The Nijmegen Questionnaire: a valid measure for hyperventilation syndrome

Running title: Validity of the Nijmegen Questionnaire

Abstract

Hyperventilation syndrome is often undiagnosed due to its multi-systemic and apparently unrelated symptoms. The Nijmegen Questionnaire is used by clinicians to assess susceptible individuals, based on self-reporting symptoms attributed to hyperventilation syndrome. However, evidence of psychometric properties of this questionnaire is lacking. This study investigated two types of validity, using interviews and Rasch analysis. Data showed that the Nijmegen Questionnaire met criteria for content validity but not for structural validity. Content validity was supported by a high matching percentage between the symptoms identified within interview data and the current items on the Nijmegen Questionnaire (94%). Reported symptoms from study participants were conceptually congruent with most of the questionnaire items with minor language inconsistencies between patients and clinicians. Rasch analysis indicated a poor fit of the Nijmegen Questionnaire to the Rasch model, demonstrating poor structural validity. This study also developed a revised version of the Nijmegen Questionnaire, which did meet criteria for structural validity. Subsequently, a conversion table was created for transforming raw total scores of the questionnaire in the clinical and research settings. Physiotherapists should use the revised 15-item Nijmegen Questionnaire for clinical and research purposes since it provides more accurate representation of the severity of patients' symptoms than the original scoring.

Keywords

Hyperventilation, Nijmegen Questionnaire, assessment, validity, Rasch analysis

Introduction

Dysfunctional breathing is an umbrella term describing breathing disorders where acute and/or chronic changes in breathing pattern result in dyspnoea and other symptoms in the absence or in excess of the magnitude of physiological, respiratory or cardiac disease (Boulding, Stacey, Niven, & Fowler, 2016). Boulding and colleagues (2016) suggested in their literature review the following classification for dysfunctional breathing patterns: hyperventilation syndrome, periodic deep sighing, thoracic dominant breathing, forced abdominal expiration, and thoraco-abdominal synchrony. Dysfunctional breathing is increasingly recognised as a costly health concern, given the involvement of various medical or surgical investigations prior to correctly identifying susceptible individuals (Chaitow, Morrison & Gilbert, 2014; Mooney & Candy, 2008). With the lack of population based cohort studies in the literature, the prevalence of dysfunctional breathing is largely an estimate (Kiesel, Rhodes, Mueller, Waninger & Butler, 2017). Two cross sectional studies based at a general practice of 7,033 clients in the United Kingdom showed that approximately 8% of adults without asthma, and who visited a general practitioner, suffered from symptoms associated with dysfunctional breathing (Thomas, McKinley, Freeman, Foy, & Price, 2005). Dysfunctional breathing was more prevalent in women than men (35% versus 20% in those with asthma; 14% versus 2% in those without asthma) and in individuals diagnosed with asthma compared to those without (29% versus 8%) (Thomas, McKinley, Freeman, & Foy, 2001; Thomas et al., 2005). However,

findings from these studies cannot be generalised to the general population, since their samples were relatively small, and participants were recruited from one semi-rural practice. Findings may be different in urban areas. In addition, clinical confirmation of dysfunctional breathing was not carried out.

The most common form of dysfunctional breathing is hyperventilation syndrome (Boulding et al., 2016) in which an individual presents with a range of apparently unrelated physiological symptoms associated with chemical changes (i.e. a reduction of carbon dioxide) in the cardiovascular / circulatory system. The reduced level of carbon dioxide within the bloodstream is the result of an acute or chronic increase in respiratory response (e.g. rate and/or volume) that exceeds the metabolic demands of the body (Lum, 1975). There is no gold standard objective assessment for the diagnosis of dysfunctional breathing/hyperventilation syndrome (Agache et al. 2012). The Nijmegen Questionnaire is used by clinicians for the assessment of symptoms attributed to hyperventilation syndrome as part of a holistic assessment. It does not provoke symptoms that could cause patient distress, in contrast to the hyperventilation provocation test (Howell, 1997). The Nijmegen Questionnaire is a self-reported 16 symptom scale, with response options never (0), rarely (1), sometimes (2), often (3), and very often (4) (Appendix A). A score above 23 out of 64 is a positive screening of hyperventilation syndrome (Garssen et al., 1984; van Doorn, Colla, & Folgering, 1983). The questionnaire is also recommended for the assessment of other dysfunctional breathing patterns (Boulding et al., 2016), however, it has not been validated in these conditions.

An assessment tool needs to be conceptually sound, valid, and reliable for application in various clinical and research settings. However, our previous literature review suggests evidence on the psychometric properties of the Nijmegen Questionnaire is limited (Li Ogilvie & Kersten, 2015). Indeed, only one study investigating structural validity was identified (van Doorn et al., 1983). Structural validity is “the degree to which scores of a measurement instrument are an adequate reflection of the dimensionality of the construct to be measured” (Mokkink et al. 2010, p743). The second identified study had methodological limitations (e.g. the methodologies and procedures used to examine the content validity and reliability of the questionnaire were unclear (van Dixhoorn & Duivenvoorden, 1985)). Content validity can be defined as “the degree to which the content of a measurement instrument is an adequate reflection of the construct to be measured” (Mokkink et al. 2010, p743). As such, there is more work needed to establish the content validity and structural validity of the Nijmegen Questionnaire. Without first establishing content validity, any other validation procedures are unlikely to yield meaningful results (Bond & Fox, 2015; McDowell, 2009). The purpose of this study therefore was to investigate the content and structural validity of the Nijmegen Questionnaire, with the research question: Is the Nijmegen Questionnaire a valid outcome measure for individuals with hyperventilation syndrome? The research findings have the potential to increase confidence in the utilisation of the Nijmegen Questionnaire among clinicians and researchers, empowering users to make relevant inferences from the questionnaire scores and facilitating the process in identifying individuals with hyperventilation syndrome for early physiotherapy intervention.

Methods

This study drew on guidelines for outcome measure development and testing, incorporating qualitative and quantitative research methods (Bowling, 2014; McDowell, 2009; Streiner, Norman, & Cairney, 2015). Content validity was investigated using qualitative descriptive methodology (Sandelowski, 2000) and structural validity was examined using Rasch analysis (Bond & Fox, 2015). The study was approved by the Auckland University of Technology Ethics Committee and the research office at the participating government funded hospital.

Content validity – Qualitative Descriptive Study

Participants and sampling

Patient participants included patients who were diagnosed by a clinician (based on their clinical diagnosis) with hyperventilation syndrome. People were eligible to take part if they were: a) 18 years or older, b) able to communicate in English (verbal and written), and c) able to provide informed consent (verbal and written). Patients were excluded if they had a known organic cardiac, neurological, and/or respiratory disease given the crossover of symptoms could propose a risk in contaminating the research findings. This was consistent with previously published studies associated with the development and validation of the Nijmegen Questionnaire (Garssen et al., 1984; van Dixhoorn & Duivenvoorden, 1985; van Doorn et al., 1983; van Doorn, Folgering, & Colla, 1982). Patient eligibility was determined by examination of their clinic records which contained such details. Clinicians were included if they had experience of working with patients with hyperventilation syndrome. Clinicians were from varied health disciplines (nursing, physiotherapy and medicine).

We intended to use purposeful sampling (Patton, 2002; Sandelowski, 2000) to select patients and clinicians, aiming to recruit individuals from different age, gender, ethnic groups, and clinical disciplines. However, after three months, only one patient had consented to participate. Given this, other recruitment strategies (distribution of study flyers via specialist services mailing list, offering flyers to patients at clinic group sessions, and snowballing sampling) were utilised (with additional ethical approval). Attempting to build on prior research (van Doorn et al., 1983) and to achieve sampling diversity, we aimed to recruit a minimum of six patients and three clinicians. Participants were identified and recruited from respiratory physiotherapy clinics in Auckland, New Zealand. A hospital administrator and physiotherapy colleague distributed or mailed the study flyers. All patient participants had knowledge of the Nijmegen Questionnaire as they had all completed this as part of the previous or ongoing treatment. We did not record how many times they had completed the questionnaire previously.

Data collection

After providing consent, each participant took part in a semi-structured interview (approximately one hour) with the researcher (first author) who is a registered physiotherapist. An interview guide was used (Table 1) to explore the symptoms attributed to hyperventilation syndrome and content validity of the Nijmegen Questionnaire. Interviews were recorded and transcribed verbatim by the researcher.

Data analysis

Interview data were analysed using conventional content analysis, in which coding categories are derived directly from the text data, and which allows the researcher to

focus on the characteristics of language used to illuminate key concepts associated with the phenomenon (Hsieh & Shannon, 2005). The researcher identified data on symptoms attributed to hyperventilation syndrome and the Nijmegen Questionnaire. Symptoms/symptom clusters identified from the interviews that had conceptual congruency with the Nijmegen Questionnaire were grouped together to form categories and sub-categories, before being compared against the Nijmegen Questionnaire items. The primary researcher kept a reflexive journal, reviewed and revised coding strategies and outcomes with co-investigators (NK and PK) throughout the analytical process to stay close to the data as the categories and sub-categories were developed, and to minimise bias.

Structural validity – Rasch Analysis

Sampling

Nijmegen Questionnaires completed by eligible patients who attended the aforementioned clinic between 02/05/2013 and 30/04/2016 were extracted from patient clinical records. For Rasch analyses, reasonably well targeted samples of 108 are reported to have 99% confidence that the estimated item difficulty is within +/-1 ½ logit of its stable value (Linacre, 1994). For poorly targeted samples, 243 are required for this level of confidence. Erring on the side of caution, we aimed to include 250 questionnaires (no upper limit was set for the number of questionnaires per patient). The individual item scores and total scores of the questionnaires made up the data set for analysis. Person characteristics (e.g. age, gender, and ethnicity) were also collected.

Data collection

The individual item scores from the questionnaires were entered into a Microsoft Access database. Total item scores were calculated by a pre-entered formula and the total item scores could not be calculated if there were any missing items. Data entry was checked against the questionnaires. Rasch analysis was carried out using RUMM2030 software (Andrich, Sherridan, & Luo, 2009).

Data analysis

Descriptive statistics for the Nijmegen Questionnaire data set (including summary statistics for personal characteristics: age, gender, and ethnicity) were calculated using IBMSPSS Statistics 22 (IBM Corp, 2013). Rasch analysis incorporated the relevant steps outlined below (Kersten & Kayes, 2011; Medvedev et al., 2017; Siegert, Tennant, & Turner-Stokes, 2010):

1. Testing of overall data fit to the Rasch model: The item-trait interaction chi-square probability should be non-significant.
2. Checking of person fit to the Rasch model: Fit residuals should be within the range of +/- 2.5, with a non-significant item fit chi-square probability, the mean fit residual should be close to zero with a standard deviation value close to one.
3. Checking of individual item fit for their fit to the Rasch model: Fit residuals should be within the range of +/- 2.5 with a non-significant item fit chi-square probability, the mean fit residual should be close to zero with a standard deviation value close to one.
4. Identifying item(s) with poor fit to the Rasch model (using fit statistics outlined under 2.)

5. Identifying local dependency / dependencies between items from the residual correlation matrix: the residual correlation should be < 0.2 above the mean residual correlation.
6. Checking if the item response categories work as intended. The validity of the five response category structure of each item was assessed by examining if the response thresholds were ordered: thresholds are the points on the scale where the probabilities of someone giving a response of either 0 or 1, and 1 or 2 (and so forth) are equally likely. When the response categories do not show a logical progression across the trait being measured disordered thresholds are observed. In such instances, responses categories can be collapsed to solve this problem.
7. Analysing Differential Item Function (DIF) for personal characteristics (e.g. age, gender, ethnicity, and assessment [time one; time two etc.]): Absence of DIF is shown if the analysis of variance (ANOVA) test is non-significant.
8. Testing of unidimensionality: Fewer than 5% of independent t-test on estimates from testlets created from items with high positive and high negative loadings on the first principal component of the residuals should be significant (the 95% Confidence Interval [CI] should include 5%).
9. (Potentially) Modifying the original scale by:
    a. deleting item(s) with poorest fit to the Rasch model;
    b. combining items with local dependencies;
    c. re-scoring item(s) with disordered threshold(s).
10. Re-testing individual item fit and overall fit to the Rasch model
11. Distribution analysis of the participant-item thresholds.

Results

Participant characteristics

Six patients (all females) aged 26 to 64 years and four clinicians (three females) aged 54-58 were interviewed. Age was undisclosed for one clinician. Ethnic identities for patients included Chinese, Māori, New Zealand European, and South African. Clinicians' ethnicities included Chinese, European, and New Zealand European.

Symptoms of hyperventilation syndrome and content validity

Table 2 presents the symptoms/symptom clusters (total of 46), symptom categories (total of 3) and sub-categories (total of 12) identified from interview data. Based on evaluation of conceptual congruency and language consistency, only one existing Nijmegen Questionnaire item (stiff fingers or arms) did not match with interview data. The other 15 items (94%) matched with interview data at a conceptual level, albeit with some inconsistencies in the language used to describe the symptoms. Table 3 contains excerpts from interview data as they relate to questionnaire items. Differences were noted between patients and clinicians in terms of the words or phrases used (e.g. [patients] "You're not breathing in a good rhythm" versus [clinician] "So the mechanics can include apical pattern of breathing, altered inspiratory expiratory ratio..."). Despite some minor discrepancies in language, these findings suggest the Nijmegen Questionnaire meets the criteria for content validity given that 94% of the items are representative of symptoms attributed to hyperventilation syndrome based on the perspectives of patients and clinicians with experience of the condition. There were symptoms identified from the interviews that were not addressed by the Nijmegen Questionnaire, 68% of which were in sub-categories with other symptoms matched by questionnaire items.

Questionnaires characteristics

Data from 239 questionnaires completed by 159 patients (1 to 5 questionnaires per patient) were extracted for the Rasch analysis. Of the 239 questionnaires, 73% were completed by females. The ethnic characteristics of the patients included: New Zealand European (41%), Asian (28%), Pacific Islander (11%), Māori (8%), and Other (12%). Age characteristics were divided into three groups: 15-46 years (40%), 47-57 years (28%), and >57 years (32%). Of the 159 patients, 72% were females. The mean age was 51 years with a standard deviation of 16 (range 15-90).

Rasch analysis and structural validity

Table 4 shows the distribution of response frequencies of the 239 questionnaires, including information on missing data. Twelve items showed a floor effect (i.e. >25% of patients scoring 0 = Never). The data did not fit the Rasch model with mean item fit residual of 0.410 and standard deviation of 1.499 (Table 5). The item-trait interaction chi-square was significant with probability of <0.001, demonstrating the lack of fit (Table 5, Analysis 1). One misfitting item (NQ14 *cold hands or feet*) was identified with an item fit residual of 4.58 (acceptable range = +/- 2.5). This item was under discriminating and shown to have uniform DIF by gender (Figure 1).

Residual correlations should be smaller than 0.2 above the average residual correlation (in this instance -0.063 + 0.2 = 0.137). High correlations between the residuals indicated local dependency between six sets of items (Table 6), suggesting that item responses of the Nijmegen Questionnaire depend not only on the severity of the symptoms of hyperventilation syndrome being measured, but on responses to other questionnaire items. The Nijmegen Questionnaire is unidimensional given that 5.1% of *t*-tests were significant (95% CI 2.3% to 7.8%, Table 5, Analysis 1).

Examination of the category probability curves indicated disordered thresholds for all 16 items.

The misfitting item NQ14 was deleted and the analysis repeated with the remaining data (Table 5, Analysis 2). The mean item fit residual was 0.39 with a standard deviation of 1.15. The item-trait interaction chi-square was not significant with probability of 0.016 (greater than the Bonferroni adjusted *p* value of 0.0033), indicating fit to the Rasch model. Item NQ9 (*bloated feeling in stomach*) had an item fit residual of 2.76, just outside the acceptable range. This item was also under discriminating, though not to the extent NQ14 was. The remaining 14 items demonstrated good fit to the Rasch model. All 15 items were invariant (i.e. unbiased, no DIF) across different age, gender, and ethnic groups, at initial and repeated assessment(s). Local dependency was found between the same clusters of items identified previously. The 15-item Nijmegen Questionnaire was found to remain unidimensional. However, as with the 16-item scale, all items had disordered thresholds. After collapsing response options (Table 5, Analysis 3) using strategies outlined in Table 7, the number of disordered thresholds were reduced over three rescoring stages. Ordered thresholds were achieved for all 15 items by combining the response categories *sometimes* and *often*. Locally dependent items were combined into new super items (testlets), removing the influence of local dependencies (Table 5, Analysis 4). Following this, the average fit residual statistics had a mean of 0.06 and standard deviation of 0.86. The item-trait interaction chi-square probability was not significant at 0.205. With only 1.8% of significant *t*-tests, the scale remained unidimensional. A conversion table (Table 8) was created, allowing the conversion of ordinal to interval data for parametric analyses and clinical use. This works by calculating the total score on a completed questionnaire,

294 excluding item 14, and then use the table to look up to convert the raw (ordinal)
295 score in column 1 to the new equivalent interval score in column 3.

296 Discussion

297 Our study evaluated the content and structural validity of the Nijmegen
298 Questionnaire. To our knowledge, this is the first study to involve patients in content
299 validity investigation for the questionnaire. It is also the first time that Rasch analysis
300 is utilised in the evaluation of structural validity of the Nijmegen Questionnaire. Our
301 study results demonstrated that 94% of the questionnaire items matched partly or
302 fully with the interview data, representing both patients' and clinicians' view on
303 symptoms of hyperventilation syndrome in relation to questionnaire content, though
304 perhaps not fully. Stiff fingers or arms was the only item (from 16) that did not map
305 onto interview data. A total of 46 symptoms/symptom clusters were identified in our
306 study, compared to a total of 45 symptoms reported by patients in the van Doorn and
307 colleague's first study (1982). We were unable to compare our additional
308 symptoms/symptom clusters with their study, however, as they only reported the
309 content of the final 16 symptoms that now make up the Nijmegen Questionnaire.
310 This study provides a point of reference for symptoms of hyperventilation syndrome,
311 as perceived by patients who experience hyperventilation syndrome first-hand, and
312 clinicians working with this population. It is worth noting that while the items were
313 conceptually congruent with interview data, there were some language
314 inconsistencies between the existing items and the symptoms/symptom clusters
315 identified. This has also been observed in the literature (Grossman & de Swart,
316 1984; Ruiter, Garssen, Rijken, & Kraaimaat, 1989; van Doorn et al., 1982). Future
317 research might involve refining item wording so that items resonate with the

language patients would use to describe their symptoms, given that any refinements would need to be tested against the Rasch model.

The Rasch analysis findings showed that the current Nijmegen Questionnaire did not fit the Rasch model and therefore did not meet criteria for structural validity. The questionnaire was not unidimensional and all 16 items demonstrated disordered thresholds. *Cold hands or feet* (NQ14) was identified as a poorly fit item, illustrating bias in its function when assessing hyperventilation syndrome between male and female patients. *Bloated feeling in stomach* (NQ9) was another item with a poor fit and under discriminating, after deleting NQ 14. However, it was retained due to the absence of bias in terms of item function in person variables. This suggested that *bloated feeling in stomach* was valid in assessing hyperventilation syndrome. The systematic rescoring of response options and the merging of items with congruent meanings into testlets resulted in the revised 15-item version of the Nijmegen Questionnaire, meeting straight criteria for structural validity. A previous study (van Dixhoorn & Duivenvoorden, 1985) utilised non-metric principal components analysis (a parametric statistical technique) to evaluate structural validity of the Nijmegen Questionnaire. However, their results cannot be compared directly with the current study results because they used parametric statistical techniques, which are not suited to ordinal data (Bond & Fox, 2015; Streiner et al., 2015). However, prior results concerning construct validity can be extrapolated and interpreted with these study results. Van Dixhoorn and Duivenvoorden (1985) identified three questionnaire components: shortness of breath, peripheral tetany, and central tetany. The identification of this underlying relationship between variables were consistent with the discovery of local dependencies among the current items of the Nijmegen Questionnaire in this study. Some of the local dependencies identified were noted

within the shortness of breath and central tetany components. This suggests that the symptoms represented by these items were scored not just based on the severity of hyperventilation syndrome related symptoms, but on the score for another item on the scale also. The locally dependent items were representing symptoms of similar nature. One item (NQ16 *feeling of anxiety*) was omitted from the van Doorn et al's (1982) validation study. This item was found to be locally dependent with *feeling tense* (NQ2) and *feeling confused* (NQ5). Van Dixhoorn and Duivenvoorden's (1985) decision to omit *feeling of anxiety* (NQ16) was not supported by our study results. *Stiff fingers or arms* (NQ12) did not match with any participant-identified symptoms. However, it was found to be locally dependent with *tingling fingers* (NQ10) which was fully conceptually and linguistically congruent with symptom identified by participants. Regardless of the lack of reporting by study participants, the fact that NQ12 was locally dependent suggests it measures something very similar to NQ10. Item NQ14 (*cold hands or feet*) was only partly congruent with interview findings. In addition, it was a misfitting item as highlighted by the Rasch analysis which resulted in it being deleted. Thus, both the interview and Rasch analysis findings from this study supported a 15-item version of the Nijmegen Questionnaire as a valid screening tool for hyperventilation syndrome.

Research and clinical implications

Interview findings revealed one existing item that appeared to be a poor match to the symptoms of hyperventilation syndrome. Additionally, a number of symptoms identified by participants are not captured by existing items of the Nijmegen Questionnaire. The reason for the mismatch between items and symptoms could be multifaceted. On the one hand, the interpretation and description of these symptoms

varied between patients and clinicians. This could cause symptoms to be missed or misinterpreted by both parties in the clinical encounter. The Nijmegen Questionnaire does contain a majority of items that reflect symptoms of hyperventilation syndrome. While the questionnaire is structurally valid for repeat assessment (as there was no bias over time points in this study), no validation process to date has proved the ability of this questionnaire in measuring change (e.g. treatment effectiveness on hyperventilation syndrome). It is important to be aware of this when interpreting results from more than one assessment for individual patients. The same caution needs to be applied when using the Nijmegen Questionnaire as an outcome measure in research.

Strengths and limitations

By involving both patients and clinicians, this study met the criteria for the evaluation of content validity as described by the Scientific Advisory Committee of the Medical Outcome Trust (2002). Studies employing the qualitative descriptive methodology are able to produce findings that are transferable to population with similar characteristics as the study participants (Sandelowski, 1995, 2000). The COSMIN checklist identifies several criteria to assess the methodological quality of measurement studies (Mokkink et al., 2010; Terwee et al., 2012). A self-assessment of the current study suggested that it meets all the criteria identified as critical to content validity, achieving an *excellent* rating (Table 9).

The interview data may be limited by the small sample size and despite the various adjustments made in the effort to recruit male participants, there was also a lack of male patient interview participants and only one in the clinician group. Although more women than men suffer from hyperventilation syndrome and more women are

393 treated at the recruitment locality, the study findings regarding content validity have
394 limited transferability to a male population. The Nijmegen Questionnaire is a
395 suggested screening tool for hyperventilation syndrome, based on reported
396 symptoms. However, these symptoms are not exclusive to individuals with
397 hyperventilation syndrome. It was not feasible to exclude patients with psychiatric
398 and/or psychological disorders due to either personal preferences or public health
399 policies around disclosure. The mental health background of patients from the study
400 was unexplored and could have affected their symptom reporting.

401

402 Conclusion

403 The revised 15-item Nijmegen Questionnaire is an outcome measure that is suitable
404 for its purpose in screening for hyperventilation syndrome in clinical and research
405 settings with standards for application in place. The utilisation of the conversion table
406 is recommended for converting ordinal raw scores to interval data when using the
407 Nijmegen Questionnaire especially when parametric testing is indicated. It should be
408 used in conjunction with other subjective and objective measures when assessing for
409 hyperventilation syndrome.

410

411 KEY POINTS

412 1. This paper demonstrates content validity of the Nijmegen Questionnaire for
413 hyperventilation syndrome, involving patients (in addition to clinicians) in the
414 validation process for the first time.

2. The structural validity of the Nijmegen Questionnaire was explored using Rasch analysis (first in the literature), in line with the principles of outcome measure development and testing for ordinal questionnaire data.
3. This paper includes a revised 15-item Nijmegen Questionnaire and a conversion table for transforming raw (ordinal) total questionnaire scores to interval scores.
4. Physiotherapists should use the revised 15-item Nijmegen Questionnaire for clinical and research purposes.

IMPACT STATEMENT

The revised Nijmegen Questionnaire is an outcome measure that is suitable for its purpose in screening for hyperventilation syndrome in clinical and research settings with standards for application in place. Physiotherapists should use this revised tool to aid communication with patients and to evaluate their services.

DISCLOSURES

Funding source: Cardiothoracic Special Interest Group, Physiotherapy New Zealand; Counties Manukau Health

Conflicts of interest: The authors have no conflicts of interest to declare.

PERMISSIONS

436 Ethics: Ethical approval was obtained from Auckland University of Technology Ethics
437 Committee (15/197). Consent was obtained from all participants who took part in
438 interviews.

439

440 ACKNOWLEDGEMENTS

441 We thank study participants for their time and contribution, and the cardiorespiratory
442 team for their assistance with recruitment.

443

444 **Appendix A** *Nijmegen Questionnaire*
445

	Never (0)	Rarely (1)	Sometimes (2)	Often (3)	Very often (4)
1. Chest pain					
2. Feeling tense					
3. Blurred vision					
4. Dizzy spells					
5. Feeling confused					
6. Faster / deeper breathing					
7. Short of breath					
8. Tight feelings in the chest					
9. Bloated feeling in the stomach					
10. Tingling fingers					
11. Unable to breathe deeply					
12. Stiff fingers or arms					
13. Tight feelings around the mouth					
14. Cold hands or feet					
15. Palpitations					
16. Feelings of anxiety					

446

447

Table 1 Interview guide

Starting questions for patients
How would you describe what it feels like to have hyperventilation syndrome?
Can you tell me about the symptoms that you associate with this condition?
How would someone know that you were experiencing hyperventilation syndrome if they were watching you?
What would they miss?
Could you think of a specific incident where you were experiencing hyperventilation syndrome and tell me about those symptoms?
Starting questions for clinicians
How would you describe the signs and symptoms of hyperventilation syndrome?
How do you determine if someone is suffering from hyperventilation syndrome?
What other symptoms would a family member / friend / support person identify from an individual with hyperventilation syndrome?
Any cases that stood out to you that are different from what you told me already?
Questions relating to the Nijmegen Questionnaire for patients and clinicians
From your perspective, what are your views on the appropriateness of the questionnaire? - appropriateness of individual complaints - appropriateness of the response options - appropriateness of the language use - any important areas that are not currently included
If you were to use this questionnaire, do you think it would give an accurate account of the symptoms associated with hyperventilation syndrome? Why?

Table 2 Symptom categories, sub-categories and symptoms

Sub-categories		Category 1: Breathing Symptoms	Interview data match with NQ* item number (F=Full; P=Partly)† Item text
Altered capacity	1	Hyperventilating / Over breathing	NQ06 (P) Faster or deeper breathing
	2	Breathing more / Deep breathing	NQ06 (P) Faster or deeper breathing
	3	Breathing fast / Shallow breathing	NQ06 (P) Faster or deeper breathing
	4	Difficulty filling lungs / taking deep breaths	NQ11 (P) Unable to breathe deeply
Altered pattern	1	Upper chest breathing	
	2	Noisy / Heavy breathing	
	3	Altered rhythm of breathing	
	4	Breath-holding	
Global changes and	1	Gasp / Pant / Puff	
difficulties	2	Short of breath	NQ07 (F) Short of breath
	3	Air hunger	
	4	Sigh / Yawn	
	5	Difficulty breathing	
		Category 2: Psychological Symptoms	

Feelings	1	Anxiety / Fear / Panic	NQ16 (F) Feeling of anxiety
	2	Aggravating / Agitated / Stressed / Rushed	
	3	Chaotic / Confused / Overwhelmed / Frustration	NO05 (P) Feeling confused
	4	Poor tolerance / Hypervigilance	
	5	Uneasy / Feeling different / Not feeling so good / Something is always at the back of your mind	
	6	Disconnected	
Thoughts	1	Out of control / balance	
	2	Worry	
		Category 3: Physical Symptoms	
Bodily regulations	1	Feeling hot / sweaty	
	2	Constipation / Irritable bowel	NQ09 (P) Bloated feeling in stomach
	3	Sleep disturbances	
Bodily sensations	1	Dizziness / Faintness / Light-headedness	NQ03 (P) Blurred vision; NQ04 (F) Dizzy spells

	2	Passing out / Physical collapse / Vision goes dark	
	3	Tiredness	
Head / face / mouth / throat	1	Headache	
	2	Pressure / Exploding feeling	
	3	Frowning / Facial expression	
	4	Pale	
	5	Tight feeling in the throat	NQ13 (P) Tight feelings around mouth
	6	Gritting teeth	NQ13 (P) Tight feelings around mouth
	7	Dry mouth	
	8	Clearing throat	
Heart / chest	1	Heart palpitations / beats fast / racing	NQ15 (F) Palpitations
	2	Chest restriction / tightness	NQ08) (F) Tight Feeling in chest)
	3	Chest pain	NQ01 (F) Chest Pain
Fingers / hands	1	Paraesthesia / Tingling	NQ10 (F) Tingling fingers
	2	Sweaty fingers / palm	NQ14 (P) Cold hands or feet

Muscle / Posture	1	Tense muscles	NQ02 (P) Feeling tense
	2	Aches and pains	
	3	Postural changes	
Speech / Voice	1	Voice changes	
	2	Talking more / faster	
	3	Poor breathing control	

***** NQ = Nijmegen Questionnaire item;

† F = Full match (consistent language, conceptually congruent); P = Part match (some discrepancy in language or not entirely conceptually congruent).

Table 3 Comparison between Nijmegen Questionnaire items and excerpts from interview data

Items	Excerpts
Chest pain	"The chest pain kind of group of symptoms." ([C]Leena)
Feeling tense	"Your muscles would tense up." ([P]Cathy)
Blurred vision	"You feel like you're going to pass out." ([P]Dora)
Dizzy spells	"Sometimes the dizziness just last despite me trying different things to calm my breathing down." ([P]Eva)
Feeling confused	"…their world feels…chaotic or confused…" ([C]Jessica)
Faster or deeper breathing	"They're breathing fast." ([C]Kelvin)
Short of breath	"I do feel like short of breath like I'm not getting enough oxygen." ([P]Eva)
Tight feelings in chest	"It's just kind of…tight, more at the bottom." ([P]Becky)
Bloated feeling in stomach	"The feeling of constipation or irritable bowel." ([C]Jessica)
Tingling fingers	"Some people have sort of tingling in their hands." ([C]Margo)
Unable to breathe deeply	"I can't take a deep breath in and I can't completely fill up my lungs." ([P]Abby)
Stiff fingers or arms	*Nil
Tight feelings round mouth	"Tightening around…your throat." ([P]Cathy)
Cold hands or feet	"I've always got…sweaty palms / fingers." ([P]Eva)
Palpitations	"…[patients] come in …saying they have palpitations." ([C]Kelvin)
Feeling of anxiety	"A general sort of sense of anxiety." ([C]Margo)

475 *No match. [C] = Clinician; [P] = Patient

Table 4 Distribution of Response Frequencies of the Nijmegen Questionnaire

		Response categories					
		Never (0)	Rare (1)	Sometimes (2)	Often (3)	Very often (4)	Missing
Item	Description	Freq (%)	Freq (%)	Freq (%)	Freq (%)	Freq (%)	Freq (%)
NQ1	Chest pain	79 (33.1)	47 (19.7)	65 (27.2)	27 (11.3)	20 (8.4)	1 (0.4)
NQ2	Feeling tense	29 (12.1)	24 (10.0)	87 (36.4)	60 (25.1)	37 (15.5)	2 (0.8)
NQ3	Blurred vision	96 (40.2)	39 (16.3)	57 (23.8)	32 (13.4)	15 (6.3)	-
NQ4	Dizzy spells	65 (27.2)	40 (16.7)	76 (31.8)	40 (16.7)	17 (7.1)	1 (0.4)
NQ5	Feeling confused	94 (39.3)	51 (21.3)	52 (21.8)	24 (10.0)	18 (7.5)	-
NQ6	Faster or deeper breathing	40 (16.7)	41 (17.2)	77 (32.2)	48 (20.1)	32 (13.4)	1 (0.4)
NQ7	Short of breath	45 (18.8)	33 (13.8)	78 (32.6)	49 (20.5)	33 (13.8)	1 (0.4)
NQ8	Tight feelings in chest	62 (25.9)	40 (16.7)	67 (28.0)	38 (15.9)	31 (13.0)	1 (0.4)
NQ9	Bloated feeling in stomach	67 (28.0)	35 (14.6)	65 (27.2)	36 (15.1)	36 (15.1)	-
NQ10	Tingling fingers	94 (39.3)	42 (17.6)	55 (23.0)	24 (10.0)	22 (9.2)	2 (0.8)
NQ11	Unable to breathe deeply	80 (33.5)	42 (17.6)	55 (23.0)	34 (14.2)	26 (10.9)	2 (0.8)
NQ12	Stiff fingers or arms	99 (41.4)	40 (16.7)	47 (19.7)	27 (11.3)	26 (10.9)	-
NQ13	Tight feelings around mouth	153 (64.0)	38 (15.9)	25 (10.5)	11 (4.6)	11 (4.6)	1 (0.4)
NQ14	Cold hands or feet	81 (33.9)	32 (13.4)	46 (19.2)	33 (13.8)	47 (19.7)	-
NQ15	Palpitations	63 (26.4)	44 (18.4)	82 (34.3)	30 (12.6)	20 (8.4)	-
NQ16	Feeling of anxiety	35 (14.6)	38 (15.9)	72 (30.1)	51 (21.3)	43 (18.0)	-

Note. Freq = frequency

476

Table 5 Summary of Fit Statistics of the Nijmegen Questionnaire to the Rasch Model

Analysis Number	Item fit residual	Person fit residual	Chi-square interaction	Chi-square Probability	PSI (without extremes)	α (without extremes)	Tests of unidimensionality
	Mean (SD)	Mean (SD)	Value (df)	*p*			Significant *t*-test (95% confidence interval)
One *	0.41 (1.50)	-0.27 (1.61)	109.4 (48)	0.000	0.880	0.890	5.1% (2.3 to 7.8)
Two †	0.39 (1.15)	-0.31 (1.58)	67.8 (45)	0.016	0.879	0.891	5.5% (2.7 to 8.3)
Three ‡	0.06 (0.97)	-0.22 (1.20)	41.9 (45)	0.604	0.826	0.869	5.8% (2.9 to 8.6)
Four §	0.06 (0.86)	-0.21 (1.02)	36.1 (45)	0.205	0.789	0.809	1.8% (1.1 to 4.6)

Note. SD = standard deviation; df = degrees of freedom; *p* = probability; PSI = Person Separation Index; α = Cronbach's alpha.

***** Fit to the Rasch model of all 16 items.

† Fit to the Rasch model after deleting item NQ14.

‡ Fit to the Rasch model after rescoring response categories for items with disordered thresholds.

§ Fit to Rasch model after merging of items.

477

Table 6 Summary of Local Dependencies of the Nijmegen Questionnaire

Analysis Number	Item		Locally dependent with:	
	Item	Description	Item	Description
One * and Two †	1	Chest pain	8	Tight feelings in chest
	2	Feeling tense	5, 16	Feeling confused, Feeling of anxiety
	3	Blurred vision	4	Dizzy spells
	6	Faster or deeper breathing	7	Short of breath
	7	Short of breath	11	Unable to breathe deeply
	10	Tingling fingers	12	Stiff fingers or arms
Three ‡	1	Chest pain	8	Tight feelings in chest
	2	Feeling tense	5, 16	Feeling confused, Feeling of anxiety
	6	Faster or deeper breathing	7	Short of breath
	10	Tingling fingers	11, 12	Unable to breathe deeply, Stiff fingers or arms
Four §	No local dependency			

***** Fit to the Rasch model of all 16 items. **†** Fit to the Rasch model after deleting item NQ14. **‡** Fit to the Rasch model after rescoring response categories for items with disordered thresholds. **§** Fit to Rasch model after merging of items.

Table 7 Rescore strategy for response categories of the Nijmegen Questionnaire

Strategy	Response options				
	Never	Rare	Sometimes	Often	Very Often
	(0)	(1)	(2)	(3)	(4)
1st rescore	0	1	1	2	3
2nd rescore	0	0	1	2	3
3rd rescore	0	0	1	1	2

479

Table 8 Conversion table for the Nijmegen Questionnaire

Raw total score	Logit	Interval score
0	-3.438	0.00
1	-2.710	4.62
2	-2.234	7.64
3	-1.923	9.62
4	-1.690	11.10
5	-1.502	12.29
6	-1.344	13.30
7	-1.207	14.17
8	-1.085	14.94
9	-0.975	15.64
10	-0.875	16.27
11	-0.782	16.86
12	-0.696	17.41
13	-0.616	17.92
14	-0.540	18.40
15	-0.469	18.85
16	-0.400	19.29
17	-0.334	19.71
18	-0.270	20.11
19	-0.208	20.51
20	-0.148	20.89
21	-0.088	21.27

480

Table 9 COSMIN checklist for Content Validity

Questions to determine if a study meets the standards for methodological quality	Excellent	Good	Fair	Poor
1 Was there an assessment of whether all items refer to the relevant aspects of the construct to be measured?	✓			
2 Was there an assessment of whether all items are relevant for the study population?	✓			
3 Was there an assessment of whether all items are relevant for the purpose of the measurement instrument?	✓			
4 Was there an assessment of whether all items together comprehensively reflect the construct to the measured?	✓			
5 Were there any important flaws in the design or methods of the study?	✓			

Note. The definition of *excellent* for different questions are: 1 = Assessed if all items refer to relevant aspects of the construct to be measured. 2 = Assessed if all items are relevant for the study population in adequate sample size (≥ 10). 3 = Assessed if all items are relevant for the purpose of the application. 4 = Assessed if all items together comprehensively reflect the construct to be measured. 5 = No other important methodological flaws in the design or execution of the study.

481

Figure 1. Differential Item Functioning for item NQ14 Cold hands or feet.

See file which we have uploaded separately.

484 References

485 Agache, I., Ciobanu, C., Paul, G., & Rogozea, L. (2012). Dysfunctional breathing
486 phenotype in adults with asthma - incidence and risk factors. *Clinical and*
487 *Translational Allergy*, 2, Article number: 18. doi:10.1186/2045-7022-2-18

488 Andrich, D., Sherridan, B., & Luo, G. (2009). For analysis assessment and attitude
489 questionnaire data RUMM2030 Perth: RUMM Laboratory Pty Ltd.

490 Bond, T. G., & Fox, C. M. (2015). *Applying the Rasch model: fundamental measurement in*
491 *the human sciences*. New York: Routledge.

492 Boulding, R., Stacey, R., Niven, R., & Fowler, S. J. (2016). Dysfunctional breathing: a
493 review of the literature and proposal for classification. *European Respiratory*
494 *Review, 25*(141), 287. doi:10.1183/16000617.0088-2015

495 Bowling, A. (2014). *Research methods in health: investigating health and health services*.
496 Maidenhead: Open University Press.

497 Chaitow, L., Morrison, D., & Gilbert, C. (2014). *Recognising and treating breathing*
498 *disorders: a multidisciplinary approach*. Edinburgh: Churchill Livingstone.

499 Garssen, B., Colla, P., van Dixhoorn, P., Folgering, H., Stoop, A., & de Swart, J. (1984).
500 Het herkennen van het hyperventilatiesyndroom (Recognising the hyperventilation
501 syndrome). *Medisch Contact, 35*, 1122-1124.

502 Grossman, P., & de Swart, J. (1984). Diagnosis of hyperventilation syndrome on the basis
503 of reported complaints. *Journal of Psychosomatic Research, 28*(2), 97-104.
504 doi:10.1016/0022-3999(84)90001-1

505 Howell, J. B. L. (1997). The hyperventilation syndrome: A syndrome under threat? *Thorax,*
506 *52*(SUPPL. 3), S30-S34. doi:10.1136/thx.52.2008.S30

507 Hsieh, H., & Shannon, S. (2005). Three approaches to qualitative content analysis.
508 *Qualitative Health Research, 15*(9), 1277-1288. doi:10.1177/1049732305276687

509 IBM Corp. (2013). IBM SPSS Statistics for Windows Version 22.0. Armonk: IBM Corp.

Kersten, P., & Kayes, N. (2011). Outcome measurement and the use of Rasch analysis, a statistics-free introduction. *New Zealand Journal of Physiotherapy, 39*(2), 92-99. ISSN: 0303-7193

Kiesel, K., Rhodes, T., Mueller, J., Waninger, A., & Butler, R. (2017). Development of a screening protocol to identify individuals with dysfunctional breathing. *The International Journal of Sports Physical Therapy, 12*(5), 774-786. doi:10.16603/ijspt20170774

Li Ogilvie, V., & Kersten, P. (2015). A critical review of the psychometric properties of the Nijmegen Questionnaire for hyperventilation syndrome. *New Zealand Journal of Physiotherapy, 43*(1), 3-10. ISSN: 0303-7193

Linacre, J. (1994). Sample size and item calibration stability. *Rasch Measurement Transactions, 7*(4), 328.

Lum, L. (1975). Hyperventilation: The tip and the iceberg. *Journal of Psychosomatic Research, 19*(5), 375-383. doi:10.1016/0022-3999(75)90017-3

McDowell, I. (Ed.) (2009). *Measuring health: a guide to rating scales and questionnaires*. New York: Oxford University Press.

Medvedev, O., Krageloh, C., Hill, E., Billington, R., Siegert, R., Webster, C., . . . Henning, M. (2017). Rasch analysis of the Perceived Stress Scale: transformation from an ordinal to a linear measure. *Journal of Health Psychology*, 1-12. doi:10.1177/1359105316689603

Mokkink, L., Terwee, C., Patrick, D., Alonso, J., Stratford, P., Knol, D., & de Swart, J. (2010). The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. *Quality of Life Research, 19*(4), 539-549. doi:http://www.biomedcentral.com/1471-2288/10/22

Mokkink, L. B., Terwee, C. B., Patrick, D. L., Alonso, J., Stratford, P. W., Knol, D. L., . . . de Vet, H. C. W. (2010). The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *Journal of Clinical Epidemiology, 63*(7), 737-745. doi:10.1016/j.jclinepi.2010.02.006

Mooney, S., & Candy, S. (2008). The real cost of effective treatment. A single case study of a patient with hyperventilation syndrome. *New Zealand Journal of Physiotherapy, 36*(2), 88.

Patton, M. (2002). *Qualitative research and evaluation methods*. Thousand Oaks: Sage.

Ruiter, C., Garssen, B., Rijken, H., & Kraaimaat, F. (1989). The hyperventilation syndrome in panic disorder, agoraphobia and generalised anxiety disorder. *Behaviour Research and Therapy, 27*(4), 447-452. doi:10.1016/0005-7967(89)90015-6

Sandelowski, M. (1995). Focus on research methods: sample size in qualitative research. *Research in Nursing & Health, 18*(2), 179-183. doi:10.1002/nur.4770180211

Sandelowski, M. (2000). Focus on research methods. Whatever happened to qualitative description? *Research in Nursing & Health, 23*(4), 334-340. doi:10.1002/1098-240X(200008)23:4<334::AID-NUR9>3.0.CO;2-G

Scientific Advisory Committee of the Medical Outcomes Trust. (2002). Assessing health status and quality-of-life instruments: attributes and review criteria. *Quality of Life Research, 11*(3), 193-205. doi:10.1023/A:1015291021312

Siegert, R., Tennant, A., & Turner-Stokes, L. (2010). Rasch analysis of the Beck Depression Inventory-II in a neurological rehabilitation sample. *Disability and Rehabilitation, 32*(1), 8-17. doi:10.3109/09638280902971398

Streiner, D., Norman, G., & Cairney, J. (2015). *Health measurement scales: a practical guide to their development and use*. New York: Oxford University Press.

Terwee, C., Mokkink, L., Knol, D., Ostelo, R., Bouter, L., & de Vet, H. (2012). Rating the methodological quality in systematic reviews of studies on measurement properties: a scoring system for the COSMIN checklist. *Qualitity of Life Research, 21*, 651-657. doi:10.1007/s11136-011-9960-1

Thomas, M., McKinley, R., Freeman, E., & Foy, C. (2001). Prevalence of dysfunctional breathing in patients treated for asthma in primary care: cross sectional survey. *British Medical Journal, 322*(7294), 1098-1100. doi:10.1136/bmj.322.7294.1098

Thomas, M., McKinley, R., Freeman, E., Foy, C., & Price, D. (2005). The prevalence of dysfunctional breathing in adults in the community with and without asthma. *Primary Care Respiratory Journal, 14*, 78-82. doi:10.1016/j.pcrj.2004.10.007

van Dixhoorn, P., & Duivenvoorden, H. (1985). Efficacy of Nijmegen Questionnaire in recognition of the hyperventilation syndrome. *Journal of Psychosomatic Research, 29*(2), 199-206. doi:10.1016/0022-3999(85)90042-X

van Doorn, P., Colla, P., & Folgering, H. (1983). A questionnaire for hyperventilation syndrome [in Dutch]. *De Psycholoog, 18*(10), 573-577.

van Doorn, P., Folgering, H., & Colla, P. (1982). Control of the end-tidal pCO2 in the hyperventilation syndrome: effects of biofeedback and breathing instructions compared. *Bullentin Europeen De Physiotherpathologie Respiratoire, 18*(6), 829-836.