



## Early View

Original article

### **Development and validation of a simple tool for the assessment of home NIV: the S<sup>3</sup>-NIV questionnaire**

Elise Dupuis-Lozeron, Grégoire Gex, Patrick Pasquina, Pierre-Olivier Bridevaux, Jean-Christian Borel, Paola Gasche-Soccal, Wolfram Windisch, Jean-Louis Pépin, Jean-Paul Janssens, Dan Adler

Please cite this article as: Dupuis-Lozeron E, Gex G, Pasquina P, *et al.* Development and validation of a simple tool for the assessment of home NIV: the S<sup>3</sup>-NIV questionnaire. *Eur Respir J* 2018; in press (<https://doi.org/10.1183/13993003.011822018>).

This manuscript has recently been accepted for publication in the *European Respiratory Journal*. It is published here in its accepted form prior to copyediting and typesetting by our production team. After these production processes are complete and the authors have approved the resulting proofs, the article will move to the latest issue of the ERJ online.

Copyright ©ERS 2018

## Development and validation of a simple tool for the assessment of home NIV: the S<sup>3</sup>-NIV questionnaire

Elise Dupuis-Lozeron<sup>1</sup>, Grégoire Gex<sup>2,3</sup>, Patrick Pasquina<sup>3</sup>, Pierre-Olivier Bridevaux<sup>2,3,5</sup>, Jean-Christian Borel<sup>4</sup>, Paola Gasche-Soccal<sup>3,5</sup>, Wolfram Windisch<sup>6</sup>, Jean-Louis Pépin<sup>7,8</sup>, Jean-Paul Janssens<sup>3,5</sup>, Dan Adler<sup>3,5</sup>

1. Division of Clinical Epidemiology, Geneva University Hospitals, Geneva, Switzerland
2. Division of Lung Diseases, Hôpital du Valais, Sion, Switzerland
3. Division of Lung Diseases, Geneva University Hospitals, Geneva, Switzerland
4. AgirADom, Grenoble, France
5. University of Geneva Medical School, Geneva, Switzerland
6. Cologne Merheim Hospital, Department of Pneumology, Kliniken der Stadt Köln, Witten/Herdecke University, Faculty of Health/School of Medicine, Germany
7. HP2 Laboratory, Inserm U1042 Unit, Grenoble Alps University, Grenoble, France
8. EFCR Laboratory, Thorax and Vessels, Grenoble Alps University Hospital, Grenoble, France

### *Corresponding author:*

Dan Adler, MD  
Division of Pulmonary Diseases  
Geneva University Hospitals  
4 Rue Gabrielle Perret-Gentil  
1211 Geneva 14 / Switzerland  
Tel: +41 22 372 9902  
E-mail: [Dan.adler@hcuge.ch](mailto:Dan.adler@hcuge.ch)

### "Take home" message:

The S<sup>3</sup>-non-invasive ventilation (NIV) questionnaire provides clinicians and patients with a simple and reliable tool to assess important domains (symptoms, sleep quality and NIV-related side-effects) as a complement to physiological monitoring.

## Abstract

Patient-centered outcomes are significantly modified by long-term home non-invasive ventilation (NIV), but a short, self-administered, specific tool for routine clinical assessment is lacking. The aim of this study was to develop and validate the S<sup>3</sup>-NIV questionnaire, a short questionnaire to measure respiratory symptoms, sleep quality and NIV-related side-effects.

Stable patients under long-term home NIV were recruited from three outpatient NIV services. Questionnaire development consisted of a selection of core items for analysis, followed by item reduction, validation and test-retest reliability.

338 patients completed a 22-item questionnaire. Eleven items were removed because of non-scalability (n=2), redundancy (n=8) and lack of fit (n=1). The final version of the S<sup>3</sup>-NIV questionnaire consisted of 11 items covering two dimensions: “respiratory symptoms” (Cronbach alpha=0.84) and “sleep & NIV-related side-effects” (Cronbach alpha=0.77). Convergent validity was high between the “respiratory symptoms” subscale of the S<sup>3</sup>-NIV questionnaire and the St. George’s Respiratory Questionnaire (rho=-0.76; p<0.001), as well as the “sleep & NIV side-effects” subscale and the Quebec Sleep Questionnaire (rho=0.51; p<0.001). The S<sup>3</sup>-NIV questionnaire had a good test-retest reliability after 4 weeks (intraclass correlation coefficient=0.72).

The S<sup>3</sup>-NIV questionnaire is a short, valid and repeatable, self-completed tool for the routine clinical assessment of patients undergoing home NIV.

## Introduction

Home non-invasive ventilation (NIV) is increasingly used as a first-line, evidence-based treatment in chronic hypercapnic respiratory failure [1, 2]. Over the past decades, the transition from acute care to the home setting has been characterized by a paradigm shift in the medical and societal approach to chronic respiratory care. Such a transition was also made possible through technological improvements in ventilators, which allowed to download data from software [3]. However, implementing home NIV services is time-consuming, labour intensive and expensive. Therefore, it is critical to document whether improvements in lung function, arterial blood gas, readmission rates and eventually survival also translate into evidence-based improvements in patient-centred outcomes, such as respiratory symptoms and sleep quality, with acceptable NIV-related side-effects.

The correlation between patient-centered endpoints and physiological parameters used to monitor NIV is notoriously low [4, 5]. Indeed, some of the most important domains affected by chronic respiratory failure remain poorly explored, whereas complex monitoring tools such as polysomnography are recommended for the adjustment of NIV [3, 6]. Working with valid, reliable and specific instruments to assess patient-centred outcomes as a complement to NIV physiological monitoring is crucial to set achievable treatment goals and to improve communication between the patient and the physician. The Severe Respiratory Insufficiency (SRI) questionnaire and the Mageri Respiratory Foundation Questionnaire-28 are most often used in NIV trials to assess patient reported outcomes [7, 8]. However, their length and complex scoring algorithms limit their use in everyday clinical practice. In addition, treatment-related side effects of NIV are not covered by these classical questionnaires. Our aim was to develop a simple, short and reliable tool to assess respiratory symptoms, sleep and comfort (or discomfort) as a complement to physiological monitoring of home NIV efficacy [9].

## Methods

A detailed Methods section with a full description of models used for item analysis and selection as well as additional figures are available in the online supplementary material. A shorter version of the Methods section is provided here.

### *Participants*

Stable patients treated at home with NIV were recruited in three French-speaking regional and university hospitals: Geneva University Hospitals (n=153) and Hôpital du Valais (n=95) in Switzerland, and Grenoble University Hospital (n= 90) in France. All participants were “non-naïve” patients established on home NIV for at least 4 months prior to inclusion. NIV was used for a primary diagnosis of chronic obstructive pulmonary disease (COPD) (n=72; 21%), obesity hypoventilation syndrome (OHS) (n= 96; 29%), central breathing disturbances during sleep (CBD)[10] (n=114; 34%), neuromuscular disorders (n=39; 10%) and restrictive disorders (n=17; 5%). Exclusion criteria were patients with a history of recent exacerbation or hospitalization (<3 months), unable to read or write French language, unable to give informed consent, or unable to understand the questionnaire. Questionnaires were self-administered during a routine medical visit. All patients provided written informed consent and the study was approved by the ethics committee of all participating centers. The research was conducted according to the principles of the Declaration of Helsinki.

### *Item selection*

We selected all items pertaining to “respiratory complaints” and “attendant symptoms and sleep” from the French translation and cultural adaptation of the SRI questionnaire [11] for further psychometric validation. The need for an additional assessment of a “comfort” dimension was agreed on by all NIV experts in our group. To this end, issues related to the experience of NIV treatment (mainly comfort and side-effects) were investigated in depth during 15 qualitative interviews with patients (D.A.). Other items were selected from two NIV comfort scales with no formal psychometric validation [12, 13]. All items were rated on a 5-point Likert scale from “strongly disagree” to “strongly agree” and were related to patient’s status during the “last 4 weeks”. Selected core items consisted of 22 questions. Of these, 8 questions were related to

the “respiratory complaints” domain, 7 to “attendant symptoms and sleep”, and 7 to NIV-related side-effects.

### *Item analysis and reduction*

An exploratory factor analysis was performed to evaluate the number of underlying dimensions [14]. Then, an IRT graded response model was used on each unidimensional subscale for item reduction. Items with a low information function ((i.e., item precision based on difficulty and discrimination for all values of the total scale) or with information that was redundant with another item were removed [15]. Item reduction also took into consideration expert opinion from investigators and other clinical experts in the field. Mean values for each subscore were transformed accordingly with a total score ranging from 0 to 10. The lowest possible score (0) corresponded to the highest impact of disease and treatment on quality of life, while the highest possible score (10) corresponded to the lowest impact of disease and treatment on quality of life.

### *Reliability*

Internal consistency of each subscale was calculated by Cronbach’s alpha coefficient [16]. Test-retest reliability of the final instrument was assessed by the intra-class correlation coefficient (ICC) in a subset of stable patients (n=43) who completed the questionnaire 4 weeks after initial testing. Differential item functioning analysis (DIF) was used to assess whether the selected items were perceived differently by patients in differing disease categories. To assess the validity of the factor structure, the fit of a confirmatory item factor model was evaluated using the M2 model fit statistic [17], the root mean square error of approximation (RMSEA) and the standardised root mean square residual (SRMR).

### *Construct and discriminant validity*

Construct and discriminant validity was assessed by evaluating correlations between the final instrument and the French version of the St. George’s Respiratory Questionnaire (SGRQ) [18] and the Quebec Sleep Questionnaire [19].

All analyses were performed with R version 3.4.4 [20] and the mirt [21] and psych packages [22].

## Results

### *Patient demographics*

A total of 338 patients were included in the item reduction phase of the S<sup>3</sup>-NIV questionnaire.

Demographics and clinical data for the overall population and each participating center are presented in table 1. The most frequent indications for home NIV across all centers were CBD, OHS and COPD, including “overlap syndrome”. Patients had been established on home NIV for a median of 45 months (interquartile range [IQR]: 21-93). Median forced expiratory volume in 1 s (FEV<sub>1</sub>) was 36% of the predicted IQR (25-45) in COPD, 68% (47-85) in OHS, 88% (75-102) in CBD, 51% (28-76) in neuromuscular disorders and 51% (42-62) in restrictive disorders. Daily adherence downloaded from built-in ventilator software was 7.8 h (IQR: 5.4-9.1) in COPD, 7.5 h (5.5-9.2) in OHS, 7.1 h (5.9-8.1) in CBD, 9.4 h (8.1-10.7) in neuromuscular disorders and 8.2 h (6.3-10.0) in restrictive disorders.

### *Item analysis and reduction*

Patients used the full range of responses for the 22 selected core items (0-4), with mean item responses on the 0-4 scale ranging from 1.32±1.46 for “I have difficulties breathing during physical exertion” to 3.38±1.01 for “My ventilator inflates my lungs with too much pressure”. The rate of missing data was very low for all items (<3%). Preliminary non parametric IRT analysis performed on the 22 core items indicated that two items, “I go to sleep easily” and “I sleep through the night easily”, were not scalable and were rejected. Based on the 20 remaining items, exploratory factor analysis suggested two latent factors that were interpreted as “respiratory symptoms” and “sleep & NIV-related side-effects” dimensions (online supplementary figure S1).

Four items pertaining to “respiratory symptoms” were rejected because their information function showed redundancy with other items: “I suffer from breathing problems even without physical exertion”, “I sometimes feel dizzy”, “I am tired during the day” and “I cough a lot”. For instance, the “cough” item was redundant with a “phlegm” item and was deleted after a round of expert opinion because cough encompasses less disease severity than the “phlegm” item, which was kept in the final tool. Four items pertaining to the “sleep & NIV-related side-effects” dimension were also rejected because of redundancy

with other items: “I often have neck pain”, “I often wake up at night”, “My ventilator doesn’t inflate my lungs enough”, and “My ventilator is too fast-paced”, and one item was rejected because of lack of fit: “My ventilator is too noisy”. The three most important issues related to the experience of NIV itself were: “The mask is uncomfortable”, “Air leaks are bothering me”, and “Air is too dry”. All three items contributed to model information and fit and were kept in the final questionnaire.

### *Psychometric properties of the finalized S<sup>3</sup>-NIV questionnaire*

The 11 items retained were used to assess the psychometric properties of the finalized version of the S<sup>3</sup>-NIV questionnaire (figure 1). The mean±standard deviation (SD) of the S<sup>3</sup>-NIV questionnaire score was 6.71±1.82. Internal consistency of each subscale was good with a Cronbach’s alpha coefficient of 0.84 for the “respiratory symptoms” dimension, and 0.77 for the “sleep & NIV-related side-effects” dimension. Test-retest reliability was equally good (ICC=0.72; 95% CI: 0.54-0.84) in 44 patients 4 weeks after the initial assessment.

The distribution of the S<sup>3</sup>-NIV score is depicted in figure 2 (panel A) and shows that the entire scaling range is used in our validation study. Eighty per cent of 338 patients used 48% of the scaling range; 10% had a score <4.32 units and 10% had a score >9.01 units. Figure 2 (panel B) shows S<sup>3</sup>-NIV total scores by disease category with no floor or ceiling effect in any disease category. The impact of disease and treatment in COPD patients was as severe as in neuromuscular patients, whereas a lower impact of disease and treatment was demonstrated for OHS and CBD patients (ANOVA,  $p=0.004$ ) (figure 2, panel B). The DIF analysis of the “respiratory symptoms dimension” identified only one item (“I am often short of breath”) as significantly different between disease categories ( $p=0.03$ ). All items from the “sleep and NIV side-effects” domain were identified as invariant across disease categories. The validity of the latent structure of the 11 remaining items was assessed using a confirmatory item factor model with two latent factors for the two domains. The fit of the model was very good ( $M_2=0.5$ ;  $df=10$  [ $p=0.400$ ];  $RMSEA=0.011$  [95% CI, 0; 0.061];  $SRMR=0.076$ ).



### *External validation*

Construct validity was assured by the correlation analysis between S<sup>3</sup>-NIV questionnaire scales and various scales of the SGRQ and Quebec Sleep Questionnaire. A high correlation was found between the “respiratory symptoms” domain of the S<sup>3</sup>-NIV questionnaire (5 items) and the symptom scale of the SGRQ (8 items) ( $\rho=-0.76$ ;  $p<0.001$ ). The correlation between the “sleep & NIV side-effects” domain of the S<sup>3</sup>-NIV questionnaire and the Quebec sleep questionnaire was  $\rho=0.51$  ( $p<0.001$ ) (online supplementary figure S2, panels A and B). Correlation of the S<sup>3</sup>-NIV questionnaire total score with the Quebec sleep questionnaire was  $\rho=0.67$  ( $p<0.001$ ) and  $\rho = -0.60$  ( $p<0.001$ ) with the SGRQ symptom scale. Patients’ perception of symptoms as assessed by the S<sup>3</sup>-NIV questionnaire was not correlated with objective pulmonary function measurement (FEV<sub>1</sub> % of predicted:  $\rho=0.12$  [ $p=0.109$ ]; forced vital capacity % of predicted:  $\rho=0.08$  [ $p=0.270$ ]) or with daily adherence downloaded from the built-in ventilator software ( $\rho=0.09$ ;  $p=0.158$ ).

## **Discussion**

This study provides a validation of a short, simple, patient-completed specific tool to monitor home NIV as a complement to the monitoring of physiological variables. It covers important patient-oriented dimensions related to NIV monitoring and treatment, i.e. respiratory symptoms, sleep quality and NIV-related side-effects. This is achieved with a limited number of items ( $n=11$ ) validated in a large international sample of patients corresponding to practices in different NIV services. Testing for internal consistency demonstrated acceptable to excellent reliability. Construct validity was also established using the SGRQ and the Quebec Sleep Questionnaire as references.

A stepwise approach combining arterial blood gas samples, pulse oxymetry with capnography (when available), data from ventilator built-in software and polygraphy or polysomnography under NIV [6, 9, 12, 23] is now recommended to adapt ventilator settings. However, there is a lack of tools to assess patient-oriented outcomes in order to inform the organization of NIV services and allow comparisons for quality of care. Patient-centered outcomes and physiological parameters used to monitor chronic respiratory failure address different dimensions of care as illustrated in our data by the weak association between the S<sup>3</sup>-NIV

questionnaire and FEV<sub>1</sub>.

We combined factor analysis and IRT with expert clinical judgement. Items with poor measurement properties were first flagged for deletion according to predetermined rules. When designing the scale, we balanced weaknesses and strengths in overall contributions before deciding whether to include or exclude any item. No single theoretical criterion was used. Clinical judgement was shared and discussed between expert co-authors throughout the entire questionnaire design process. The final S<sup>3</sup>-NIV questionnaire consists of 11 items formatted as a 5-point Likert scale, thus making the questionnaire easy for patients to understand and to complete. We consider that the S<sup>3</sup>-NIV questionnaire may be the most suitable tool currently available as it has been specifically developed for clinical practice in NIV services. Before the wider application of the S<sup>3</sup>-NIV questionnaire is supported by evidence, its use should be restricted to the assessment of clinically stable, non-naïve patients admitted for routine follow-up of NIV. Another application of a short, patient-oriented, self-administered tool can be anticipated with the exponential development of home NIV tele-monitoring [24]. In this particular setting, it could be very beneficial to merge the data downloaded from the built-in software with a brief self-administered questionnaire in an integrated e-health platform tailored to patients' needs and NIV services. This paradigm shift in care planning has already been tested with success in cancer patients [25], as well as in severe COPD [26] and sleep apnea [27].

The 11 items were selected to cover a wide range of disease severity in a large cohort with the intention of demonstrating persistent discriminant power across different settings, practices and underlying diseases. This is illustrated by the distribution of the S<sup>3</sup>-NIV questionnaire score over the entire scaling range for the overall study population, including in specific disease subgroups, as demonstrated by DIF analysis. Indeed, both moderately severe ambulatory patients and individuals with a high NIV dependency, such as those in late stages of neuromuscular disorders, were included in the study population to reflect current trends in home NIV indications [28]. Conversely, excellent measurement properties have been demonstrated for the S<sup>3</sup>-NIV questionnaire across all patient subgroups, which allows to use it widely, increases practicability for the organization of home NIV services, and facilitates the training of home care

providers. Overlap syndrome (i.e. COPD+OSAS) was merged with the COPD subgroup. However, a post hoc analysis (data not shown) of the S<sup>3</sup>-NIV questionnaire subscores in COPD and overlap syndrome demonstrated that even if the total scores were the same, patients diagnosed with overlap syndrome exhibited a trend to have higher scores on the respiratory domain and lower scores on the sleep domain compared to pure COPD patients. Further studies with an appropriate sample size are needed to address this specific question.

Our study has some limitations. It presents cross-sectional data, similar to the original SRI questionnaire [7]. Thus, a prospective longitudinal study will be required to assess the minimal clinically important difference, as well as the sensitivity of our tool to changes over time or changes induced by NIV settings/interface modifications. Another limitation stems directly from the questionnaire development methodology. We only included items pertaining to respiratory symptoms, sleep quality and side-effects related to NIV in the preselected core of items before the reduction process. We then eliminated items that were infrequent, redundant, or with poor scaling properties with the aim to develop a short practical tool with the least number of items. This may have two potential undesirable consequences. First, the S<sup>3</sup>-NIV questionnaire can only be used as a clinical tool for NIV monitoring and is not a surrogate measure of general health status or quality of life. In contrast, the original longer SRI questionnaire may be more appropriate if a systematic analysis of health status is required or for interventional randomized controlled trials. Second, we may have underestimated important items for a small group of patients that would have been retained using a different methodology based on clinical importance. For instance, noisy alarms may seriously disturb sleep and, in turn, alter daytime symptoms in a few patients using conservative alarm settings. Therefore, our instrument may not be appropriate for a small number of very dependent patients. Specific studies need to be dedicated to these subgroups. Finally, S<sup>3</sup>-NIV questionnaire was only tested in a French speaking international cohort with a thorough item reduction methodology starting from the original SRI questionnaire. This was made possible because the French translation and cultural adaptation of the original questionnaire was already available[11]. Moreover, as most items pertaining to respiratory symptoms and sleep have been already translated and formally validated in English[29], we argue that the S<sup>3</sup>-NIV questionnaire does not need further English cultural validation to be relevant and

used in stable patients treated with home NIV.

In summary, the S<sup>3</sup>-NIV questionnaire provides clinicians and patients with a simple and reliable tool to assess three important domains related to home NIV as a complement to physiological monitoring of home NIV. This will also facilitate shared decision-making in home NIV services. Although the S<sup>3</sup>-NIV questionnaire is short with a simple scoring algorithm, its content and layout can also serve as a backbone during consultations to identify key areas to be further explored in a personalized interview in order to optimize care delivery in this patient population.

### **Acknowledgements**

The authors would like to thank all patients who participated in the study. The authors wish to thank Rosemary Sudan for editorial assistance.

### **Funding**

This study was partly supported by the *Fonds de la Loterie Romande 2014* and by the *Fonds de la Ligue Pulmonaire Genevoise 2014*. JLP is supported by the French National Research Agency in the framework of the *Investissements d'avenir* programme (grant no. ANR-15-IDEX-02) and the Chair of Excellence of the *E-health and integrated care in chronic disease-Université Grenoble Alpes foundation*.

## References:

1. Lloyd-Owen SJ, Donaldson GC, Ambrosino N, Escarabill J, Farre R, Fauroux B, Robert D, Schoenhofer B, Simonds AK, Wedzicha JA: **Patterns of home mechanical ventilation use in Europe: results from the Eurovent survey.** *Eur Respir J* 2005, **25**(6):1025-1031.
2. Janssens J-P, Cantero C, Adler D, Pasquina P, Younossian AB, Gasche P: **Home non-invasive ventilation in the Canton of Geneva: ventilators, modes, interfaces.** *European Respiratory Journal* 2017, **50**(suppl 61).
3. Janssens JP, Borel JC, Pepin JL: **Nocturnal monitoring of home non-invasive ventilation: the contribution of simple tools such as pulse oximetry, capnography, built-in ventilator software and autonomic markers of sleep fragmentation.** *Thorax* 2011, **66**(5):438-445.
4. Curtis JR, Deyo RA, Hudson LD: **Pulmonary rehabilitation in chronic respiratory insufficiency. 7. Health-related quality of life among patients with chronic obstructive pulmonary disease.** *Thorax* 1994, **49**(2):162-170.
5. Jones PW: **Issues concerning health-related quality of life in COPD.** *Chest* 1995, **107**(5 Suppl):187s-193s.
6. Medicine NTTotAAoS: **Best Clinical Practices for the Sleep Center Adjustment of Noninvasive Positive Pressure Ventilation (NPPV) in Stable Chronic Alveolar Hypoventilation Syndromes.** *Journal of Clinical Sleep Medicine : JCSM : Official Publication of the American Academy of Sleep Medicine* 2010, **6**(5):491-509.
7. Windisch W, Freidel K, Schucher B, Baumann H, Wiebel M, Matthys H, Petermann F: **The Severe Respiratory Insufficiency (SRI) Questionnaire: a specific measure of health-related quality of life in patients receiving home mechanical ventilation.** *J Clin Epidemiol* 2003, **56**(8):752-759.
8. Carone M, Bertolotti G, Anchisi F, Zotti A, Donner C, Jones P: **Analysis of factors that characterize health impairment in patients with chronic respiratory failure. Quality of Life in Chronic Respiratory Failure Group.** *European Respiratory Journal* 1999, **13**(6):1293-1300.
9. Adler D, Rabec C, Janssens J-P: **Home monitoring and follow-up of long-term NIV.** *ERS Practical Handbook of Noninvasive Ventilation* 2015:265.
10. Randerath W, Verbraecken J, Andreas S, Arzt M, Bloch KE, Brack T, Buyse B, De Backer W, Eckert DJ, Grote L: **Definition, discrimination, diagnosis and treatment of central breathing disturbances during sleep.** *European respiratory journal* 2016:ERJ-00959-02016.
11. Cuvelier A, Lamia B, Molano LC, Muir JF, Windisch W: **[The French translation and cultural adaptation of the SRI questionnaire. A questionnaire to assess health-related quality of life in patients with chronic respiratory failure and domiciliary ventilation].** *Rev Mal Respir* 2012, **29**(5):705-713.
12. Adler D, Perrig S, Takahashi H, Espa F, Rodenstein D, Pepin JL, Janssens JP: **Polysomnography in stable COPD under non-invasive ventilation to reduce patient-ventilator asynchrony and morning breathlessness.** *Sleep Breath* 2012, **16**(4):1081-1090.
13. Janssens JP, Metzger M, Sforza E: **Impact of volume targeting on efficacy of bi-level non-invasive ventilation and sleep in obesity-hypoventilation.** *Respir Med* 2009, **103**(2):165-172.
14. Horn JL: **A rationale and test for the number of factors in factor analysis.** *Psychometrika* 1965, **30**(2):179-185.
15. Nering ML, Ostini R: **Handbook of polytomous item response theory models:** Taylor & Francis; 2011.
16. Cronbach LJ: **Coefficient alpha and the internal structure of tests.** *psychometrika* 1951, **16**(3):297-334.
17. Cai L, Hansen M: **Limited-information goodness-of-fit testing of hierarchical item factor models.** *British Journal of Mathematical and Statistical Psychology* 2013, **66**(2):245-276.

18. Bourbeau J, Maltais F, Rouleau M, Guimont C: **French-Canadian version of the Chronic Respiratory and St George's Respiratory questionnaires: an assessment of their psychometric properties in patients with chronic obstructive pulmonary disease.** *Can Respir J* 2004, **11**(7):480-486.
19. Lacasse Y, Bureau MP, Series F: **A new standardised and self-administered quality of life questionnaire specific to obstructive sleep apnoea.** *Thorax* 2004, **59**(6):494-499.
20. Core Team R: **R: A language and environment for statistical computing.** R Foundation for Statistical Computing, Vienna, Austria. 2016. In.; 2017.
21. Chalmers RP: **mirt: A multidimensional item response theory package for the R environment.** *Journal of Statistical Software* 2012, **48**(6):1-29.
22. Revelle W: **psych: Procedures for Personality and Psychological Research.** Northwestern University, Evanston, Illinois, USA 2018.
23. Gonzalez-Bermejo J, Perrin C, Janssens JP, Pepin JL, Mroue G, Leger P, Langevin B, Rouault S, Rabec C, Rodenstein D: **Proposal for a systematic analysis of polygraphy or polysomnography for identifying and scoring abnormal events occurring during non-invasive ventilation.** *Thorax* 2012, **67**(6):546-552.
24. Ambrosino N, Vitacca M, Dreher M, Isetta V, Montserrat JM, Tonia T, Turchetti G, Winck JC, Burgos F, Kampelmacher M *et al*: **Tele-monitoring of ventilator-dependent patients: a European Respiratory Society Statement.** *European Respiratory Journal* 2016, **48**(3):648-663.
25. Basch E, Deal AM, Dueck AC, *et al*.: **Overall survival results of a trial assessing patient-reported outcomes for symptom monitoring during routine cancer treatment.** *JAMA* 2017, **318**(2):197-198.
26. Kessler R, Casan-Clara P, Koehler D, Tognella S, Viejo JL, Dal Negro RW, Díaz-Lobato S, Reissig K, Rodríguez González-Moro JM, Devouassoux G *et al*: **COMET: a multicomponent home-based disease-management programme <em>versus</em> routine care in severe COPD.** *European Respiratory Journal* 2018, **51**(1).
27. Malhotra A, Crocker ME, Willes L, Kelly C, Lynch S, Benjafield AV: **Patient Engagement Using New Technology to Improve Adherence to Positive Airway Pressure Therapy: A Retrospective Analysis.** *Chest* 2018, **153**(4):843-850.
28. Cantero C, Pasquina P, Vienne F, Adler D, Younossian A, Prella M, Gasche P, Janssens J-P: **The Geneva Lake Study Update: A Descriptive and Cross-Sectional Multicenter Study Regarding the Current Trends of Long-Term Home Non-Invasive Ventilation: ASV Population.** In: *B68 DIAGNOSIS AND TREATMENT OF SLEEP DISORDERED BREATHING*. American Thoracic Society; 2018: A3982-A3982.
29. Ghosh D, Rzehak P, Elliott MW, Windisch W: **Validation of the English Severe Respiratory Insufficiency Questionnaire.** *Eur Respir J* 2012, **40**(2):408-415.

TABLE 1: Demographics and clinical characteristics by recruiting centers				
	All centers	Geneva	Valais	Grenoble
<b>Subjects (%)</b>	338	153 (45)	95 (28)	90 (27)
<b>Gender (male %)</b>	252 (75)	104 (68)	85 (90)	63 (70)
<b>COPD, n (%)</b>	72 (21)	38 (25)	14(15)	20 (22)
<b>OHS, n (%)</b>	96 (28)	51 (33)	21 (22)	24 (27)
<b>CBD, n (%)</b>	114 (34)	35 (23)	57 (60)	22 (24)
<b>Neuromuscular disorders, n (%)</b>	39 (12)	15 (10)	3 (3)	21 (23)
<b>Restrictive disorders, n (%)</b>	17 (5)	14 (9)	0 (0)	3 (3)
<b>Established on home NIV for n months, median (IQR)</b>	45 [21;93]	57 [21;93]	57 [21;105]	45 [33;63]
<b>Age (years), median (IQR)</b>	69 [61;75]	71 [61;77]	69 [62;73]	67 [60;72]

NIV: non-invasive ventilation; COPD: chronic obstructive pulmonary disease; OHS: obesity-hypoventilation syndrome; CBD: central breathing disturbances during sleep; IQR: interquartile range.

TABLE 2: S <sup>3</sup> -NIV questionnaire score and subscores according to disease category			
	S <sup>3</sup> - NIVquestionnaire total score, mean± SD	Respiratory symptoms subscore, mean± SD	Sleep & NIV- related side effects subscore, mean ± SD
<b>All subjects</b>	6.71±1.82	6.28±2.42	7.07±1.97
<b>COPD</b>	6.09±1.58	5.01±2.21	6.99±1.81
<b>OHS</b>	6.72± 1.92	6.37 ± 2.31	6.99± 2.02
<b>CBD</b>	7.09± 1.78	7.09± 2.29	7.08± 2.14
<b>Neuromuscular disorders</b>	6.50± 1.90	5.80± 2.58	7.12± 1.75
<b>Restrictive disorders</b>	7.22± 1.61	6.80± 2.13	7.57± 1.84

COPD: chronic obstructive pulmonary disease; OHS: obesity-hypoventilation syndrome; CBD: central breathing disturbances during sleep; SD: standard deviation.



## Figure legends

Figure 1: The final version of the S3-NIV questionnaire.

The total score can be computed as the average of all answered items multiplied by 2.5. The lowest possible score (0) corresponds to the highest impact of disease and treatment, while the highest possible score (10) corresponds to the lowest impact of disease and treatment.

Figure 2: Panel A. Empirical cumulative distribution of the S3-NIV questionnaire total score in 338 patients treated with home NIV. Panel B. Distribution of the S3-NIV total score by disease category. Dots represent individual data.

COPD: chronic obstructive pulmonary disease; OHS: obesity hypoventilation syndrome; CBD: Central breathing disturbances during sleep.

Figure 1

	always true	mostly true	sometimes true	mostly untrue	completely untrue	Score
1. I suffer from breathing problems when I eat.	<input type="text" value="0"/>	<input type="text" value="1"/>	<input type="text" value="2"/>	<input type="text" value="3"/>	<input type="text" value="4"/>	
2. I often have a headache.	<input type="text" value="0"/>	<input type="text" value="1"/>	<input type="text" value="2"/>	<input type="text" value="3"/>	<input type="text" value="4"/>	
3. I wake up at night with breathing difficulties.	<input type="text" value="0"/>	<input type="text" value="1"/>	<input type="text" value="2"/>	<input type="text" value="3"/>	<input type="text" value="4"/>	
4. I am often short of breath.	<input type="text" value="0"/>	<input type="text" value="1"/>	<input type="text" value="2"/>	<input type="text" value="3"/>	<input type="text" value="4"/>	
5. I have trouble breathing when I speak.	<input type="text" value="0"/>	<input type="text" value="1"/>	<input type="text" value="2"/>	<input type="text" value="3"/>	<input type="text" value="4"/>	
6. There is often mucus in my airways.	<input type="text" value="0"/>	<input type="text" value="1"/>	<input type="text" value="2"/>	<input type="text" value="3"/>	<input type="text" value="4"/>	
7. I have difficulties breathing during physical exertion.	<input type="text" value="0"/>	<input type="text" value="1"/>	<input type="text" value="2"/>	<input type="text" value="3"/>	<input type="text" value="4"/>	
8. I am disturbed by leaks	<input type="text" value="0"/>	<input type="text" value="1"/>	<input type="text" value="2"/>	<input type="text" value="3"/>	<input type="text" value="4"/>	
9. My mask is uncomfortable	<input type="text" value="0"/>	<input type="text" value="1"/>	<input type="text" value="2"/>	<input type="text" value="3"/>	<input type="text" value="4"/>	
10. I receive too much air from my ventilator	<input type="text" value="0"/>	<input type="text" value="1"/>	<input type="text" value="2"/>	<input type="text" value="3"/>	<input type="text" value="4"/>	
11. I suffer from nasal or oral dryness	<input type="text" value="0"/>	<input type="text" value="1"/>	<input type="text" value="2"/>	<input type="text" value="3"/>	<input type="text" value="4"/>	

TOTAL

Total divided by 11 x 2.5

S<sup>3</sup>-NIV Score

Figure 2 Panel A

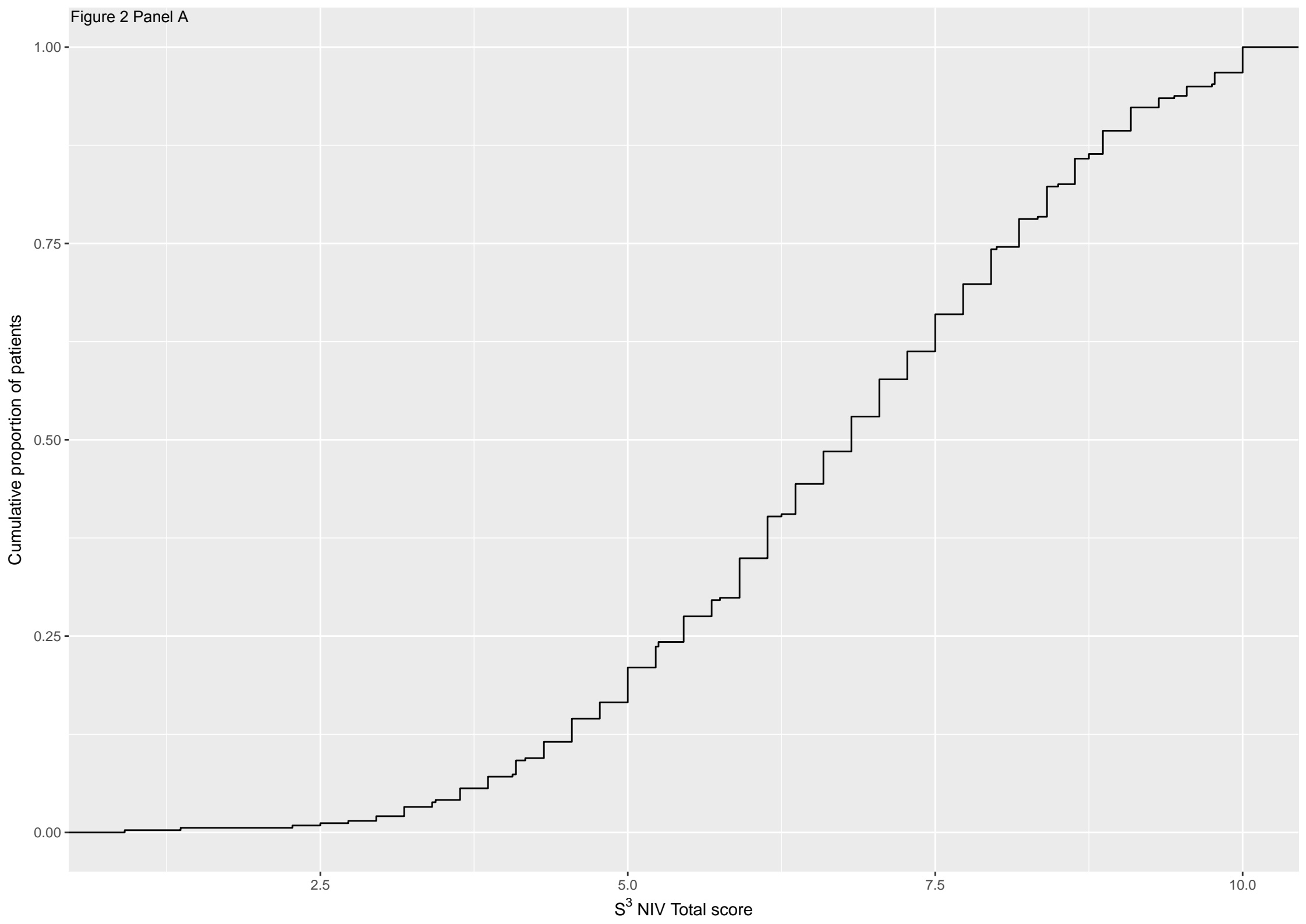
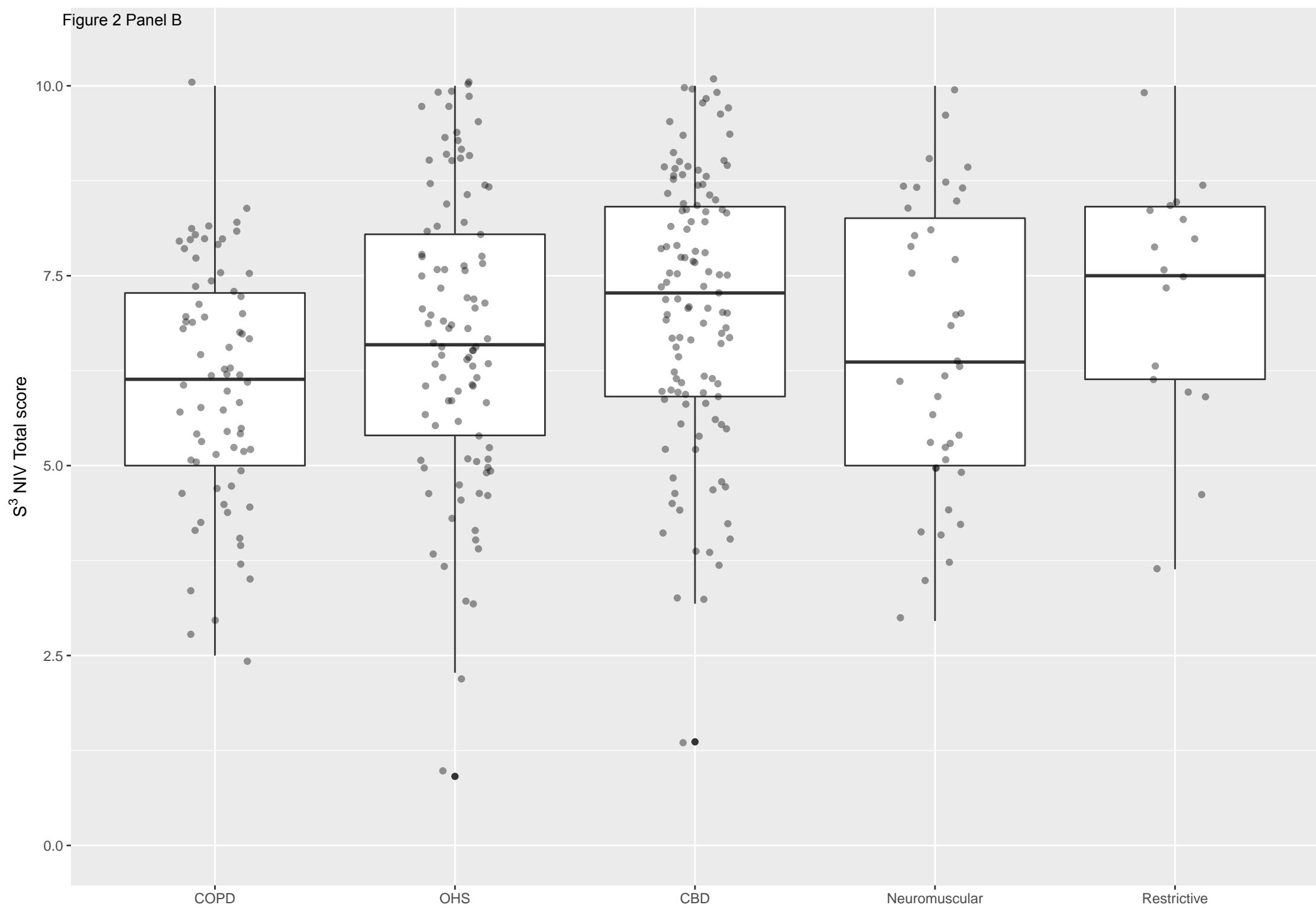


Figure 2 Panel B



## **Supplementary material**

### **Methods**

#### *Item selection*

The main objective was to develop a clinical tool consisting of only a few items to capture a wide range of respiratory, sleep and NIV treatment experiences, but with valid and reliable measurement properties. For this purpose, we selected all items pertaining to “respiratory complaints” and “attendant symptoms and sleep” from the French translation and cultural adaptation of the SRI questionnaire [1] for further psychometric validation. The need for an additional assessment of a “comfort” dimension was agreed on by all NIV experts in our group. To this end, issues related to the experience of NIV treatment (mainly comfort and side-effects) were investigated in depth during 15 qualitative interviews with patients (D.A.). Items were also selected from two NIV comfort scales used in previous studies by our group, but with no formal psychometric validation [2, 3]. All items were rated on a 5-point Likert scale from “strongly disagree” to “strongly agree” and were related to patient’s status during the “last 4 weeks”. Selected core items consisted of 22 questions. Of these, 8 questions were related to the “respiratory complaints” domain, 7 to “attendant symptoms and sleep”, and 7 to NIV-related side-effects.

#### *Item analysis and reduction*

Concepts of psychometric analysis are summarised in table S1. The Mokken scale analysis (non-parametric Item Response Theory (IRT) analysis) was first performed to check the assumptions of unidimensionality, local independence, and latent monotonicity that are required by parametric IRT models. To evaluate whether the unidimensionality assumption was plausible, an exploratory factorial analysis using the polychoric correlations and the number of underlying dimensions was performed as a parallel analysis [4]. Factor loadings were then used to investigate the meaning of each dimension. At this stage, numerous IRT models (graded rating scale model, the graded response model, and the generalized partial credit mod) were tested on each unidimensional subscale for item reduction. The Akaike information criterion was applied to choose the IRT model that best fitted the data. The graded rating scale model was used to assess item information (i.e., item precision based on difficulty and discrimination for all values of the total scale). Items with a low information function or with information that was redundant with another item were removed [5]. Item fit statistics (signed chi-square test) were

checked for each item. Item reduction also took into consideration expert opinion from investigators and other clinical experts in the field. Mean values for each subscore were transformed accordingly with a total score ranging from 0 to 10. The lowest possible score (0) corresponded to the highest impact of disease and treatment on quality of life, while the highest possible score (10) corresponded to the lowest impact of disease and treatment on quality of life.

### ***Reliability***

Internal consistency of each subscale was calculated by Cronbach's alpha coefficient using data from all patients. Values  $>0.7$  are considered acceptable for group comparison and values  $>0.8$  are considered good or excellent [6]. Test-retest reliability of the final instrument was assessed by the intra-class correlation coefficient (ICC) in a subset of stable patients ( $n=43$ ) who completed the questionnaire 4 weeks after initial testing. A value of 0.7 indicates good reliability. To assess the validity of the factor structure, the fit of a confirmatory item factor model was evaluated using the M2 model fit statistic [7], the root mean square error of approximation (RMSEA) and the standardised root mean square residual (SRMR). Differential item functioning analysis (DIF) was used to assess whether the selected items were perceived differently by patients in differing disease categories.

### ***Construct and discriminant validity***

Construct and discriminant validity was assessed by evaluating correlations between the final instrument and the French version of the St. George's Respiratory Questionnaire (SGRQ) [8] and the Quebec Sleep Questionnaire [9].

All analyses were performed with R version 3.4.4 [10] and the mirt[11] and psych packages[12].

#### Online supplementary table S1: Summary of psychometric concepts used in this study

**Loading:** measurement of the correlational relationship between the observed and the latent variable/s (also called “factor”).

**Unidimensionality assumption:** a scale is said to be unidimensional if it measures only one underlying construct (or one latent variable).

**Local independence:** this assumption implies that the response to any items is unrelated to any other item when the latent variable level is controlled.

**Monotonicity (latent monotonicity):** this means that the item step response functions are non-decreasing functions of the underlying construct, i.e. patients at a higher level of the underlying construct have a higher probability of scoring higher for an item.

**Item response theory (IRT)** is a group of specialised methods for the development and validation of rating scales with items of ordinal or nominal type. IRT considers a class of latent variable models that link mainly dichotomous and polytomous observed variables to a single latent variable.

**Mokken scale analysis:** an analytical method for ordinal data to assess the dimensionality and scalability of psychometric measures. It can be viewed as a non-parametric IRT model.

**Differential item functioning** occurs when subgroups of a sample respond in different ways to a particular item, despite having the same level of the underlying trait being measured.

### **Legend to online supplementary figures**

**Figure S1:** Visual illustration of the two-factor model in 20 of 22 selected core items.

Ovals indicate latent factors MR1 and MR2 regressed on all items (boxes) in the exploratory factor analysis. The value linking MR1 to MR2 is the correlation between the two factors. The values in the single-headed arrows are the factor loadings standardised to both the latent factor and the variances observed among the variables of the factor analysis model. The 22 cores items were the following ones:

- Q1: My ventilator doesn't inflate my lungs enough.
- Q2: I am disturbed by leaks.
- Q3: My mask is uncomfortable.
- Q4: My ventilator is too fast-paced.
- Q5: I receive too much air from my ventilator.
- Q6: My ventilator is too noisy.
- Q7: I suffer from nasal or oral dryness
- Q8: I suffer from breathing problems when I eat.
- Q9: I suffer from breathing problems even without physical exertion.
- Q10: I often have a headache.
- Q11: I sometimes feel dizzy.
- Q12: I wake up at night with breathing difficulties.
- Q13: I often have neck pain.
- Q14: I often wake up at night.
- Q15: I am often short of breath.
- Q16: I have trouble breathing when I speak.
- Q17: I cough a lot.
- Q18: There is often mucus in my airways.
- Q19: I have difficulties breathing during physical exertion.
- Q20: I am tired during the day.
- Q21: I go to sleep easily (not included in factor analysis).
- Q22: I sleep through the night easily (not included in factor analysis).



**Figure S2:** Panel A. Correlation between the “respiratory symptoms” S<sup>3</sup>-NIV questionnaire subscore and the SGRQ symptoms score. Panel B. Correlation between the “sleep and NIV side-effects” S<sup>3</sup>-NIV questionnaire subscore and the Quebec Sleep Questionnaire Score.

SGRQ: St. George's Respiratory Questionnaire.

## Supplementary references:

1. Cuvelier A, Lamia B, Molano LC, Muir JF, Windisch W: **[The French translation and cultural adaptation of the SRI questionnaire. A questionnaire to assess health-related quality of life in patients with chronic respiratory failure and domiciliary ventilation]**. *Rev Mal Respir* 2012, **29**(5):705-713.
2. Adler D, Perrig S, Takahashi H, Espa F, Rodenstein D, Pepin JL, Janssens JP: **Polysomnography in stable COPD under non-invasive ventilation to reduce patient-ventilator asynchrony and morning breathlessness**. *Sleep Breath* 2012, **16**(4):1081-1090.
3. Janssens JP, Metzger M, Sforza E: **Impact of volume targeting on efficacy of bi-level non-invasive ventilation and sleep in obesity-hypoventilation**. *Respir Med* 2009, **103**(2):165-172.
4. Horn JL: **A rationale and test for the number of factors in factor analysis**. *Psychometrika* 1965, **30**(2):179-185.
5. Nering ML, Ostini R: **Handbook of polytomous item response theory models**: Taylor & Francis; 2011.
6. Cronbach LJ: **Coefficient alpha and the internal structure of tests**. *psychometrika* 1951, **16**(3):297-334.
7. Cai L, Hansen M: **Limited-information goodness-of-fit testing of hierarchical item factor models**. *British Journal of Mathematical and Statistical Psychology* 2013, **66**(2):245-276.
8. Bourbeau J, Maltais F, Rouleau M, Guimont C: **French-Canadian version of the Chronic Respiratory and St George's Respiratory questionnaires: an assessment of their psychometric properties in patients with chronic obstructive pulmonary disease**. *Can Respir J* 2004, **11**(7):480-486.
9. Lacasse Y, Bureau MP, Series F: **A new standardised and self-administered quality of life questionnaire specific to obstructive sleep apnoea**. *Thorax* 2004, **59**(6):494-499.
10. Core Team R: **R: A language and environment for statistical computing**. R Foundation for Statistical Computing, Vienna, Austria. 2016. In.; 2017.
11. Chalmers RP: **mirt: A multidimensional item response theory package for the R environment**. *Journal of Statistical Software* 2012, **48**(6):1-29.
12. Revelle W: **psych: Procedures for Personality and Psychological Research**. Northwestern University, Evanston, Illinois, USA 2018.

Figure S1

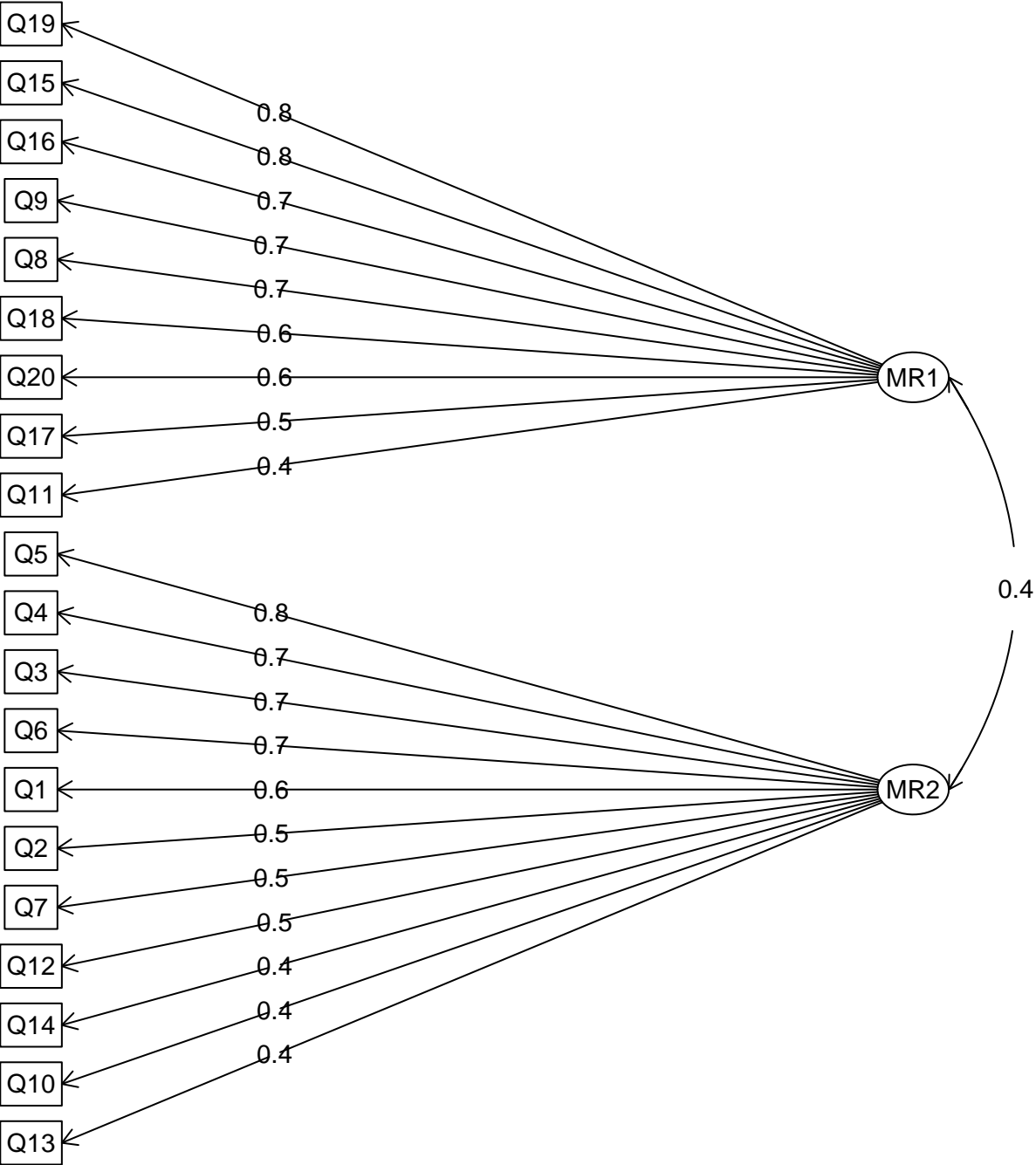


Figure S2 Panel A

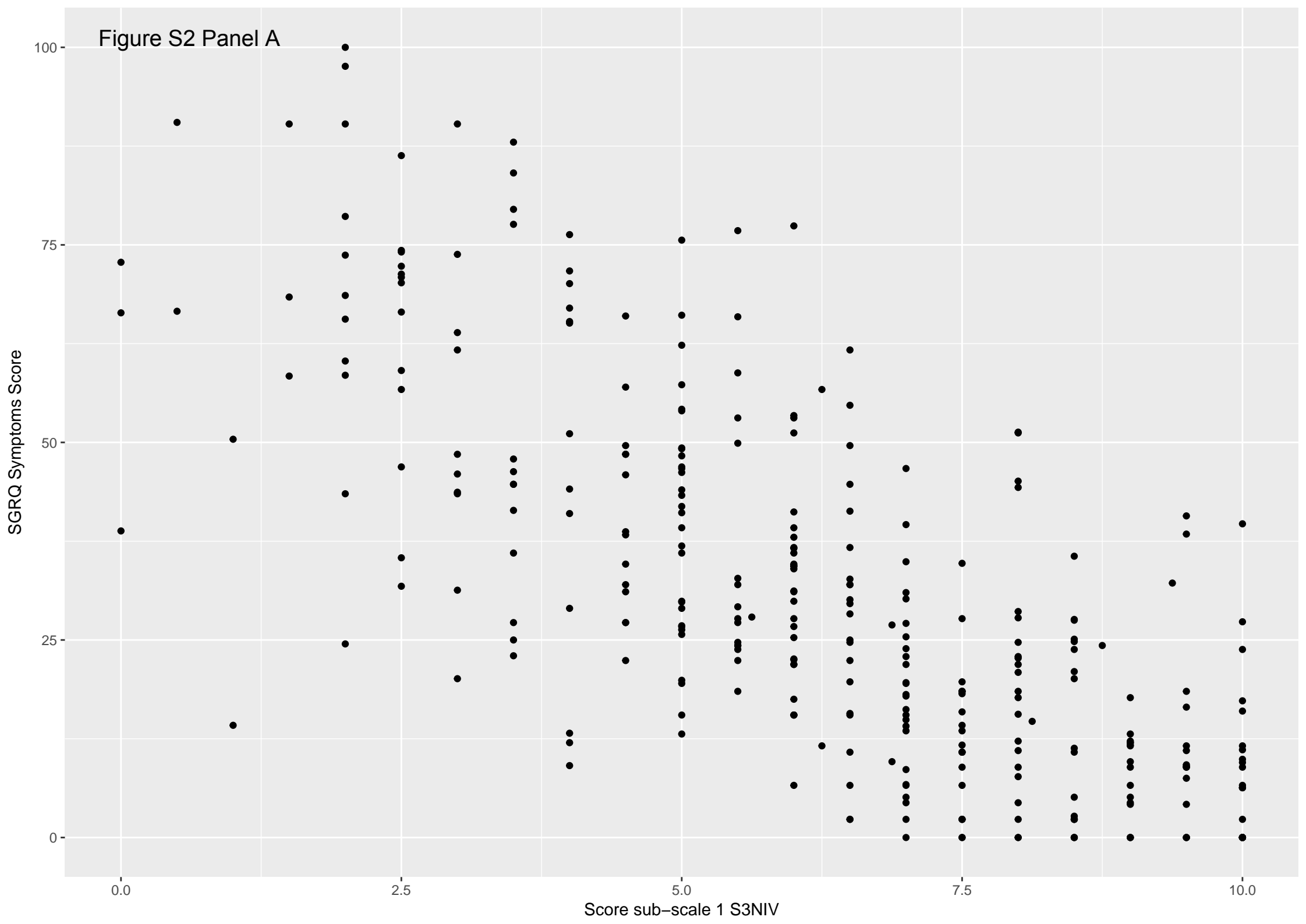


Figure S2 Panel B

