Measuring upper limb disability for patients with neck pain: evaluation of the feasibility of the Single Arm Military Press (SAMP) test

ALRENI, Ahmad Salah Eldin, ABDO ABOALMATY, Heba Roohy, DE HERTOGH, Willem, MEIRTE, Jill, HARROP, Deborah <http://orcid.org/0000-0002-6528-4310> and MCLEAN, Sionnadh <http://orcid.org/0000-0002-9307-8565>

Measuring upper limb disability for patients with neck pain: Evaluation of the feasibility of the single arm military press (SAMP) test

Ahmad Salah Eldin Alreni [b,] *, Heba Roohy Abdo Aboalmaty [a], Willem De Hertogh [b], Jill Meirte [b], Deborah Harrop [c], Sionnadh Mairi McLean [c]

[a] Department of Sports Training and Kinesiology, Tanta University El-Gaish Street, Tanta Qism 2, Tanta, Garbia Government, Egypt Faculty of Medicine and Health Science,
[b] University of Antwerp, Campus Drie Eiken, Universiteitsplein 1, D.R.319, 2610, Wilrijk, Belgium Centre for Health and Social Care Research, Faculty of Health and Wellbeing,
[c] Sheffield Hallam University, Sheffield, United Kingdom

Abstract

Background: Non-specific neck pain (NSNP) is frequently associated with upper limb disability (ULD). Consequently, evaluation of ULD using an outcome measure is necessary during the management of patients with NSNP. The Single Arm Military Press (SAMP) test is a performance-based ULD measure developed for populations with neck pain. During the SAMP test, patients are asked to repeatedly lift a weight above their head for 30 s. The number of repetitions is counted. Its clinical utility in a patient group is still unknown.

Objective: This study investigates the feasibility of the SAMP test from patients and clinicians' perspectives. Methods: Seventy female patients with NSNP were randomly allocated into one of three groups. Participants in each group completed the SAMP test using one of three proposed weights (½kg, 1 kg or 1½kg). The feasibility of the SAMP test was established using structured qualitative exit feedback interviews for patients and administrating clinicians.

Results: Participants using ½kg achieved the highest number of repetitions, but a high proportion reported the weight as extremely light, whereas those who tested using the 1½kg achieved the lowest number of repetitions and participants reported the weight as being heavy. Participants tested using 1 kg achieved an average number of repetitions and a high proportion reported the weight as acceptably heavy. Clinicians and patients reported that the SAMP test was efficient and convenient.

Conclusion: The 1 kg SAMP test is feasible for use in female patients with NSNP. The measurement properties of the SAMP test should be determined in a patient group.

Keywords:

Neck pain Upper limb function Disability Outcome measure Feasibility

1. **Background**

Neck pain is one of the most common musculoskeletal conditions that causes pain and disabilities (Hogg-Johnson et al., 2008). It affects approximately 70% of the population with an annual prevalence estimated at between 30% and 50% (Hoy et al., 2010). Neck pain is second to low back pain in healthcare cost, work absenteeism, and loss of productive capacity and therefore poses a substantial socioeconomic burden for patients, employers, insurers and society (Hoy et al., 2014). For the majority (85%) of patients with neck pain, a pathoanatomical cause cannot be identified and as a result it is classified as non-specific neck pain (NSNP) (Walker-Bone et al., 2003; Moffett and McLean, 2006; Binder, 2007). Females tend to have higher rate of seeking physiotherapy treatment for neck pain with 35–49-year age group being at higher risk of developing neck pain (Freburger et al., 2005; Hoy et al., 2010; Hoy et al., 2014).

NSNP is frequently associated with upper limb disability (ULD) (Frank et al., 2005; Bot et al., 2005; Rasmussen et al., 2008; Feleus et al., 2008; Huisstede et al., 2009). ULD is defined here as the limitation an individual may have when performing physical activity using the upper limbs such as carrying, lifting and overhead activity (World Health Organization (WHO), 2001). It is estimated that over 80% of patients with NSNP report difficulties with daily activities that involve functional loading of the upper limbs (Osborn and Jull, 2013). The mechanisms which cause these conditions to co-exist are not clear but may relate to the mechanical attachment between the neck and upper limb via skeletal, muscular and neural structure (McLean et al., 2011). For example, mechanical loading or repetitive overhead movement of the upper limb may increase the mechanical load to the articular and ligamentous structure of the neck which may provoke neck pain or create protective neck muscle spasms (Gorski and Schwartz, 2003). Another possible mechanism is that patients with NSNP may limit the functional use of their upper limbs because of neck provocation, fear avoidance and/or poor pain self-efficacy (McLean et al., 2007; McLean et al., 2011). Consequently, a deconditioning effect may occur leading to a reduction in cardiovascular capacity and reduced strength and endurance in the neck/upper limb muscles (Smeets et al., 2006; McLean et al., 2011).

Optimal management of NSNP for many patients will include ULD rehabilitation and ongoing evaluation of the upper limb functional capacity using a suitable ULD outcome measure (McLean et al., 2011; Osborn and Jull, 2013). A recent systematic review critically reviewed the measurement properties of all available ULD outcome measures for populations with neck pain (Alreni et al., 2017). The review identified five potentially suitable measures but highlighted a lack of high-quality evidence and methodological and quality issues for all measures. Of these five measures, the 3 kg Single Arm Military Press (SAMP) test was the only performance-based ULD measure that was designed specifically for populations with neck pain (McLean et al., 2010). Preliminary investigations support the reliability and validity of the SAMP test in female non-patient populations (Alreni et al., 2017). However, it is unknown whether the SAMP test is feasible for use by clinicians for patients, which limits its clinical utility (de Vet et al., 2011). Consequently, after careful consideration of the strength and physical capacity of patients with NSNP, the aim of this study was to investigate the feasibility of the SAMP procedure using lower weights (½kg, 1 kg, 1½kg) in a clinical population.

2. Methods

2.1 Study design

This cohort study investigated the feasibility of the SAMP test procedure from both patients and clinicians' perspectives. Feasibility is described here as the difficulties and burden that patient and clinician may encounter during the administration and interpretation of the SAMP performance (Fitzpatrick et al., 1998). The study was conducted in accordance with the International Society for Quality of Life Research (ISOQOL) checklist for the development or evaluation of measurement and practical properties of outcome measures (Reeve et al., 2013). The study also adhere to the COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN) recommendation (Mokkink et al., 2010; Terwee et al., 2012, 2018) and is reported in accordance with the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) Statement: guidelines for reporting observational-cohort studies (von Elm et al., 2007).

2.2 Study sample and recruitment

Participants were recruited and female patients were included if they had acute, sub-acute or chronic NSNP with/without referred symptoms in the head or upper limbs, were at least 18 years of age, able to travel independently to the testing venue and scored at least 10 (out of 100) in the Neck Disability Index (NDI) questionnaire. Patients were excluded if they had any potentially serious conditions (e.g. systemic disease, progressive or worsening neurological disorders, inflammatory conditions or major trauma), a neck condition that requires urgent treatment or previous traumatic injury to the neck (e.g. Whiplash Associated Disorder 'WAD', Cervical Radiculopathy), upper limbs or shoulder girdle that resulted in current or prolonged disability.

A list of patients was obtained and potential participants were invited to attend an assessment and testing session. On arrival, patients completed the NDI (Vernon and Mior, 1991; Shaheen et al., 2013), after which stratification based on their NDI score (mild: 10–29, moderate: 30–49, severe: 50–100) was utilised to ensure accurate proportional representation of the sample and balance of NSNP severity between testing groups. This was followed by a subjective examination, in which standardised clinical questions were used. Patients that met the eligibility criterion were asked to sign a consent form. To prevent selection bias, a randomisation procedure with concealed allocation sequence in each of the stratified groups was then used to allocate patients into one of the final three testing groups (either ½kg, 1 kg or 1½kg), of which patients and testing examiners were blinded to the group allocation and the testing weight.

2.3 Outcome measures

2.3.1 The SAMP test

The SAMP test measures the strength and endurance of the upper limb with the expectation that difficulty in sustaining overhead activity within 30 s would discriminate between patients with varying degrees of ULD. It is conducted with the patient in a standing position with their feet positioned at shoulder width. The patient is asked to carry a dumbbell and lift it, using their dominant hand, to shoulder level (see Fig. 1A). The patient is requested to raise their hand with the dumbbell directly overhead by extending the elbow (see Fig. 1B) and repeat this process as fast and as frequently as possible for 30 s (McLean et al., 2010). The SAMP

test is quick and easy to use since it only requires one dumbbell that is available and inexpensive. Scoring the SAMP test involves counting the number of correctly performed repetitions completed within 30 s. The 3 kg SAMP test demonstrated excellent measurement properties, reliability and validity, in a population of female non-patient with and without NSNP (Alreni et al., 2017).

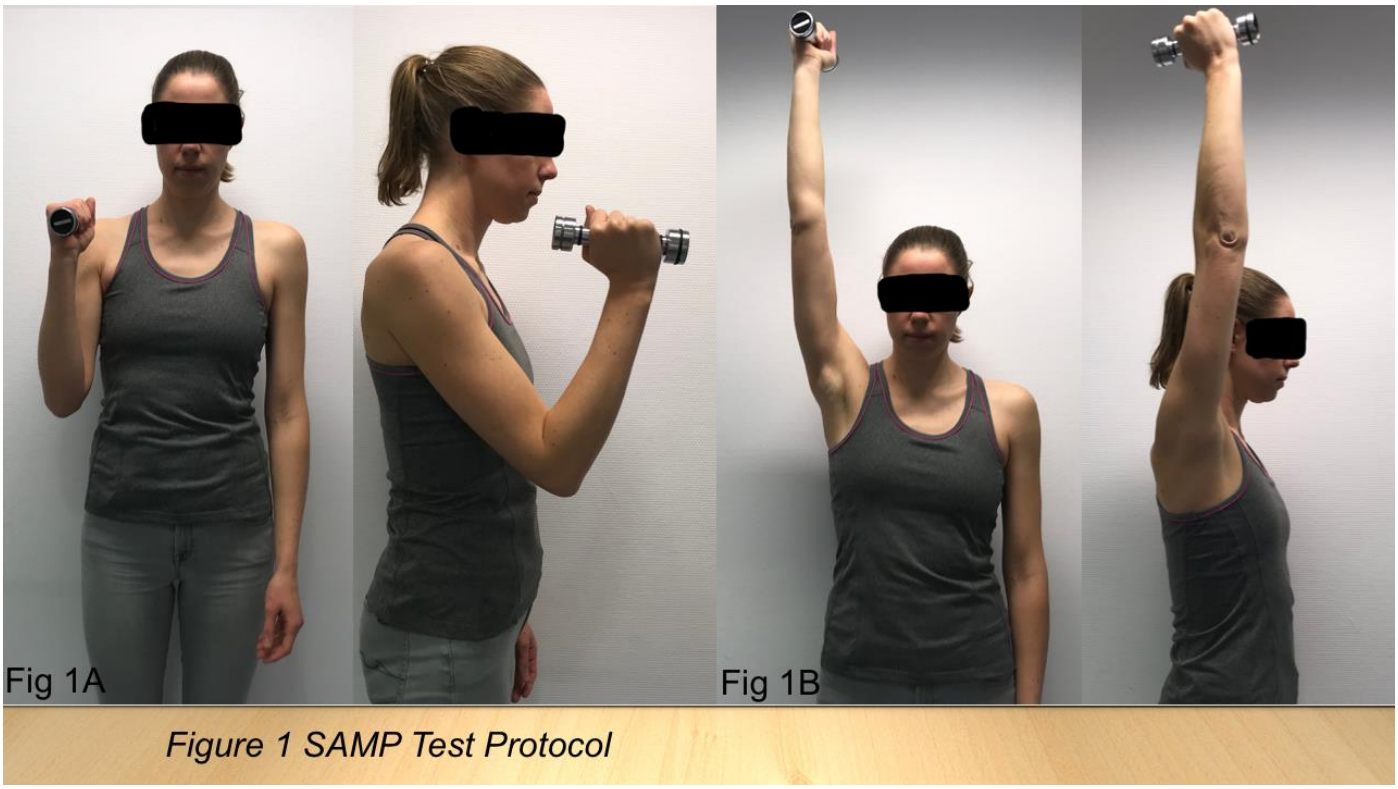

Figure 1 SAMP Test Protocol

2.3.2 The neck Disability Index (NDI) questionnaire

The NDI is a standardised patient-reported outcome measure (PROM) that was developed and extensively evaluated to measure a patient's disability due to neck pain (Vernon and Mior, 1991). It has 10 items; 7 items related to activities of daily living, 2 items related to pain, and 1 item related to concentration. Each item is scored from 0 to 5 and a total score is expressed as percentage score, with higher scores indicating greater disability. Due to its excellent measurement properties, the NDI is the most commonly used PROM for patients with neck pain in clinical and research practice (Linton, 2000; Sterling et al., 2003; Dunckley et al., 2005; Abrams et al., 2006; Bot et al., 2007; Cote et al., 2008; de Koning et al., 2008; Skeat and Perry, 2008; Nordin et al., 2009; MacDermid et al., 2009, 2013). The NDI was translated and culturally-adapted in Arabic and its reliability and validity were determined in Arabic-speaking patients with neck pain (Shaheen et al., 2013).

2.3.3 The Visual Analogue Scale (VAS)

The VAS with 0–10 response categories, where 0 indicates no symptoms and 10 indicates the worst possible symptoms, was used to measure the NSNP and ULD severity for the participants pre-testing, immediately post testing and 24 h post testing. The VAS (0–10) scale has been extensively evaluated and found to be reliable, valid and responsive in measuring symptom severity in patients with various musculoskeletal conditions including neck pain and upper limb dysfunction (Bond and Lader, 1974; Scott and Huskisson, 1977; Remington

et al., 1979; McCormack et al., 1988; Wewers and Lowe, 1990; Jaeschke et al., 1990; Bowling, 1995; van Dijk et al., 2002).

2.3.4 The Likert scale

To ensure accuracy and precision when assessing the feasibility of the SAMP test, a Likert scale with nine response categories was used for each question in a qualitative exit feedback interview with participants and examiners (Avis et al., 1994; Fitzpatrick et al., 1998). The interview questions for participants explored their experience of the weight used, the difficulties of understanding the instruction and the time and effort required. The interview questions for examiners measured the difficulty they encountered pre and post the administration of the SAMP procedure and included questions about the length and complexity of the overall testing procedure alongside the resources required (cost and time) (Feeny and Torrance, 1989; Aaronson, 1992; Lansky et al., 1992; Erickson et al., 1995). Likert scales using seven or more response categories are accurate and precise (Avis et al., 1994; Fitzpatrick et al., 1998) and reliable, valid and responsive in measuring satisfaction and feasibility (Sprangers et al., 1993; Bowling, 1995; Bolton and Wilkinson, 1998; Vickers, 1999; van Dijk et al., 2002).

2.4 Testing procedure and data collection

A total of three examiners, physicians, with at least 3-years of experience of working with musculoskeletal patients were involved in the data collection. An additional experienced fourth clinician carried out the stratification and randomisation of patients. Meanwhile, the lead author conducted the subjective examination; NSNP and ULD measuring pre and post testing; and the feedback interviews for patients and examiners.

For SAMP testing, participants were randomly assigned to one of the three proposed weight groups (½kg, 1 kg or 1½kg) to be tested once only and each group was led by one of the three examiners. This was to eliminate chances of fatigue or soreness to patients, which could lead to drop-out and also to avoid the Hawthorne effect (de Vet et al., 2011). NSNP and ULD severity were measured immediately pre-testing using a VAS scale. After this, the SAMP testing was conducted for each participant individually and started with a brief warm-up followed by description and demonstration of the SAMP procedure by the examiner. The participant was then asked to perform the SAMP test for 30 s.

Examiners completed a data collection sheet (see appendix 1). The participants NSNP and ULD severity were remeasured immediately post-testing and participants were requested to provide feedback in relation to the weight used and the SAMP procedure in a structured qualitative exit feedback interview (see Appendix 2). (Fitzpatrick et al., 1998; Reeve et al., 2013) Participants were then discharged and telephoned the following day to monitor and measure the severity of their NSNP and ULD (24 h post-testing). Following completion of all the SAMP testing on the final day, each examiner was requested to provide qualitative feedback regarding the SAMP test procedure (e.g. length, complexity, resources required) (see Appendix 3).

2.5 Data analysis

Data screening and entry were performed by the lead author and SPSS (IBM SPSS Statistical Software, version 24.0–26.0) was used for all analysis. Data were checked for normality for

each testing group using the SPSS (Skewness & Kurtosis z-values; Histograms, Normal Q-Q plots and box plots (Cramer, 1998; Doane and Seward, 2011; Razali and Wah, 2011).

Descriptive statistics (frequencies, percentages, mean, standard deviation SD) were used to present demographic information, the severity of the NSNP and ULD (pre-testing, immediately and 24 h post-testing) and the NDI scores. The SAMP test scores using the three proposed weight (½kg, 1 kg, 1½kg) and both participants and examiners' experience regarding the SAMP testing procedure were descriptively analysed to determine the most optimal weight and feasibility of the SAMP procedure. Analysis of variance (ANOVA) was used to test for differences between the three testing groups to determine the feasibility of the weight used.

3. Results

The flow of the participants through each stage is presented in Fig. 2. A list of 80 female patients was obtained from Tanta University Teaching Hospital. Following the telephone screening, 70 out 80 patients were eligible to participate. Eight patients were ineligible and 2 patients declined to participate. Following the subjective examination, all 70 patients were eligible for SAMP testing, consented in writing and were randomised to one of the SAMP testing groups.

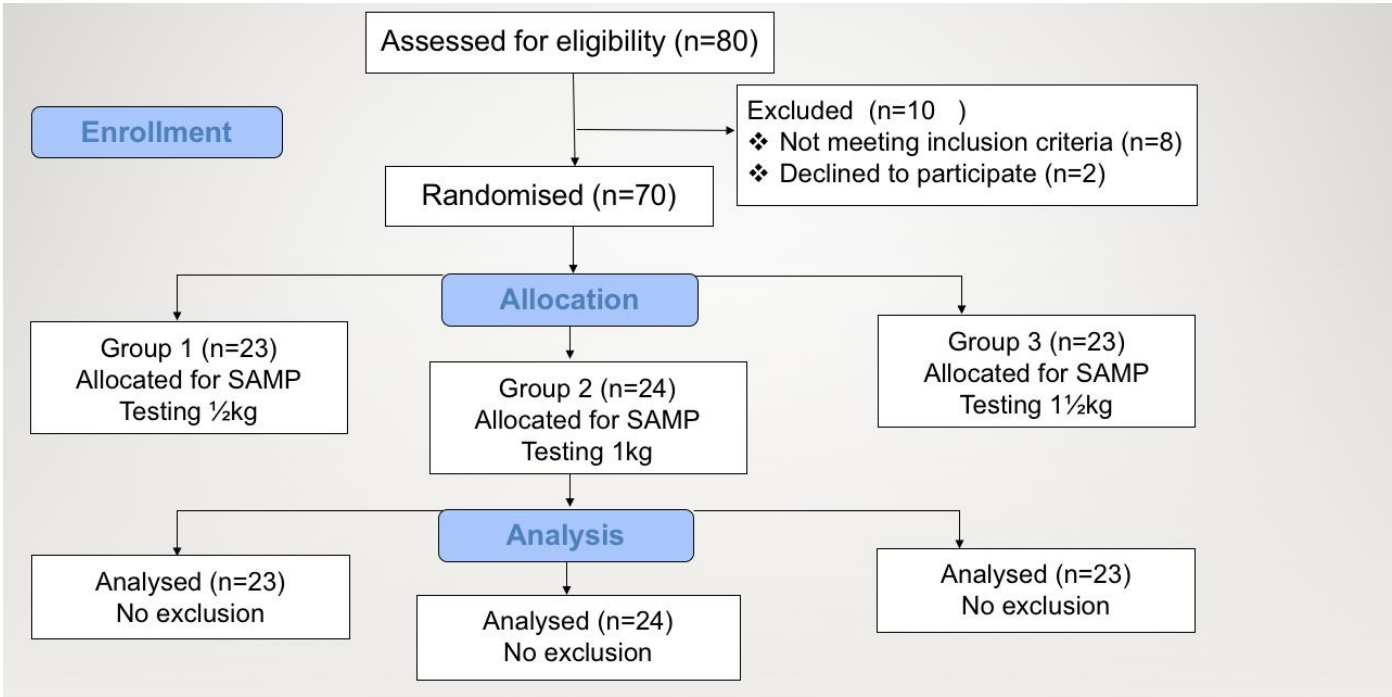

Figure 2 Flow-Chart of Participants in the Feasibility Study

3.1 Demographic characteristics

Participants characteristics, NDI scores and NSNP acuteness phases are presented in Table 1. A Shapiro-Wilk's test ($P > 0.05$) and a visual inspection of their histograms, normal Q-Q plots and box plots showed that the data for the ½kg, 1 kg and 1½kg were approximately normally distributed. The sample was balanced across the three testing groups regarding age,

weight, height and resultant BMI, and NSNP acuteness. The BMI range was high across the 3 testing groups. Group 1 has a significantly lower mean NDI score compared with group 2 and 3 ($P < 0.05$). In addition, group 1 had a slightly higher proportion of participants with acute NSNP (13%) compared with those in group 2 (0%) and 3 (0%). Group 3 had slightly higher proportion of participants with chronic NSNP (78.2) compared with group 1 (56.7%) and group 2 (62.5%). Nevertheless, the distribution of mild, moderate and severe NDI categories was equally spread across the 3 groups indicating that the stratified randomisation was broadly effective.

Table 1
Demographic Characteristics

Variables	Group 1 SAMP ½kg n=23	Group 2 SAMP 1kg n=24	Group 3 SAMP 1½kg n=23
Age (Years): Mean ± SD	39.1 ± 4.6	39.9 ± 4.4	40.9 ± 4.8
Weight (kg): Mean ± SD	90.7 ± 6.2	91.79 ± 4.5	92.7 ± 4.1
Height (cm): Mean ± SD	161.1 ± 2.3	161.25 ± 1.8	159.5 ± 2.6
BMI: Mean	35.0	35.4	36.2
NDI Score: Mean ± SD	35 ± 15.3	45 ± 17.1	43 ± 15.1
NDI categories: (frequencies)			
Mild	n=7	n=7	n=6
Moderate	n=10	n=9	n=9
Severe	n=6	n=8	n=8
NSNP phase: Frequencies (%)			
Acute	3 (13)	0 (0)	0 (0)
Sub-acute	7 (30.3)	9 (37.5)	5 (21.8)
Chronic	13 (56.7)	15 (62.5)	18 (78.2)

SD: Standard Deviation, NDI: Neck and Disability Index, NSNP: Non-Specific Neck Pain.

3.2 Feasibility of the SAMP test

Descriptive statistics for the SAMP scores, the administration and completion time and the NSNP and ULD severity immediately pre and post testing and 24 h post-testing are presented in Table 2. There were statistically significant differences ($p < 0.05$) between the mean SAMP scores of the three testing groups. The SAMP scores for those who used the ½kg hand weight (mean ± SD, 21 ± 7.1) were significantly higher compared to the scores of those used the 1 kg weight (16 ± 7.4). The SAMP scores for participants who used the 1 kg hand weight were significantly higher compared to the scores of those who used the 1½kg weight (10 ± 5.6). NSNP and ULD severity were significantly ($p < 0.01$) higher across the three groups immediately post testing; 24 h post-testing NSNP and ULD for the ½kg and 1 kg hand weight had reduced to pre-testing levels, whilst those for the 1½kg weight remained elevated. The completion time was less than 2 min across the 3 groups.

Table 2
Descriptive statistics (SAMP scores, NSNP and ULD measurements).

Variable	½ kg n=23	1kg n=24	1½kg n=23	*F*	*P* value
SAMP Scores: Mean ± SD	21 ± 7.1	16 ± 7.4	10 ± 5.6	15.04	0.01
SAMP Procedure (S): Mean ± SD					
Administration Time	56 ± 4.1	55 ± 4.3	47 ± 8.5	25.21	0.01
Completion Time	116 ± 4.1	115 ± 4.3	107 ± 8.5	25.21	0.01
NSNP (VAS): Mean ± SD					
Immediately Pre SAMP	4 ± 0.7	5 ± 0.7	4 ± 0.7	1.84	0.17
Immediately Post SAMP	5 ± 1.0	6 ± 0.8	6 ± 0.8	8.96	0.01
24 Hours Post SAMP	3 ± 0.9	4 ± 0.8	5 ± 0.9	21.78	0.01
ULD (VAS): Mean ± SD					
Immediately Pre SAMP	2 ± 0.9	2 ± 1.1	2 ± 0.7	1.22	0.30
Immediately Post SAMP	4 ± 0.9	5 ± 1.1	6 ± 0.7	34.96	0.01
24 Hours Post SAMP	2 ± 1.0	2 ± 1.1	5 ± 1.0	70.63	0.01

SD: Standard Deviation, SAMP: Single Arm Military Press, S: Second, ULD: Upper Limb Disability, NSNP: Non-Specific Neck Pain, VAS: Visual Analogue Scale, F: One-Way ANOVA.

Table 3 presents the patients and examiners' experiences regarding the weight used and the SAMP procedure. The majority of participants in group 1 reported that the ½kg hand weight was extremely light (median ± SD, 1.0 ± 0.8), whereas those in group 3 reported that the 1½kg weight was very heavy (7.0 ± 1.6). The majority of participants in group 2 reported that the 1 kg was neither light nor heavy (4.0 ± 1.2). Participants across the groups stated that the SAMP procedure was extremely easy regarding instruction and performance (1.0 ± 0.1, 0.3). In addition, examiners involved in the testing process reported that the SAMP procedure was extremely easy regarding administration and completion (1.0 ± 0.1, 0.2, 0.3). They also confirmed that the SAMP test procedure was highly appropriate regarding the resources required (1.0 ± 0.2, 0.3, 0.4). Therefore, the 1 kg SAMP test procedure was deemed to be feasible by patients and clinicians for use in a population of female patients with NSNP.

Table 3
Patients & Examiners' experiences about the weight used and the SAMP procedure.

Patients feedback	½ kg n=23	1kg n=24	1½kg n=23
Weight used: Median ± SD Likert scale 1-9 (1=extremely light / 9=extremely heavy)	1.0 ± 0.8	4.0 ± 1.2	7.0 ± 1.6
Willingness and ability: Median ± SD Likert scale 1-9 (1=extremely easy / 9=extremely difficult)	1.0 ± 1.7	4.0 ± 1.7	7.0 ±1.2
Instruction and performance: Median ± SD Likert scale 1-9 (1=extremely easy / 9=extremely difficult)	1.0 ± 0.3	1.0 ± 0.1	1.0 ± 0.1
Time and effort: Median ± SD Likert scale 1-9 (1=highly suitable / 9=completely unsuitable)	1.0 ± 0.9	2.0 ± 0.9	3.0 ± 0.9
Examiners feedback	**Examiner A**	**Examiner B**	**Examiner C**
Explanation, demonstration and instruction to patients: Median ± SD Likert scale 1-9 (1=extremely easy / 9=extremely difficult)	1.0 ± 0.1	1.0 ± 0.2	1.0 ± 0.2
Administration and completion: Median ± SD Likert scale 1-9 (1=extremely easy / 9=extremely difficult)	1.0 ± 0.3	1.0 ± 0.1	1.0 ± 0.2
Resources required (e.g. time and cost): Median ± SD Likert scale 1-9 (1=highly appropriate / 9=completely inappropriate)	1.0 ± 0.2	1.0 ± 0.3	1.0 ± 0.4

SD: Standard Deviation.

4. Discussion

4.1 Summary and discussion of the main findings

The study aimed to determine the feasibility of the SAMP procedure using lower weights (½kg, 1 kg, 1½kg) in female patients with NSNP. Nearly all patients and clinicians involved in the study confirmed that regardless of the weight used, the SAMP test procedure was simple, quick, inexpensive and extremely easy to administer and score. The reasons for the high level of feasibility may be because the SAMP test is convenient, since it can be efficiently administered by clinicians with varying experience in any setting using the minimum of equipment:1 dumbbell, which is readily available and inexpensive. Furthermore, it is time effective as it takes less than 2 min for administration and completion.

The 1 kg SAMP test procedure was deemed feasible by patients and clinicians for use in a population of female patients with NSNP. A majority of participants in group 1 reported that the ½kg hand weight was too light. Using a hand weight that is too light in the practical application of a physical performance measure risks missing out on identifying patients with subtle/mild pain and/or disability due to ceiling effect (Fitzpatrick et al., 1998; Reeve et al., 2013). This suggests that the ½kg hand weight is unsuitable for use in the practical application of the SAMP test procedure for female patients with NSNP. Conversely, the 1½kg hand weight was considered too heavy by the majority of participants, with a number of participants being unable to lift the weight or experiencing an increase in NSNP and ULD immediately post-testing which did not resolve 24 h post-testing. Using a heavy hand weight in the practical application of a physical performance measure may distress patients and risk aggravating their pain for a longer period post-testing, which could lead to patient fear and avoidance of physiotherapy and exercise and non-adherence to rehabilitation (Fitzpatrick et al., 1998; Reeve et al., 2013; Ahuga, 2015). Furthermore, the findings regarding the 1 kg hand weight are consistent with those of other studies which used a 1 kg hand weight when measuring the functional capacity of the upper limb for Canadian patients with neck pain and/or shoulder pathology (MacDermid et al., 2007; Kumta et al., 2012; Constand and MacDermid, 2013).

The findings of this study suggest that the SAMP test with a 1 kg hand weight could potentially be a clinically useful measure of ULD for patients with NSNP. The 3 kg SAMP test has been shown to have excellent reliability and validity in female non-patient populations (McLean et al., 2010; Alreni et al., 2017), however this has not yet been reported for the 1 kg SAMP test. Therefore, further investigation of the 1 kg SAMP test is required to establish the reliability and validity of the measure, prior to use in a clinical setting, and its responsiveness to change over time in a patient population with NSNP and other types of neck pain.

4.2 Strengths of the study

This study was conducted, analysed and interpreted in accordance with both ISOQOL checklist (Reeve et al., 2013) and the COSMIN recommendations for developing and evaluating health-related outcome measures (Mokkink et al., 2010; Terwee et al., 2012, 2018). The study was reported in accordance with the STROBE statement (von Elm et al., 2007). Patients and clinicians were involved which ensured relevance and feasibility of the measure for these key stakeholders (Fitzpatrick et al., 1998; de Vet et al., 2011; Reeve et al., 2013; Terwee et al., 2018). A larger than recommended sample size coupled with broad

inclusion and exclusion criteria ensured that the participants were representative of female patients with NSNP (Fitzpatrick et al., 1998; de Vet et al., 2011). Broadly effective stratified randomisation and the use of valid measures ensured that robust processes were used to investigate the feasibility of the SAMP test (Donovan et al., 1993; Fitzpatrick et al., 1998; Hasson and Arnetz, 2005; de Vet et al., 2011; Brokelman et al., 2012; Reeve et al., 2013).

4.3 Limitations of the study

This study was conducted on a population of female patients with a higher BMI range (Badran and Laher, 2011) and therefore the findings may potentially be different in females with lower BMI values (Ylinen et al., 2004; Svedmark et al., 2016). Hence, further investigation may be required to examine the feasibility of the 1 kg SAMP test in other female with lower BMI range. Although the distribution into severity categories was similar across all 3 groups, the lower mean NDI in the ½kg group may have influenced this group's estimate that the weight was too light and easy, since they may have had less upper limb problems than the participants in the 1 kg and 1½kg groups. Each of the 3 testing groups was led by three different examiners. In spite of the training to ensure consistency of approach between the examiners, it is still possible that examiner bias may have influenced the findings in each group. Because this study was conducted on female patients, it is not possible to generalise these findings to male patients with NSNP. It is well known that due to larger body size and muscle cross-sectional area, males are generally stronger than females (Cheng et al., 2003; Hunter, 2009). Consequently, the 1 kg hand weight may be too light and further investigation is required to establish the optimum feasible weight of the SAMP test for male patients with NSNP. This study focused on participants with NSNP aged between 39 and 41, which may limit the generalisability of this test to older women whose upper limb strength may be diminished (McLean et al., 2010) and patients with other forms of neck pain (e.g. cervical radiculopathy or Whiplash Associated Disorder "WAD") who may have greater severity of neck pain and ULD (Jull et al., 2008). Finally, the researcher who designed the study was also the clinician who collected the patients and examiners qualitative input post testing, which may present assessor bias.

5. Conclusion

This study established that the 1 kg SAMP test procedure is feasible for use in female patients with NSNP. The SAMP test is convenient, efficient and inexpensive and therefore has the potential to be useful in clinical practice and research although further testing of reliability and validity is required prior to use in a clinical setting. The measurement and practical properties of the SAMP test should also be tested in male patients, populations with other neck and upper limb dysfunctions and female populations in other national settings.

Ethical approval
Ethical approval for this study was obtained from Sheffield Hallam University Research Committee (SHUREC) and Tanta University Teaching Hospital.

Funding
None declared.

Funding acknowledgement
This research was unfunded.

Declaration of competing interest
None declared.

Acknowledgments
Dr Karen Kilner from Sheffield Hallam University for her valuable and constructive suggestions during the planning and development of this research work. Prof Dr Ali Eldeeb and all the staff in the Rheumatology and Physical Therapy Department at Tanta University Teaching Hospital for their assistance throughout the recruitment and examination of patients.

Appendix A. Supplementary data

Supplementary data to this article can be found online at
https://doi. org/10.1016/j.msksp.2020.102254.

References

Aaronson, N.K., 1992. Assessing the quality of life of patients in cancer clinical trials: common problems and common-sense solutions. Eur. J. Canc. 28 (8–9), 1304–1307. Abrams, D., Davidson, M., Harrick, J., Harcourt, P., Zylinski, M., Clancy, J., 2006.

Monitoring the change: current trends in outcome measure usage in physiotherapy. Man. Ther. 11 (1), 46–53.

Ahuga, D., 2015. Determinants of Rehabilitation Adherence in Outpatient Musculoskeletal Physiotherapy: A Mixed Methods PhD Project. Sheffield Hallam University, Sheffield, p. 426.

Alreni, A.S.E., Harrop, D., Lowe, A., Potia, T., Kilner, K., McLean, S.M., 2017. Measures of upper limb function for people with neck pain. A systematic review of measurement and practical properties. Musculoskel. Sci. Pract. 29, 155–163.

Avis, N.E., Smith, K.W., 1994. Conceptual and methodological issues in selecting and developing quality of life measures. In: Fitzpatrick, R. (Ed.), Advances in Medical Sociology. JAI Press Inc, London, pp. 255–280.

Badran, M., Laher, I., 2011. Obesity in Arabic-speaking countries. J. Obes. https://doi.org/10.1155/2011/686430.

Binder, A., 2007. The diagnosis and treatment of nonspecific neck pain and whiplash. Eur. Medicophys. 43 (1), 79–89.

Bolton, J.E., Wilkinson, R.C., 1998. Responsiveness of pain scales: a comparison of three pain intensity measures in chiropractic patients. J. Manip. Physiol. Ther. 21 (1), 1–7.

Bond, A., Lader, M., 1974. The use of analogue scales in rating subjective feelings. Br. J. Med. Psychol. 47, 211–218.

Bot, S.D., van der Waal, J.M., Terwee, C.B., van der Windt, D.A., Scholten, R.J., Bouter, L.M., et al., 2005. Predictors of outcome in neck and shoulder symptoms: a cohort study in general practice. Spine 30 (16), 459–470.

Bot, S.D.M., Terwee, C.B., van der Windt, D.A., van der Beek, A.J., Bouter, L.M., Dekker, J., 2007. Work-related physical and psychosocial risk factors for sick leave in patients with neck or upper extremity complaints. Int. Arch. Occup. Environ. Health 80 (8), 733–741.

Bowling, A., 1995. Measuring Disease. Open University Press, Buckingham, p. 374.

Brokelman, R.B.G., Haverkamp, D., van Loon, C., Hol, A., van Kapen, A., Veth, R., 2012. The validation of the visual analogue scale for patient satisfaction after total hip arthroplasty. Eur. Orthop. Traumatol. 3, 101–105.

Cheng, A., Ditor, D.S., Hicks, A.L., 2003. A comparison of adductor pollicis fatigue in older men and women. Can. J. Physiol. Pharmacol. 81 (9), 873–879.

Constand, M.K., MacDermid, J.C., 2013. Effect of neck pain on reaching overhead and reading: a case-control study of long and short neck flexion. BMC Sports Sci. Med. Rehabil. 5, 21.

Cote, P., van der Velde, G., Cassidy, J.D., Carroll, L.J., Hog-Johnson, S., Holm, L.W., et al., 2008. The burden and determinants of neck pain in workers: results of the Bone and Joint Decade 2000–2010 Task Force on Neck Pain and its associated disorders. Spine 33, 60–74.

Cramer, D., 1998. Fundamental Statistics for Social Research. Routledge, London.

de Koning, C.H., van den Heuvel, S.P., Staal, J.B., Smits-Engelsman, B.C., Hendriks, E.J., 2008. Clinimetric evaluation of active range of motion measures in patients with non-specific neck pain: a systematic review. Eur. Spine 17 (7), 905–921.

de Vet, H.C.W., Terwee, C.B., Mokkink, L.B., Knol, D.L., 2011. Measurement in Medicine: A Practical Guide. Cambridge University Press, Cambridge.

Doane, D.P., Seward, L.E., 2011. Measuring skewness. J. Stat. Educ. 19 (2), 1–18. Donovan, J.L., Frankel, S.J., Eyles, J.D., 1993. Assessing the need for health status measures. J. Epidemiol. Community Health 47 (2), 158–162.

Dunckley, M., Aspinal, F., Addington-Hall, J.M., Hughes, R., Higginson, I.J., 2005. A research study to identify facilitators and barriers to outcome measure implementation. Int. J. Palliat. Nurs. 11 (5), 218–225.

Erickson, P., Taeuber, R.C., Scott, J., 1995. Operational aspects of quality-of-life assessment: choosing the right instrument: review article. Pharmacoeconomics 7, 39–48.

Feeny, D.H., Torrance, G.W., 1989. Incorporating utility-based quality-of-life assessment measures in clinical trials. Two examples. Med. Care 27 (3), 190–204.

Feleus, A., Bierma-Zeinstra, S.M., Miedema, H.S., Bemsen, R.M., Verhaar, J.A., Koes, B.W., 2008. Incidence of nontraumatic complaints of arm, neck and shoulder in general practice. Man. Ther. 13 (5), 426–433.

Fitzpatrick, R., Davey, C., Buxton, M.J., Jones, D.R., 1998. Evaluating patient-based outcome measures for use in clinical trials. Health Technol. Assess. 2 (14), 1–74.

Frank, A.O., De Souza, L.H., Frank, C.A., 2005. Neck pain and disability: a cross-sectional survey of the demographic and clinical characteristics of neck pain seen in a rheumatology clinic. Int. J. Clin. Pract. 59 (2), 173–182.

Freburger, J.K., Carey, T.S., Holmes, G.M., 2005. Management of back and neck pain: who seeks care from physical therapists? Phys. Ther. 85, 872–886.

Gorski, J.M., Schwartz, L.H., 2003. Shoulder impingement presenting as neck pain. J. Bone Joint Surg. 85 (4), 635–638.

Hasson, D., Arnetz, B.B., 2005. Validation and findings comparing VAS vs. Likert scales for psychosocial measurements. Int. Electron. J. Health Educ. 8, 178–192.

Hogg-Johnson, S., van der Velde, G., Carroll, L.J., Holm, L.W., Cassidy, J.D., Guzman, J., et al., 2008. The burden and determinants of neck pain in the general population: results of the bone and joint decade 2000–2010 task force on neck pain and its associated disorders. Spine 33 (4), 39–51.

Hoy, D.G., Protani, M., De, R., Buchbinder, R., 2010. The epidemiology of neck pain. Best Pract. Res. Clin. Rheumatol. 24 (6), 783–792.

Hoy, D.G., March, L., Woolf, A., Blyth, F., Brooks, P., Smith, E., et al., 2014. The global burden of neck pain: estimates from the global burden of disease 2010 study. Ann. Rheum. Dis. 73 (7), 1309–1315.

Huisstede, B.M., Feleus, A., Bierma-Zeinstra, S.M., Verhaar, J.A., Koes, B.W., 2009. Is the disability of arm, shoulder, and hand questionnaire (DASH) also valid and responsive in patients with neck complaints. Spine 34 (4), 130–138.

Hunter, S.K., 2009. Sex differences and mechanisms of task-specific muscle fatigue. Exerc. Sport Sci. Rev. 37 (3), 113–122.

Jaeschke, R., Singer, J., Guyatt, G.H., 1990. A comparison of seven-point and visual analogue scales. Data from a randomized trial. Contr. Clin. Trials 11 (1), 43–51.

Jull, G., Sterling, M., Falla, D., Treleaven, J., O'leary, S., 2008. Whiplash, Headache, and Neck Pain: Research-Based Direction for Physical Therapies. Churchill Livingstone.

Kumta, P., MacDermid, J.C., Mehta, S.P., Stratford, P.W., 2012. The FIT-HaNSA demonstrates reliability and convergent validity of functional performance in patients with shoulder disorders. J. Orthop. Sports Phys. Ther. 42 (5), 455–464.

Lansky, D., Butler, J.B.V., Waller, F.T., 1992. Using health status measures in the hospital setting: from acute care to 'outcomes management'. Med. Care 30 (5), 57–73. Linton, S.J., 2000. A review of psychological risk factors in back and neck pain. Spine 25(9), 1148–1156.

MacDermid, J.C., Ghobrial, M., Quirion, K.B., St-Amour, M., Tsui, T., Humphreys, D., et al., 2007. Validation of a new test that assesses functional performance of the upper extremity and neck (FIT-HaNSA) in patients with shoulder pathology. BMC Muscoskel. Disord. 8, 42.

MacDermid, J.C., Walton, D.M., Avery, S., Blanchard, A., Etruw, E., McAlpine, C., et al., 2009. Measurement properties of the neck disability index: a systematic review. J. Orthop. Sports Phys. Ther. 39 (5), 400–417.

MacDermid, J.C., Walton, D.M., Cote, P., Santaguida, P.L., Gross, A., Carlesso, L., et al., 2013. Use of outcome measures in managing neck pain: an international multidisciplinary survey. Open Orthop. J. 7, 506–520.

McCormack, H.M., Horne, D.J., Sheather, S., 1988. Clinical applications of visual analogue scales: a critical review. Psychol. Med. 18 (4), 1007–1019.

McLean, S.M., May, S., Moffett, J.K., Sharp, D.M., Gardiner, E., 2007. Prognostic factors for progressive non-specific neck pain. Phys. Ther. Rev. 12 (3), 207–220.

McLean, S.M., Taylor, J., Balassoubramien, T., Kulkarni, M., Patekar, P., Darne, R., et al., 2010. Measuring upper limb disability in non-specific neck pain: a clinical performance measure. Int. J. Physiother Rehabil. 1, 44–52.

McLean, S.M., Moffett, J.K., Sharp, D.M., Gardiner, E., 2011. An investigation to determine the association between neck pain and upper limb disability for patients with non-specific neck pain: a secondary analysis. Man. Ther. 16 (5), 434–439.

Moffett, J., McLean, S., 2006. The role of physiotherapy in the management of nonspecific back pain and neck pain. Rheumatology 45 (4), 371–378.

Mokkink, L.B., Terwee, C.B., Patrick, D.L., Alonso, J., Stratford, P.W., Knol, D.L., et al., 2010. The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. Qual. Life Res. 19 (4), 539–549.

Nordin, M., Carragee, E.J., Hogg-Johnson, S., Weiner, S.S., Hurwitz, E.L., Peloso, P.M., et al., 2009. Assessment of neck pain and its associated disorders: results of the bone and joint decade 2000-2010 task force on neck pain and its associated disorders. J. Manip. Physiol. Ther. 32 (2), 111–140.

Osborn, W., Jull, G., 2013. Patients with non-specific neck disorders commonly report upper limb disability. Man. Ther. 18 (6), 492–497.

Rasmussen, C., Leboeuf-Yde, C., Hestbaek, L., Manniche, C., 2008. Poor outcome in patients with spine-related leg or arm pain who are involved in compensation claims: a prospective study of patients in the secondary care sector. Scand. J. Rheumatol. 37 (6), 462–468.

Razali, N.M., Wah, Y.B., 2011. Power comparisons of shapiro-wilk, kolmogorv-smirnov, lilliefors and anderson-darling tests. J. Stat. Model. Anal. 2 (1), 21–33.

Reeve, B.B., Wyrwich, K.W., Wu, A.W., Velikova, G., Terwee, C.B., Snyder, C.F., et al., 2013. ISOQOL recommends minimum standards for patient-reported outcome measures used in patient-centred outcomes and comparative effectiveness research. Qual. Life Res. 22 (8), 1889–1905.

Remington, M., Tyrer, P.J., Newson-Smith, J., Cicchetti, D.V., 1979. Comparative reliability of categorical and analogue rating scales in the assessment of psychiatric symptomatology. Psychol. Med. 9 (4), 765–770.

Scott, P.J., Huskisson, E.C., 1977. Measurement of functional capacity with visual analogue scales. Rheumatol. Rehabil. 16 (4), 257–259.

Shaheen, A.A., Omar, M.T., Vernon, H., 2013. Cross-cultural adaptation, reliability, and validity of the Arabic version of neck disability index in patients with neck pain. Spin 38 (10), 609–615.

Skeat, J., Perry, A., 2008. Exploring the implementation and use of outcome measurement in practice: a qualitative study. Int. J. Lang. Commun. Disord 43 (2), 110–125.

Smeets, R.J., Wade, D., Hidding, A., Van Leeuwen, P.J., Vlaeyen, J.W., Knottnerus, J.A., 2006. The association of physical deconditioning and chronic low back pain: a hypothesis-oriented systematic review. Disabil. Rehabil. 28 (11), 673–693.

Sprangers, M.A., Cull, A., Bjordal, K., Groenvold, M., Aaronson, N.K., 1993. The European organization for research and treatment of cancer. Approach to quality of life assessment: guidelines for developing questionnaire modules. EORTC study group on quality of life. Qual. Life Res. 2, 287–295.

Sterling, M., Kenardy, J., Jull, G., Vicenzino, B., 2003. The development of psychological changes following whiplash injury. Pain 106 (3), 481–489.

Svedmark, A., Djupsjobacka, M., Hage, C., Jull, G., Bjorklund, M., 2016. Is tailored treatment superior to non-tailored treatment for pain and disability in women with non-specific neck pain? A randomized controlled trial. BMC Muscoskel. Disord. 17(1), 408.

Terwee, C.B., Mokkink, L.B., Knol, D.L., Ostelo, R., Bouter, L.X., de Vet, H.C.W., 2012. Rating the methodological quality in systematic reviews of studies on measurement properties: a scoring system for the COSMIN checklist. Qual. Life Res. 21 (4), 651–657.

Terwee, C.B., Prinsen, C.A.C., Chiaeotti, A., Westerman, M.J., Patrick, D.L., Alonso, J., et al., 2018. COSMIN methodology for evaluating the content validity of patient-reported outcome measures: a Delphi study. Qual. Life Res. 27 (5), 1159–1170.

van Dijk, M., Koot, H.M., Saad, H.H., Tibboel, D., Passchier, J., 2002. Observational visual analog scale in paediatric pain assessment: useful tool or good riddance? Clin. J. Pain 18 (5), 310–316.

Vernon, H., Mior, S., 1991. The Neck Disability Index: a study of reliability and validity. J. Manip. Physiol. Ther. 14 (7), 409–415.

Vickers, A.J., 1999. Comparison of an ordinal and a continuous outcome measure of muscle soreness. Int. J. Technol. Assess. Health Care 15 (4), 709–716.

von Elm, E., Altman, D.G., Egger, M., Pocock, S.J., Gotzsche, P.C., Vandenbroucke, J.P., 2007. The strengthening the reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies. Ann. Intern. Med. 147 (8), 573–577.

Walker-Bone, K.E., Palmer, K.T., Reading, I., Cooper, C., 2003. Soft-tissue rheumatic disorders of the neck and upper limb: prevalence and risk factors. Semin. Arthritis Rheum. 33 (3), 185–203.

Wewers, M.E., Lowe, N.K., 1990. A critical review of visual analogue scales in the measurement of clinical phenomena. Res. Nurs. Health 13 (4), 227–236.

World Health Organization (WHO), 2001. In: International Classification of Functioning, Disability and Health: ICF. http://www.who.int/.

Ylinen, J., Takala, E.P., Kautiainen, H., Nykanen, M., Hakkinen, A., Pohjolainen, T., et al., 2004. Association of neck pain, disability and neck pain during maximal effort with neck muscle strength and range of movement in women with chronic nonspecific neck pain. Eur. J. Pain 8 (5), 473–478.