamic and CPET parameters. Though a bit disappointing for the CPET method, this might be caused by the heterogeneity of both groups. This is illustrated by the proportion of 84% right heart catheterization in ILD vs. 54% in no-ILD what indicates on a pre-test bias. This pre-selection also prevents an interpretation of the higher (51%) proportion with pulmonary hypertension in the no-ILD vs. 32% in the ILD-group. We included the main aspects in the results section and discussion. Since we are restricted in the number of table, we suggest publishing the detailed tables as supplements.

We discussed in detail the calculation of alternative control groups, as proposed by the reviewer. The most interesting assessment would be an additional comparison with non-SSc ILD and non-SSc PAH patients, but we think this is beyond the scope of our study, even for our working group as a multiple center team.

"2) In the patients who underwent RHC (a minority), what was the reason for the RHC - were there any specific criteria that led to RHC? And what percentage had left-heart disease during the CPET (since it is becoming more apparent that SSc is a likely risk factor for development of diastolic dysfunction)? In the group who had PAH at RHC, were they treated and, if so, did this change prognosis? Lastly, were there differences in the predictive value of the metrics in patients who underwent RHC vs. those that did not?"

Answer

RHC was indicated in responsibility of each center due to the criteria of the ESC and ERS if clinical symptoms and echocardiographic criteria suggested a possible pulmonary hypertension. We applied the criteria defined by the expert consent group [Avouac J 2014] which summarize clinical findings (progressive or unexplained dyspnoea, signs of right heart failure), echocardiography (RVsys > 45 mmHg, right ventricle dilation) and DLCO (< 50%). The parameters of our patients reflect these criteria: The subgroup with RHC had a higher proportion of extensive ILD, tricuspid regurgitation, estimated RVsys, and a lower DLCO, FVC, 6-MWD. Moreover, nearly all CPET parameters in the RHC group were worse than in the non-RHC group (e.g. lower peakVO2 and higher VE/VCO2-slope), but this due to a selection bias because all study centers mostly assessed CPET parameters as indication criteria for the performance of the RHK. We extended the discussion section to clarify this point.

We agree with the reviewer that more detailed RHC data will support our results and the discussion. In addition to the analysis lcSSc vs. dcSSc (results section, 2nd paragraph), we suggest a comparison of patients with and without RHC in a new table S2. RHC data were available in 136 patients, of which 52 had PH, including a subgroup of 38 patients with PAH. None of the PAH patients had a PAWP > 15
mmHg, and even the total number (n=9) of patients with diastolic dysfunction (PAWP>15 mmHg) was rather small. Although the majority of PAH patients was on specific PAH treatment, the total number was too small to calculate an effect on prognosis. Similarly, a predictive value for the subgroups with PH and even more with diastolic dysfunction could not be calculated. We have inserted a summary of these statements in the results.

"3) Likewise, only a minority of patients underwent 6MWT. Was there a reason that some patients did and the majority did not? And, if so, how were these patients different? Were there differences in the predictive value of the metrics in patients who underwent 6MWT vs. those that did not?"

Answer
The inclusion criterion for our study was CPET, so that the additional performance of 6-MWT was within discretion of the respective study center. This led to the smaller number of patients with 6MWT. Moreover, the correlation between 6-MWD and hemodynamic parameters is well documented [Sanges S 2017], so that we did not focus on this context. However, we confirmed the prognostic value of the 6-MWD and calculated a cut-off (413m) that separates prognostic relevant groups. We depict this calculation in figure 1c.

"4) CREST is no longer considered a separate group by most Rheumatologists - it is a subgroup of limited SSc."

Answer
We certainly agree that CREST was regarded as a separate form of SSc for a long time, but was meanwhile defined as a part of the limited SSc [Knobler R 2017]. We considered this in the methods (3rd paragraph) and in figure 1b and summarized in the results/mortality already CREST and lcSSc without separate mention of CREST.

"5) The literature on the prognosis of patients with SSc-ILD is very controversial - there are much data to suggest that ILD is detrimental and much data suggesting that is is not predictive - this discussion should be better balanced. The difference between the older and newer literature may be aggressive treatment of the ILD with cyclophosphamide, mycophenolate, etc."

Answer
The controversial literature on ILD and prognosis is a very justified point. We included and discussed more balanced references and appended some sentences on newer therapies like MMF, Nintedanib and Rituximab. The related paragraph now reads:

......and with a meta-analysis that found the degree of interstitial changes to be an independent prognostic variable for mortality in SSc [43]. A recent study differentiated subforms of ILD and showed that a manifestation as usual interstitial pneumonia (UIP) has a 2.3fold risk of mortality compared to a manifestation as non-specific interstitial pneumonia (NSIP) [Mango RL 2018]. Moreover, new drugs seem to offer an effective therapy of ILD, as there are Rituximab [Thiebaut M 2018] [Narvaez J 2019], Mycophenolate [Volkmann 2017], their combination [Fraticelli P 2018], and Nintedanib [Distler O 2019]. These therapies improved parameters of pulmonary function that are related to prognosis as DLCO, DLCO/FVC and TLC [Caron 2018] [Thiebaut]5 [Barnes H]. These improved parameters contrast with the fact only one study actually proved a better survival in patients treated with the new immunosuppressive agents [Volkmann 2019]. Hence, there is a need of new parameters that better estimate long time survival under immunosuppression [Kouranos 2018].
"6) Ultimately, are the authors suggesting that all patients with SSc should undergo CPET? This would be difficult (impractical) to say the least and there are many SSc patients who cannot or will not ride a bike. Lastly, even if this were ever possible, it would have to be shown that changing these values with treatment altered the outcomes in these patients. In other words, what is the clinical utility of this findings above what is already known about scleroderma patients?"

Answer

We are aware of a longer debate on the therapeutic implications of both 6-MWT and CPET, the more that all contributing centers have been researching and publishing in the field of CPET. Recently, Rizzi M et al. discussed CPET as an alternative to 6-MWT. We think that just the integration of different cardiac, muscle and pulmonary pathologies in CPET parameters allows prognostication. This is in part reflected by the precedent questions 1-5, indicating that we have to subsume different etiologies and stages (ILD, PH, PAH, CREST, lcSSc…) under the generic term “SSc”. CPET captures these different patients and allows differentiation between more cardiac and more pulmonary manifestation, and is in this way superior to 6MWT. Moreover, CPET increases the pre-test probability for PH [Dumitrescu D et al. 2017] and in this way CPET may induce specific therapy.

We like to take up the comment and therefore we have justified our assessment of CPET more detailed in the introduction section, and formulated in discussion and summary more clearly that our study was the first to show practicability and prognostic relevance in a clinical routine setting. It would be very interesting to evaluate whether an improvement in CPET parameters could indicate on an improved prognosis. We think we should broaden the discussion in this point and we now refer to the study on 6-MWT and prognosis [Farber 2014] in the discussion paragraph.

Reda Girgis (Reviewer 2)

Comments to Author

"This is a large study assessing the prognostic value of CPET in SSc. The finding of strong relationship between various CPET parameters, as well as 6 MWD, is important, as it demonstrates the utility of exercise testing in this population.

Specific Comments

1. With the multitude of clinical parameters examined, it is not clear from what point survival was calculated. Under statistical analysis section it is stated that "For group status the follow-up time was calculated based at the time of diagnosis; for the other variables the time of first examination defined the starting point." What does group status refer to? Was survival taken from the time of CPET or other time point? If CPET and the other assessments were performed at different times, how was the survival impact of each variable assessed?"

Answer

We thank the reviewer for this comment, because the follow-up times really need more clarification. We had 3 different follow-up times and depicted them now explicitly in the text and figures:
1) At time of diagnosis for the comparison between dcSSc and lcSSc (groups 1 and 2), and for demographic data like age and gender.

2) At time of CPET for all other analyses except RHC data.

3) At time of RHC, especially for the prognostic value of the RVsys.

"2. It would be valuable to know if CPET was abnormal and predicted survival in patients without overt pulmonary involvement with ILD or PAH. The authors speculate that CPET may serve as early risk marker."

Answer
We took up the proposal of both reviewers that differentiation in subgroups with and without ILD or PAH would add valuable information. We calculated two new tables (table S1 and S2).

As to be assumed, patients with and without ILD showed significant differences in all parameters that characterize pulmonary restriction and diffusion, had more often a RHC and a higher proportion of patients with pulmonary limitation at exercise (VE/MVV&gt;80%). There were no significant differences in co-morbidity, echocardiography, hemodynamic and CPET parameters. The total number of PAH patients (n=38) was too small to calculate an effect on prognosis. Similarly, a predictive value for the subgroups with PH and even more with diastolic dysfunction could not be calculated.

The proportion of 84% right heart catheterization in ILD vs. 54% in no-ILD indicates on a pre-test bias. This pre-selection also prevents an interpretation of the higher (51%) proportion with pulmonary hypertension in the no-ILD vs. 32% in the ILD-group. We included the main aspects in the results section and discussion. Since we are restricted in the number of table, we suggest publishing the detailed tables as supplements.

Against this background, our last sentence in the discussion section needs more precision. As criticized by the reviewer, we should avoid the speculation that our data define CPET as an early risk marker. This is a summary of both literature [Dumitrescu D 2010] and our findings, that peakVO2 and VE/VCO2-slope can be sensitive to early vasculopathic manifestations.

"3. It is also not clearly stated if the models shown represent risk factors that are all independent of each other. e.g Does CPET add prognostic information to DLCO or KCO?"

Answer
We agree to the reviewer that the additional prognostic value of CPET should be clarified. We have therefore included the Harrel’s C value and a new calculation restricted to KCO, TLC and peakVO2 for the model in the text. The paragraph now reads:

“In addition to age, in this model FVC, KCO and peakVO2 in mL·kg 1·min 1 were significantly linked to survival (Harrel’s C =0.96). Exclusion of peakVO2 impaired the predictive value of the model (Harrel’s C =0.84). In a calculation restricted to KCO, TLC and peakVO2, only peakVO2 remained as parameter associated to survival.”
Please find the additional calculation in the table below. If necessary, this table could also be published as supplemental.

| Haz.  | Ratio     | Std. Err. | z     | P>|z| | 95% Conf. Interval |
|-------|-----------|-----------|-------|------|-------------------|
| TLC   | .9739016  | .0220484  | -1.17 | 0.243| .9316323 - 1.018089|
| KCO   | .9899941  | .0218806  | -0.45 | 0.649| .9480245 - 1.033822|
| peakVO2%pred | .9442787 | .0170444 | -3.18 | 0.001| .9114563 - .978283|

"4. Line 3 in Results section seems like it should say that dcSSc had a "lower" proportion of women that in lcSSc"

Answer
This is certainly completely correct, and we have corrected this in the text.