Review CO Diffusion

Major comment
The paper has improved however it is not there yet. Especially the comparison between DLNO and DLCO need further statistical analysis. All results should be expressed in Z scores and correlations should be done in Z scores to include the variability of the measurement. I’m curious to see how the comparison will be after taking variability into account. In the tables Z score should be added in addition figures should show Z scores instead of % predicted.

Other major comments

Abstract
Please add the following to the abstract:
In this cross sectional study patients with CF, DLNO and DLCO reflected CT-morphological alterations of the lung better than other measures. Thus the combined diffusing capacity for NO and CO may play a future role for the non-invasive, functional assessment of structural alterations of the lung in CF.

Introduction
‘ The diffusing capacity for nitric oxide (DLNO) can be used to directly describe pulmonary membrane diffusing capacity (Dm), without interfering with pulmonary capillary blood volume (Vc). In combination with DLCO, Vc and Dm can be determined in a single maneuver, based on the equation for the serial connection of resistances [7,8]. Diffusion properties of NO are similar to those of CO; however, its rate of reaction with red blood cells is much greater [9].’

Q: To what refers its. Please specify, before the authors state that Dm does not depend on pulmonary blood volume. Hence it must be DLco? Always specify when possible

‘Patients were excluded if severe exacerbations, resulting in hospitalisation, were recorded.’
Q: This issue was already raised in the first review. Specify the time span when these exacerbations lead to exclusion

‘Dm was expressed as DLNO divided by 1.97, and Vc was derived as previously described [7].’

Q: Not clear, specify why this was done

‘The order in which CTs were scored was not randomized.’

Q: Specification was requested by my as part of the first review. Clearly this absence of randomisation introduces the potential for bias, and is a substantial weakness of the study. This shortcoming should be addressed as weakness in the discussion

recently published reference value equations using gender, age, and height as predictors, respective DLNO %pred and DLCO %pred could be calculated as: 71±19 and 82±17 %pred [16]; 60±17 and 75±16 %pred [17]; 63±18 and 86±19 %pred [18].

Q: Sentence is not clear do you mean:

recently published reference value equations using gender, age, and height as predictors, respective mean DLNO %pred and DLCO %pred for the patient group were: 71±19 and 82±17 %pred [16]; 60±17 and 75±16 %pred [17]; 63±18 and 86±19 %pred [18].

Please go through the manuscript to take out these kind of unspecified sentences

‘15 subjects (71%) were below the 95% Lower Limit of Normal concerning DLNO compared to only 9 subjects (43%) concerning DLCO.’

Q: The idea of specifying the number of subjects below normal reference values is interesting. However the readership does not have the sd in their heads. These are very high for CO are they even higher for DNO? Comparison between %predicted do not make sense since it does not show the variability of the measurement.. Better is to express in SD or Z scores since Please do so throughout manuscript, tables etc. This is an important issue that is needed for correct comparison

‘There are, however, contradictory reports, suggesting that DLCO is a useful marker of disease severity in CF [2,20].’

Q: Be explicit: Based on the conflicting results it is unlikely that Dlco is a useful marker of disease severity in CF…..

‘Thus, mainly due to the higher sensitivity, the measurement of DLNO seems to have an advantage over DLCO in the assessment of CF related changes’.
Q: Higher sensitivity is not correct! You showed only correlations. There is no proof in these cross sectional data that it is sensitive to detect change. Please correct

‘Thus, both KCO and KNO reflect different aspects of lung morphology and functionality. In CF, VA and/or TLC may be reduced. This renders KNO, being less related to VA than KCO, superior for a noninvasive assessment of structural alterations independent of volume-dependent effects’.

Q: This still does not make sense.

‘Thus, both KCO and KNO reflect different aspects of lung morphology and functionality. In CF, VA and/or TLC may be reduced. This renders KNO, being less related to VA than KCO, superior for a noninvasive assessment of Dm of structural alterations independent of volume-dependent effects’.

‘Correlation coefficients were comparable or slightly higher when using %predicted values instead of absolute values for DLNO and DLCO, probably reflecting the fact that CT scores do not depend on anthropometric characteristics, in contrast to lung function indices.’

Please change to: Correlation coefficients were comparable or slightly higher when using %predicted values instead of absolute values for DLNO and DLCO, probably reflecting the fact that lung function indices depend on anthropometric characteristics, in contrast to CT scores

‘Moreover, CT scans were retrospectively investigated, with a maximum time interval between scan and measurements of 3 years and a maximum interval of 1 year in 15 patients.’

Q: Please state explicitly in the beginning of this paragraph that this is a major weakness of the study. Not doing the scan on the same day will affect the correlations substantially and introduces a potential bias. The paragraph addressing this issue can be shortened as well

‘In conclusion, our findings indicated that in patients with CF NO diffusing capacity was the functional measure that was suitable to quantify structural changes of the lung as assessed by CT scores. It outperformed measures derived from spirometry or bodyplethysmography. The suitability of NO diffusing capacity is in line with previous observations in healthy subjects or patients with other diseases than CF [23] and renders it a challenging question of whether the combined diffusing capacity for NO and CO has a potential to be included in the assessment and monitoring of CF.’

Q: This conclusion is by far to strong. First you can make comparisons between the accuracy of the various PFT measurements only when expressed in Z scores. It should be stated that this is a cross sectional study and that a longitudinal study is needed. Please change conclusion in such a way that it reflects adequately the paper with its shortcomings.