Received: 27 March 2023 | Revised: 29 July 2023 | Accepted: 31 July 2023

DOI: 10.1111/aogs.14665

AOGS Acta Obstetricia et Gynecologica Scandinavica

ORIGINAL RESEARCH ARTICLE

Amsterdam complex pelvic pain symptom scale with subscales: Based on a Norwegian translation, psychometric assessment and modification of the Amsterdam hyperactive pelvic floor scale

Susan Saga[1] | Signe Nilssen Stafne[1,2] | Vilde Dons[3] | Ingvild Bratlie[3] | Siri Spetalen[3] | Cecilie Therese Hagemann[4,5]

[1]Department of Public Health and Nursing, Faculty of Medicine and Health Sciences NTNU, Norwegian University of Science and Technology, Trondheim, Norway

[2]Clinic of Rehabilitation, St. Olav's Hospital, Trondheim University Hospital, Trondheim, Norway

[3]Faculty of Medicine and Health Sciences NTNU, Norwegian University of Science and Technology, Trondheim, Norway

[4]Department of Clinical and Molecular Medicine, Faculty of Medicine and Health Sciences NTNU, Norwegian University of Science and Technology, Trondheim, Norway

[5]Department of Obstetrics and Gynecology, St. Olav's Hospital, Trondheim University Hospital, Trondheim, Norway

Correspondence
Susan Saga, Department of Public Health and Nursing, Faculty of Medicine and Health Sciences NTNU, Norwegian University of Science and Technology, Fakultet for medisin og helsevitenskap, Institutt for samfunnsmedisin og sykepleie, Postboks 8905, N-7491, Trondheim, Norway.
Email: susan.saga@ntnu.no

Abstract

Introduction: Women with an abnormally high pelvic floor muscle tone may have a clinical presentation that is complex, involving urinary, anorectal and/or sexual dysfunction, genital/pelvic pain and psychological distress. The Amsterdam Hyperactive Pelvic Floor Scale (AHPFS) is a Dutch 30-item condition-specific self-report questionnaire developed to measure these complex pelvic pain symptoms. The aim of this study was to translate the Dutch version into Norwegian, to assess the psychometric properties, and to present a valid factor structure.

Material and methods: Translation, back-translation and a review of the back-translated version were performed. Thereafter, a pilot test including feedback from six clinical experts and cognitive interviews with 11 patients from the target group was conducted. Next, a field test was performed among women who were (1) patients at the gynecological outpatient clinic/pelvic floor physiotherapist at St. Olav's Hospital, (2) members of the Vulvodynia or the Endometriosis Patient Associations or (3) female students and employees from the Faculty of Medicine and Health Science, the Norwegian University of Science and Technology, in a web-based survey. To ensure a sample with symptomatic women, only women who scored ≥11 according to the Dutch prespecified factor structure were included in the statistical analyses ($n = 232$).

Results: Content/face validity demonstrated that the questionnaire was perceived as relevant, comprehensive and understandable. Some adjustments in the instructions of the questionnaire and the response categories were made, which lead to the Norwegian translation ACPPS-30. Assessment of the questionnaire's dimensionality revealed a five-factor structure similar to the original Dutch Amsterdam Hyperactive Pelvic Floor Scale (AHPFS) but without the Urinary tract infection factor and seven

Abbreviations: ACPPS, Amsterdam Complex Pelvic Pain Symptom Scale; AHPFS, Amsterdam Hyperactive Pelvic Floor Scale; CFA, confirmatory factor analysis; df, degrees of freedom; FSFI, Female Sexual Function Index; ICC, intraclass correlation coefficient; MI, modification indices; NTNU, Norwegian University of Science and Technology; PCA, principal component analysis; ρc, composite reliability.

other items. The translated and modified ACPPS-16 total score and subscales correlated as expected with scales measuring similar conditions. Test–retest reliability demonstrated good stability for scales (intraclass correlation coefficient 0.85–0.93) and single items (weighted kappa values from 0.34 to 0.90).

Conclusions: A modified Norwegian version ACPPS-30 was presented, in addition to a shorter version with only 16 of the translated items distributed among five factors similar to the original Dutch version (ACPPS-16). Both versions proved to be valid, stable and reliable tools to investigate complex pelvic pain symptoms possibly due to an abnormally high-toned pelvic floor muscle.

KEYWORDS
AHPFS, hyperactive pelvic floor, overactive pelvic floor, pelvic floor dysfunctions, pelvic pain, PROM, psychometric evaluation, questionnaire, validation

1 | INTRODUCTION

For women with complex pelvic pain – often inadequately termed a hyperactive pelvic floor – this is a disabling and overlooked condition. The condition is associated with urological, gynecological and gastrointestinal symptoms, sexual problems and chronic pelvic pain and has a huge impact on quality of life.[1] A hyperactive pelvic floor has been defined as "a condition in which the pelvic floor muscles do not relax or may even contract when relaxation is functionally needed, for example during micturition or defecation. This condition is based on symptoms such as voiding problems, obstructed defecation or dyspareunia and on signs like the absence of voluntary pelvic floor muscle relaxation".[2]

An intact and well-functioning pelvic floor is characterized by the connective tissue of the ligaments and fascias and the pelvic floor muscles acting together to counteract the impact of any increase in intra-abdominal pressure and ground reaction forces, thus keeping the pelvic organs in place with little downward movements and little or no opening of the levator hiatus area or the urethra.[3] This is important in providing support for internal organs, as well as preventing urinary and anal incontinence, a satisfying sexual life, and possibilities for movement and exercise. Muscular tone exists on a continuum and any change in muscular tone depending on intra-abdominal pressure should be automatic in women with a well-functioning pelvic floor. However, tone is a dynamic physiological state modulated by many inputs, such as the spinal cord, cortex, brainstem relays, stretch reflexes and cutaneous receptors, visceromotor reflex pathways, emotions and pain (anticipation or experience of pain).[4] The muscle tone in the pelvic floor can be abnormally low or abnormally high, resulting in various symptoms of pelvic floor disorders such as urinary or anal incontinence, pelvic organ prolapse, sensory and emptying abnormalities of the lower urinary tract, defecatory dysfunction, sexual dysfunction and chronic pain syndromes.[5]

Multiple possible etiologies exist for the onset of pelvic floor disorders associated with an abnormally high-toned pelvic floor, and it is often difficult to identify the exact cause.[1] Possible etiologic factors are chronic pelvic pain, psychological distress (eg anxiety, fear of pain), psychosocial/psychosexual disturbances (eg adverse relationships, sexual trauma or abuse), abnormal behaviors/patterns of pelvic floor muscle use (eg prolonged holding to delay voiding or defecation), direct trauma or pathology/disorder causing tissue changes within the pelvic region, postural abnormalities in the region of the spine, pelvis, and/or lower extremities (eg faulty sitting and standing postures, and prolonged lack of motion and/or repetitive activities, structural/skeletal asymmetries).[1] A recent systematic review and meta-analysis found that women with persistent pelvic pain conditions had higher pelvic floor muscle tone than women without pelvic pain.[6] Persistent pelvic pain, also known as chronic pelvic pain, is defined as persistent pain lasting longer than 6 months or recurrent episodes of abdominal/pelvic pain, hypersensitivity or discomfort often associated with elimination changes, and sexual dysfunction often in the absence of organic etiology.[7] The Working Group of the International Continence Society Standardization Steering Committee on Chronic Pelvic Pain has stressed that chronic pelvic pain is a complex syndrome consisting of concurrent symptoms and signs.[8] In this study, we will therefore use the term complex pelvic pain. A consideration of all potential initiating and/or contributing factors is therefore essential in the assessment and treatment of women presenting with complex pelvic pain symptoms.

The prevalence of women reporting complex pelvic pain symptoms possibly due to an abnormally high pelvic floor tone is uncertain, as it is a heterogeneous state and the underlying causes

Key message

A Norwegian modified 30-item version and a shorter 16-item version with five factors was presented. Though shorter, the Norwegian version has proven to be a valid, stable, and reliable tool to investigate complex pelvic pain symptoms.

are complex and individual.[9] Mentioning pelvic floor disorders is still largely taboo, and many women suffer silently due to a lack of information about prevention and treatment. Health professionals also find this topic difficult to address in clinical settings. Therefore, questions concerning pelvic floor dysfunction are often left out in consultations with women, which means that these ailments are difficult to discover.[10] There are validated scales to measure symptom-specific pelvic floor dysfunctions, eg urinary incontinence (ICIQ-UI short form);[11] pelvic floor distress, including micturition, defecation, and pelvic organ prolapse (PFDI-20);[12] anal incontinence (Wexner score);[13] and constipation (ODSS).[14] Further, validated scales exist that address sexual dysfunction in different settings, eg the Female Sexual Function Index (FSFI)[15] and the PISQ-IR.[16] No pelvic floor symptom-specific validated questionnaires include questions on widespread pain, fatigue, anxiety or depression.

A research group from the Netherlands led by Laan and van Lunsen developed the Amsterdam Hyperactive Pelvic Floor Scale (AHPFS), a 30-item self-report questionnaire intended to cover the broad range of symptoms related to abnormally high pelvic floor tone.[17] Scores are given on a 5-point Likert scale ranging from "never" (score 1) to "very often" (score 5). Among the 30 items, 25 are divided into six subscales: (1) provoked vulvodynia (six items); (2) irritable bowel syndrome (four items); (3) lower urinary tract symptoms (four items); (4) urinary tract infection (two items); (5) rectal problems (Rectal, three items); and (6) physiological symptoms of general stress/tension (Stress, six items). The remaining five items are unscored (Appendix S1). A mean score for each of these subscales is calculated, and a total score for the AHPFS is calculated by a sum of the mean of the subscales giving a range from 6 to 30. A report on the validity of the Dutch version has never been published but it has been described in an article by Postma et al.[17] The questionnaire has never before been translated into Norwegian or tested in a Norwegian setting. The aim of this study was therefore to translate the AHPFS into Norwegian to assess the validity and reliability in a Norwegian sample of women with symptoms associated with abnormally high pelvic floor muscle tone, and to present a valid factor structure.

2 | MATERIAL AND METHODS

In this translation of the Dutch questionnaire AHPFS into Norwegian and the validation study, we have used a triangulation of methods involving the opinions of experts, cognitive interviews and a quantitative survey. The study has been performed in accordance with the Consensus-based Standards for the Selection of Health Measurement Instruments (COSMIN) guidelines).[18,19]

2.1 | Translation and cultural adaptation

First, three bilingual Norwegian–Dutch health personnel independently translated the AHPFS from Dutch to Norwegian. In the second step, the multidisciplinary research group (the authors) discussed the wording, possible sources of misunderstanding, cultural aspects, similarities and differences between the three translated Norwegian versions, and adapted them to form a single preliminary Norwegian version (Appendix S1). Third, a back-translation to Dutch was performed by a professional bilingual translator, and a review of the back-translation was conducted by the developers of the Dutch questionnaire for a comparison with the original tool after the translation process. The back-translation revealed no problematic aspects regarding the items.

In step four, the Norwegian preliminary AHPFS version of the translated 30 items was pilot-tested through feedback from an invited group of clinical experts and through cognitive interviews with patients from the target group, which the assignment was to assess the questionnaire's comprehensiveness, readability and equivalence.[20] A paper version of the questionnaire was used in this phase. Amendments were made between each step, and gradually a comprehensible 30-item Norwegian version, named Amsterdam Complex Pelvic Pain Symptom Scale (ACPPS-30) [in Norwegian: Spørreskjema om sammensatte underlivssmerter – ASUS-30] emerged ready for field-testing of its psychometric properties in the fifth step (Figure 1).

2.2 | Participants and data collection

In our study, three groups were recruited. The first was a sample of six clinical experts consisting of three gynecologists, two pelvic floor physiotherapists, and one sexologist. They were recruited to evaluate the preliminary Norwegian AHPFS's (Appendix S1) comprehensiveness, relevance, and wording one-by-one. The correspondence was via email.

Secondly, a sample of 11 female patients with pelvic floor dysfunction were recruited from the outpatient clinic at St. Olav's Hospital to evaluate the meaning and wording of the translated items. Patients received information about the study prior to an appointment at the outpatient clinic and were interviewed by a study nurse upon arrival. Patients who participated in the cognitive interviews were not enrolled in the subsequent part of the study.

Thirdly, to test the translated, slightly modified and the psychometric properties of the renamed ACPPS-30, three different groups of women were invited to answer the electronic questionnaire: (1) women referred to the gynecological outpatient clinic or to a pelvic floor physiotherapist at St. Olav's Hospital, (2) members of the Vulvodynia or the Endometriosis Patient Associations, and (3) female students and employees at the Faculty of Medicine and Health Sciences, Norwegian University of Science and Technology (NTNU). Inclusion criteria were women over the age of 18 years, nonpregnant, able to consent to participation, and able to read, write and understand Norwegian. To identify women with complex pelvic pain complaints in this broad sample of women, only women with symptoms, eg an ACPPS-30 total score of ≥11

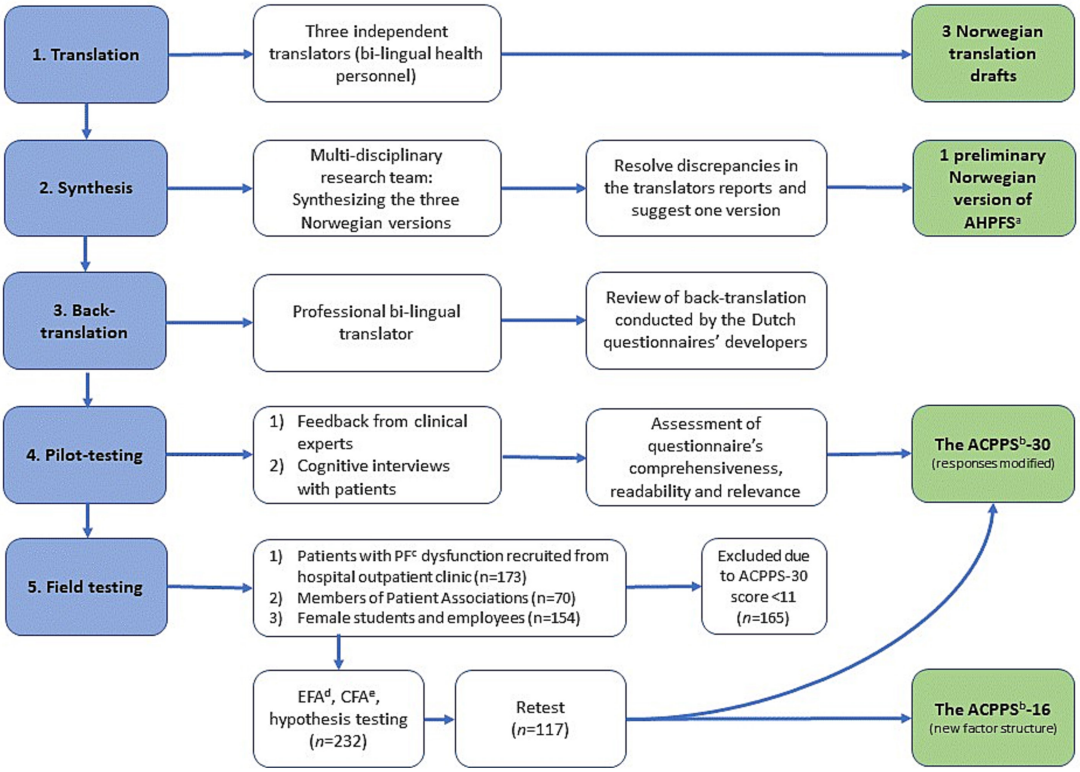

FIGURE 1 Flowchart of five steps from translation to psychometric testing and analyses. [a]Amsterdam Hyperactive Pelvic Floor Scale; [b]Amsterdam Complex Pelvic Pain Symptom Scale; [c]Pelvic Floor; [d]Exploratory Factor Analysis; [e]Confirmatory Factor Analysis.

according to the 25-item prespecified factor structure from the developers of the Dutch questionnaire, were included in the subsequent analysis (Figure 1). This cutoff has been used in a previous study to discriminate between patients with and without symptoms but has not been validated.[17] Eligible patients were invited to participate in the study from March to May 2022. Patients from St. Olav's Hospital were invited by a short message included in the hospital's SMS call-in routines, encouraging them to log on to an electronic version of the questionnaire. A reminder was sent a couple of days before the given appointment at the hospital. In addition, nurses and secretaries at the outpatient clinic encouraged the patients to participate upon arrival. Posters and screens were on hand with study information and a QR code leading directly to the questionnaire in the hospital waiting area. Members from the Vulvodynia and Endometriosis Patient Associations were recruited through their open-access social media groups and webpages. Female students and employees at the NTNU were invited to participate in the study through the intranet, email and social media groups.

After 2 weeks, a retest was automatically sent to all responders. The retest contained only the ACPPS-30. In addition, participants were asked if their condition had changed during the interim period[20] with the question, "Compared with the first time you completed the questionnaire, has your pelvic floor condition changed?" If the answer was "Yes", women were excluded from the test–retest validation.

2.3 | Measures

The Amsterdam Hyperactive Pelvic Floor Scale is a Dutch self-reported instrument which in this study was translated into Norwegian and assessed for its psychometric properties. The instrument contains 30 items concerning complex pelvic floor symptoms associated with an abnormally high pelvic floor muscle tone.[17] As demonstrated in Figure 1, a preliminary Norwegian version (Appendix S1) was used in pilot-testing of multidisciplinary clinical experts and women from the target group; a final Norwegian version, renamed ACPPS-30, was used in the field-testing (Appendix S2).

Background and gynecological variables such as age, mother tongue, educational level, employment, marital status, prior deliveries and menopausal status were collected. The following measures were used in hypothesis testing.

Pelvic Floor Distress Inventory (PFDI-20) is a self-reported instrument divided into three subscales: Pelvic Organ Prolapse Distress Inventory (POPDI-6), Colorectal–Anal Distress Inventory (CRADI-8) and Urinary Distress Inventory (UDI-6), of which only the last two were used in the hypothesis testing. Subscale scores are converted to a range from 0 to 100. In all scales, a higher score indicates greater distress.[12]

Female Sexual Function Index (FSFI) is a self-reported instrument with 19 items divided into six subscales: Desire, Arousal, Lubrication, Orgasm, Satisfaction and Pain, in which only the last subscale was used in the hypothesis testing. FSFI-Pain subscale scores range from 0 to 6.[15]

Mental Health Inventory-5 (MHI-5) is a brief, reliable and valid international instrument for assessing mental health in adults.[21] The MHI-5 has a score of 0–100, with a score of 100 representing optimal mental health.

Fibromyalgia Survey Criteria 2016 contains two parts: (1) the Widespread Pain Index summarizes the number of body areas (among 19 possible) in which the patient experiences pain and gives a score from 0 to 19, and (2) the Symptom Severity Score gives a score based on the extent of memory problems, fatigue, sleep quality, depression, cramps in lower abdomen and headache, with a score from 0 to 12.[22]

2.4 | Statistical analyses

Descriptive statistics, principal component analysis (PCA) and correlation analyses were performed with IBM SPSS version 28. Confirmatory factor analysis (CFA) was performed using STATA 17.0.[23] Due to settings in the electronic questionnaire, there were no missing data in the ACPPS-30 questionnaire. However, the Dutch AHPFS items 1, 4, 5 and 6 had "not applicable" (NA) options. According to scoring instructions from the Dutch developers, items answered "NA" were replaced with the mean of the remaining responses belonging to this item. No replacements were made for missing data in the other questionnaires.

Content validity was assessed in two ways. The first method was a review of the preliminary Norwegian AHPFS version by clinical experts and cognitive interviews with patients (Appendix S1). In the second method, the floor and ceiling effects of the new factors in ACPPS-16 were calculated and perceived as problematic if more than 15% of respondents achieved the highest- or lowest-possible scores (Appendix S2).[24]

Construct validity was assessed through structural validity where the dimensionality of the ACPPS-30 was assessed, and by hypothesis-testing comparing the constructs with other relevant existing questionnaires (Figure 1). First, structural validity was assessed through exploratory factor analysis and CFA. We performed exploratory factor analysis using PCA with varimax rotation, which reduces and enables dimensions to identify correlations in our dataset.[25] Thereafter, a CFA analysis was performed. The CFA is a hypothesis-driven approach to a possible factor structure. In this case the factor structure had already been prespecified by the questionnaire's Dutch developers. Analysis was performed to assess the covariance captured by the factors, evaluating goodness of fit to see how well the model fit the observed data. In CFA, a high loading (preferably >0.4) of an item indicates that the factor and the respective item share a common variance.[26,27] Range of fit indices were used to assess the relation between the observed data and the theoretical model, that is, the fit of the measurement model: χ^2/degrees of freedom (df) χ^2/df (≤2 good fit, ≤3 acceptable).[27] Skewness and kurtosis were present when inspecting the assumption of normality, and the Satorra–Bentler-corrected χ^2 was applied as recommended when analyzing nonnormal continuous endogenous variables.[26] Furthermore, the root mean square error of approximation (RMSEA) (≤0.05 good fit, ≤0.10 acceptable) and the standardized root mean square residual (SRMR) (≤0.05 good fit, ≤0.10 acceptable), the comparative fit index (CFI) (≥0.95 good fit, ≥0.90 acceptable), and the Tucker–Lewis index (TLI) (≥0.95 good fit, ≥0.90 acceptable) were used.[28] Correlations were calculated between the ACPPS-16 (total score and four of its subscales) and other validated scales, and coefficients were considered low (<0.30), moderate (0.30–0.59) or high (≥0.60).

Reliability of the questionnaire was assessed for internal consistency and stability over time. Internal consistency is a part of the assessment of the questionnaire's dimensionality and reliability coefficients of Cronbach's alpha (α) and composite reliability (ρc), with values ≥0.7 considered to be good.[20] Stability over time with a test–retest was assessed using intraclass correlation coefficients (ICC) for scales (ACPPS-16) and weighted kappa values for single items (ACPPS-30).[24] ICCs ≤0.70 and kappa values ≤0.6 are considered good.[20]

2.5 | Ethics statement

Approval was granted by the Regional Committee for Medical and Health Research Ethics on June 21, 2021 (#245815), the Norwegian Centre for Research Data on March 8, 2022 (#607016) and the institutional review board at the Department of Obstetrics and Gynecology at St. Olav's Hospital, Trondheim, on December 14, 2021 (#2021/14758). Permission was also granted by the developers of the Dutch AHPFS instrument.

Written informed consent was obtained from all patients participating in the cognitive interviews. Women participating in the survey received information about the study electronically and consented to participate by pushing the consent button of the web-based questionnaire. Respondents were informed that participation in the study was voluntary and that they could withdraw their data from the study at any time.

3 | RESULTS

The study population included in the statistical analysis of the field-testing of psychometric properties comprised 232 women (Figure 1), of whom 46% were patients from St. Olav's Hospital ($n = 107$), 26% women from the Vulvodynia and Endometriosis Patient Associations ($n = 61$) and 28% students and employees at the NTNU ($n = 64$). Their ages ranged between 19 and 82 years, with a mean of 37.5 years. See Table 1 for further demographic characteristics.

3.1 | Content validity

Feedback from clinical experts and cognitive interviews with patients from the target group were conducted on the 30-item

TABLE 1 Demographic characteristics of participants, $n = 232$.

Variables	*n*	%
Recruited (total $n = 232$) from		
St. Olav's hospital	107	46.1
Patient Association[a]	61	26.3
NTNU[b] student or employee	64	27.6
Education (total $n = 231$)		
Primary	7	3.0
Secondary	88	38.1
Higher	136	58.9
Language (total $n = 231$)		
Norwegian	223	96.5
Other	8	3.5
Marital status (total $n = 228$)		
Not in relationship	55	24.1
In relationship or one partner	173	75.9
Parity (total $n = 231$)		
Nullipara	109	47.2
One or two	85	36.8
Three or four	34	14.7
Five-seven	3	1.3
Menopause (total $n = 231$)		
Yes	46	19.9
No	185	80.1

[a]Patient Association for women with Vulvodynia and for Endometriosis.
[b]Norwegian University of Science and Technology.

preliminary Norwegian version of AHPFS in the pilot testing (Figure 1). Based on the feedback from the clinical experts, the preliminary Norwegian version of AHPFS's response categories were improved and quantified from "Never" to "Almost never or never", from "Sometimes" to "Occasionally (less than half of the time)", from "Regularly" to "Sometimes (half the time)", from "Often" to "Usually (more than half of the time)" and from "Very often" to "Almost always or always". Additionally, in the questions regarding sexual function (items 1, 4, 5 and 6), the response category NA was replaced with "No sexual activity". "No menstruation" was added as a sixth category to item 9 (Menstruation pain) (Appendix S2). These improvements were made in agreement with the original developer, Ellen T. M. Laan. Based on results from the cognitive interviews with patients from the target group, a clarifying sentence was added to the questionnaire instructions to inform users that some items were about sexual function and others about more general complaints. Based on feedback from both groups, the preliminary Norwegian version of the AHPFS was regarded as relevant, comprehensive and easy to read.

Floor and ceiling effects were measured on the five constructs after the dimensional structure had been assessed and revised. Only one ACPPS-16 construct, the Rectal symptoms construct, had the lowest possible score, occurring with a frequency of more than 15% (eg 17.7%). Hence, a minor floor effect was found in the total score distributions of one construct. No construct had the highest possible score with more than 15% frequency, hence, there were no ceiling effects.

3.2 | Construct validity

To explain as much of the total variance as possible with as few factors as possible, the 30 items of ACPPS-30 were subjected to PCA. The Kaiser–Meyer–Olkin measure of sampling adequacy was 0.76, exceeding the recommended value of 0.60, and Bartlett's test of sphericity showed statistical significance ($P < 0.0001$), supporting the factorability of the correlation matrix.[29] When subjecting the ACPPS-30 to PCA, we searched for the cleanest factor structure with loadings >0.32.[29] The developers of the original Dutch AHPFS questionnaire had provided an (unpublished) scoring instruction containing six dimensions including 25 of the 30 items; therefore, we expected a six-dimensional structure with correlated factors.

Nine factors with eigenvalue ≥1.0 were extracted (Table 2), showing factor loadings ranging from 0.32 to 0.83 and explaining 63.1% of the variance.

This PCA-suggested solution revealed nine factors comprising between two and four items each and several cross-loadings. Cronbach's alpha coefficients ranged between 0.38 and 0.75. Two factors had a Cronbach's alpha of ≥0.70 and seven factors had a Cronbach's alpha of <0.70 (Table 2). Since the Dutch AHPFS developers had already removed five items (items 3, 8, 9, 20 and 24) from the scales' factor structure, a similar 25-item PCA was also tested, still with an unsatisfactory result (results not shown). Hence, the factor structure seemed indecisive, and we proceeded with CFA.

In CFA, we tested the six-dimensional original proposed factor structure involving 25 of the 30 items suggested by the questionnaire developers and used in other studies.[17] This model would not converge in STATA as the latent factor Urinary tract infection contained only two items. The model was therefore simplified by excluding the Urinary tract infection factor, with the intention of re-inserting it after an easier-to-fit model was produced. The starting model, termed Model-1, therefore involved 23 of 25 items and 5 factors. This model exposed factor loadings (λ) ranging between 0.36 and 0.84, with squared multiple correlations (R^2-values) ranging from 0.13 (item 13) to 0.70. The fit was bad (Satorra–Bentler $\chi^2 = 486.341$, $df = 220$, $\chi^2/df = 2.21$, $P = 0.0001$, RMSEA = 0.073, *P*-value for test of close fit = 0.0001, CFI = 0.79, TLI = 0.76, and SRMR = 0.084). The estimated χ^2 value was acceptable, whereas the other fit indices indicated misspecification. Reliability assessed by the composite reliability coefficient was good for all five dimensions (Table 3).

Scrutinizing factor loadings, residuals and modification indices (MI) revealed no significant residuals. However, several other possible changes for our model – for instance, three factor loadings <0.4, where item 13 had the lowest (0.36) loading, in addition to the lowest R^2-value (0.13), which needed to be dealt with. Item 13 concerned *constipation or feelings of incomplete defecation*, which

TABLE 2 Principal component analysis with varimax rotation[a] of the ACPPS-30. Estimates for factor loadings, extraction sums of squared loadings and Cronbach's alpha, $n = 232$.

	Component								
Items	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
15. Having to push or strain to be able to urinate	0.83								
17. Poor stream, hesitancy, terminal dribbling or incomplete voiding at micturition	0.83								
14. Frequent urination (more than 10 times per 24 hours)	0.66								
19. Bladder pain	0.43	0.34						0.39	
10. Abdominal pain not related to menstruation		0.71							
12. Alternating periods of diarrhea and constipation		0.71							
11. Abdominal cramps		0.70							
13. Constipation or feelings of incomplete defecation		0.53				0.37			
3. Pain or a burning sensation at the skin of the vaginal entrance while sitting/biking/wearing tight clothing			0.81						
2. Painful, burning, stinging spots or tears at the skin of the vaginal entrance			0.80						
7. Persistent feelings of genital swelling	0.35		0.58						
20. Pain in the area between vagina and anus (perineum)		0.32	0.47						
30. Neck/shoulder pain				0.75					
29. Headache				0.734					
25. Lower back pain		0.35		0.57					
24. Pain in the tailbone				0.44					
6. Pain upon deep thrusting of the penis during sexual intercourse					0.78				
5. Pain in the genitals during or after orgasm					0.71				
1. Pain with vaginal penetration and/or a burning sensation after sexual intercourse			0.54		0.61				
4. A lack of vaginal lubrication during sexual intercourse					0.50				
23. Anal fissures or tears						0.80			
22. Hemorrhoids						0.79			
21. Anal pain		0.32				0.63			
27. Teeth grinding							0.80		
28. Feelings of tightness in jaw muscles				0.38			0.70		
18. Urinary tract infections			0.43					0.67	
26. Hyperventilation							0.44	0.55	

(Continues)

TABLE 2 (Continued)

Items	Component 1	2	3	4	5	6	7	8	9
16. Sudden, compelling urge to urinate (urinary urgency)	0.48							0.55	
8. Cumbersome vaginal discharge									−0.75
9. Menstrual pain									0.71
Eigenvalues	19.75	7.64	7.25	6.88	5.42	4.68	4.28	3.74	3.46
Cumulative % of total variance explained	19.75	27.39	34.64	41.51	46.93	51.61	55.89	59.63	63.09
Cronbach's Alpha (number of items)	0.74 (4)	0.69 (4)	0.75 (4)	0.68 (4)	0.67 (4)	0.69 (3)	0.66 (2)	0.45 (3)	0.38 (2)

[a]Rotation converged in 9 iterations.

is a known symptom of Irritable bowel syndrome. Irritable bowel syndrome is, however, also characterized by *alternating periods of diarrhea and constipation*, which is expressed in item 12 instead of constipation alone. Thus, item 13 seemed redundant, so we removed item 13 and ran CFA once more. This solution, termed Model-2 (now including 22 items), showed only a slightly improved fit (Table 3).

Again, guided by the factor loadings, modification indices, residuals and the nuances of the actual construct, we found that item 27 (*Teeth grinding*) shared a very high MI with item 28 (*Feelings of tightness in jaw muscles*), signifying that these items shared error variance, which is logical, as both items are dealing with tenseness in the jaw. Item 27 had a low factor loading (0.37) and a low R^2-value (0.14), suggesting lower reliability. Therefore, we removed item 27 and ran the CFA once more. Nevertheless, Model-3 (now including 21 items) still revealed a poor fit (Table 3). Furthermore, MIs revealed that item 7 cross-loaded strongly on other factors (Lower urinary tract symptoms and Stress). Item 7 dealt with *the persistent feelings of genital swelling*. By contrast, other items of the Provoked vulvodynia factor are concerned with incidents that would provoke pain in the vulvar area. Hence, item 7 was removed, giving us Model-4 with a slightly improved fit (Table 3).

Thus far, we had removed items 13, 27 and 7. Nevertheless, though the χ^2/df was good, the fit remained poor. Item 2 *(Painful, burning, stinging spots or tears at the skin of the vaginal entrance)* showed very high MIs, with several items (1, 5 and 6), signifying that these items shared error variance, which is logical: individuals experiencing painful or burning fissure/soreness in the vagina will also experience this *during or after penetration* (item 1), *during or after orgasm* (item 5) and *upon deep thrusting of the penis during sexual intercourse* (item 6). Therefore, we dismissed item 2 and ran the CFA once more in Model-5, with a further improved fit (Table 3). Again, MIs revealed that item 25 cross-loaded with several other factors (Provoked vulvodynia, Irritable bowel syndrome and Lower urinary tract symptoms). Item 25 (*Lower back pain*) may theoretically be connected to stress but could also have several other explanations and we therefore removed it. This gave us Model-6, which was closer to an acceptable fit (Table 3). However, item 4 demonstrated an unacceptably low factor loading (0.29) and R^2 (0.08), and was conceptually different from the remaining items in the Provoked vulvodynia factor. Item 4 deals with *lack of vaginal lubrication during sexual intercourse*, and the remaining items in this factor focus on provoked pain during sexual activity. Item 4 was therefore removed, giving us Model-7 with a further improved fit (Table 3). Nevertheless, item 26 (*Hyperventilation*) had a poor factor loading of 0.38 and an R^2-value of 0.15, and was therefore removed, resulting in the final model, Model-8, with an acceptable fit (Table 3).

The latent factor Urinary tract infection with its two items was reinserted into the adapted model but was still unable to converge and was therefore removed again. At this stage, we had removed seven items, in addition to the Urinary tract infection factor, and the model fit was considerably improved. We had a 16-item model with five factors containing three to four items in each factor. All items had substantial loadings significant at the 0.001 level. The standardized loadings ranged from 0.44 to 0.83 (Table 4). Although composite reliability is expected to decrease with fewer items, the factors demonstrated a good composite reliability between 0.70 and 0.77 (Table 4). Hence, Model-8 with five factors and 16 items provided the most parsimonious model with a good fit (Figure 2).

The constructs of this modified factor structure, named ACPPS-16, were further assessed through hypothesis testing where the new constructs were compared with similar scales. The correlations between the ACPPS-16 total score and subscales, with FSFI-Pain, CRADI-8, UDI-6, Fibromyalgia Survey Criteria 2016 and MHI-5, supported hypotheses H1– H5 (Table 5), all concerning construct convergent validity of the ACPPS-16 subscales.

3.3 | Reliability

In total, 117 women with a score ≥11 according to the 25-item prespecified factor structure from the Dutch developers responded to

TABLE 3 Goodness-of-fit measures for measurement models. Confirmatory Factor Analysis for Model-1 to Model-8, $n = 232$.

Fit measure	Model-1 *5-factors 23 items*	Model-2 *5 factors 22 items*	Model-3 *5 factors 21 items*	Model-4 *5 factors 20 items*	Model-5 *5 factors 19 items*	Model-6 *5 factors 18 items*	Model-7 *5 factors 17 items*	Model-8 *5 factors 16 items*
χ^2	486.341	433.179	370.138	313.138	235.465	195.263	168.181	135.099
P-value	0.0001	0.0001	0.0001	0.0001	0,0001	0.0001	0.0001	0.0001
	2.21 (df[a] = 220)	2.17 (df = 199)	2.07 (df = 179)	1.96 (df = 160)	1.66 (df = 142)	1.56 (df = 125)	1.54 (df = 109)	1.44 (df = 94)
RMSEA[b]	0.073	0.072	0.068	0.065	0.053	0.049	0.49	0.044
P-value (close fit test)	0.0001	0.0001	0.0001	0.004	0.166	0.352	0.388	0.629
SRMR[c]	0.084	0.083	0.082	0.078	0.072	0.067	0.065	0.062
CFI[d]	0.79	0.81	0.84	0.86	0.90	0.92	0.93	0.95
TLI[e]	0.76	0.78	0.81	0.84	0.88	0.91	0.92	0.94
ρc[f]	0.70–0.76	0.70-0.76	0.70–0.76	0.70–0.76	0.67–0.76	0.67–0.76	0.70–0.76	0.70–0.76

Note: Model-1: 23 items 5-factor solution, Model-2: 22-items 5-factor solution (item 13 removed), Model-3: 21 items 5-factor solution (item 13 and 27 removed), Model-4: 20-items 5-factor solution (items 13, 27 and 7 removed), Model-5: 19-items 5-factor solution (items 13, 27, 7 and 2 removed), Model-6: 18-items 5-factor solution (items 13, 27, 7, 2 and 25 removed. Model-7: 17-items 5-factor solution (items 13, 27, 7, 2, 25 and 4 removed. Model-8: 16-items 5-factor solution (items 13, 27, 7, 2, 25, 4 and 26 removed.

[a]Degrees of freedom.

[b]Root mean square error of approximation.

[c]Standardized root mean square residual.

[d]Comparative Fit Index.

[e]Tucker–Lewis Index.

[f]Composite reliability, Raykov's factor reliability coefficient $\rho c = \frac{(\sum \lambda)^2}{\left[(\sum \lambda)^2 + \sum(\theta)\right]}$.

TABLE 4 Model-8[a]: the best fitting measurement model, $n = 232$.

Item	Parameter	STATA estimate[b]	*t*-value[c]	R^{2}[d]
PVD[e] symptoms				
ACPPS-1	λx 1,1	0.51	7.97	0.26
ACPPS-5	λx 5,1	0.83	16.04	0.69
ACPPS-6	λx 6,1	0.64	12.68	0.41
IBS[f] symptoms				
ACPPS-100	λx 10,2	0.78	16.73	0.61
ACPPS-11	λx 11,2	0.81	20.28	0.65
ACPPS-12	λx 12,2	0.47	7.42	0.22
LUTS[g] symptoms				
ACPPS-14	λx 14,3	0.52	9.77	0.27
ACPPS-15	λx 15,3	0.82	21.93	0.67
ACPPS-16	λx 16,3	0.44	6.92	0.20
ACPPS-17	λx 17,3	0.83	19.90	0.68
Rectal symptoms				
AHPFS-21	λx 21,4	0.54	8.31	0.29
AHPFS-22	λx 22,4	0.63	9.48	0.40
AHPFS-23	λx 23,4	0.81	13.36	0.26
Stress symptoms				
ACPPS-28	λx 28,5	0.52	9.26	0.42
ACPPS-29	λx 29,5	0.69	13.19	0.53
ACPPS-30	λx 30,5	0.80	17.44	0.66
Factor				
ρ_c[h] PVD	ρ_c		0.70	
ρ_cIBS	ρ_c		0.73	
ρ_cLUTS	ρ_c		0.76	
ρ_cRectal	ρ_c		0.70	
ρ_cStress	ρ_c		0.70	

[a]Five-factor solution including 16 items (UTI factor, and items 13, 27, 7, 2, 25, 4, 26 removed).
[b]Completely standardized factor loadings.
[c]Significant at the 1% level.
[d]Bentler–Raykov squared multiple correlation coefficient.
[e]Provoked Vulvodynia.
[f]Irritable Bowel Syndrome.
[g]Lower Urinary Tract Symptoms.
[h]Composite reliability $\rho c = \frac{(\sum \lambda)^2}{(\sum \lambda)^2 + \sum(\theta)}$.

the retest in the field-testing (Figure 1). The test–retest stability revealed ICCs between 0.85 and 0.93 (Table 6). Concerning the stability of single items, the weighted kappa revealed seven items with values between 0.80 and 0.90, 21 items with values between 0.60 and 0.79, and two items <0.6.

4 | DISCUSSION

In this study, the Dutch questionnaire AHPFS was translated into Norwegian, and its psychometric properties were tested among women with symptoms associated with abnormally high pelvic floor muscle tone. Content/face validity demonstrated that the questionnaire was perceived as relevant and understandable, but some adjustments were made in the instructions and wording of response categories. The Norwegian version was given a Norwegian name, which directly translated to English is: "Amsterdam Complex Pelvic Pain Symptom Scale - 30 (ACPPS-30)." Assessment of the questionnaire's dimensionality revealed a five-factor structure with 16 items similar to the original Dutch AHPFS factor structure, but without the Urinary tract infection factor and another seven items. Hence, a modified and shortened version of the Dutch AHPFS factor structure was presented in the ACPPS-16 (Appendix S2). The ACPPS-16 total score and subscales correlated as expected with

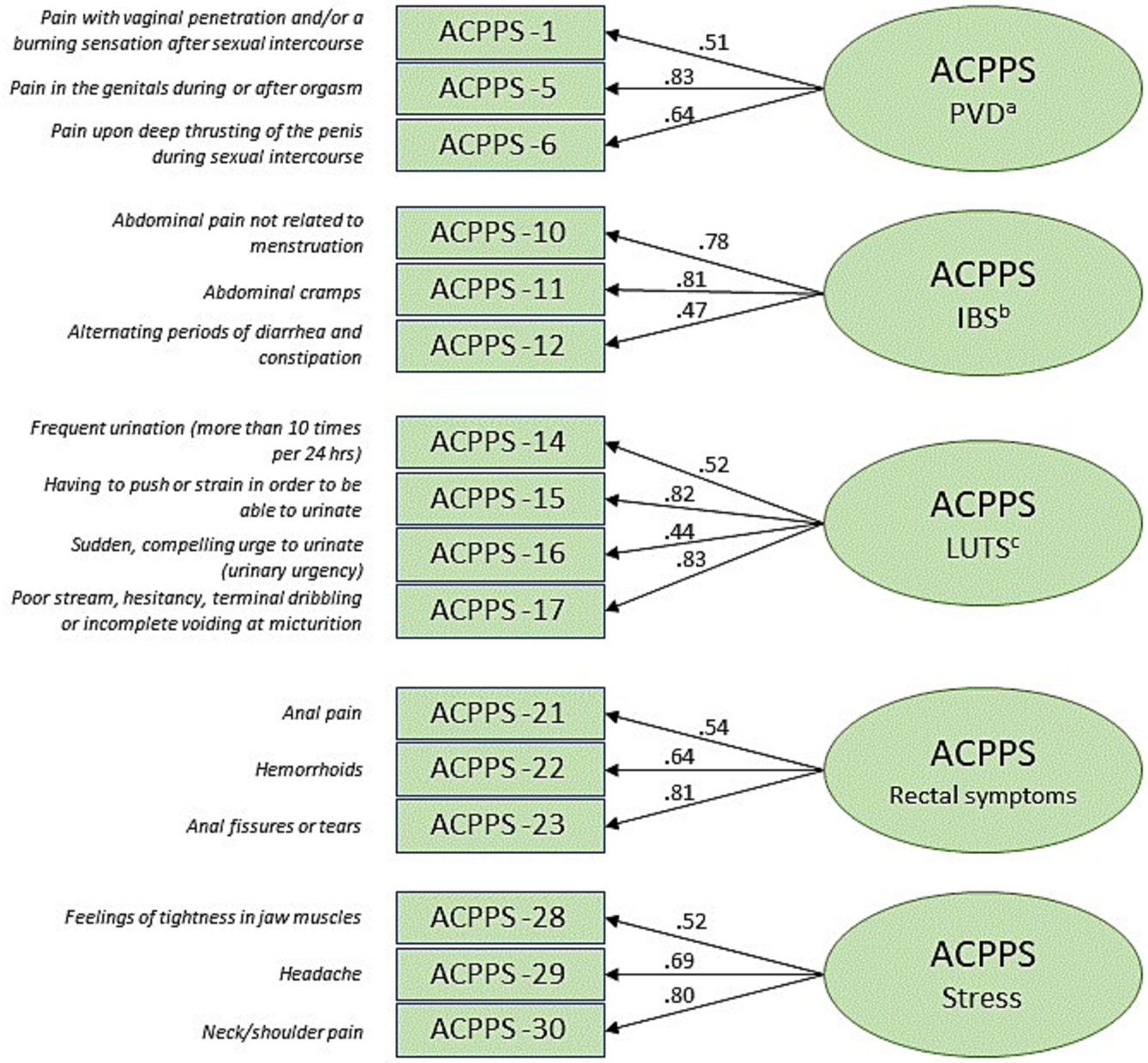

FIGURE 2 The best fitting measurement model of the Amsterdam Complex Pelvic Pain Symptom Scale (ACPPS). [a]Provoked Vulvodynia; [b]Irritable Bowel Syndrome; [c]Lower Urinary Tract Symptoms.

scales measuring similar conditions. Test–retest reliability demonstrated good stability for scales (ACPPS-16) and single items (ACPPS-30).

Assessment of the translated questionnaire's face and content validity and subsequent amendments gave a modified Norwegian version, ACPPS-30, with items corresponding well with the constructs intended to be measured.[20] The original Dutch questionnaire's items are based on theory and evidence about an abnormally high pelvic floor tone and its symptoms and characteristics, and the items were explored through an exploratory factor analysis that concluded with 25 items and 6 factors.[17] However, our assessment of the ACPPS-30 dimensional structure suggested that these results may have been premature and that further assessment was necessary.

Hence, we performed a new exploratory factor analysis to validate the factor structure of the Norwegian version, ACPPS-30. Our PCA suggested nine factors, with six substantial dimensions involving three to four items, accompanied by three weaker factors involving two to three items each. Moreover, our PCA revealed several cross-loadings. Since the original Dutch AHPFS version contained six factors, the dimensionality now seemed indecisive. However, conclusions should not be drawn based solely on a PCA. Therefore, a CFA was performed. However, the model seemed troublesome, indicating misspecification. First of all, we had to remove the Urinary tract infection factor, since the six-factor model would not converge in STATA with the factor included. The Urinary tract infection factor had only two items, whereas there should be a minimum of three

TABLE 5 Testing of hypothesis H1-H5 concerning correlation (*r*) between ACPPS[a]-16 Total score and subscales with validated scales, $n = 232$.

Hypothesis tested			
Expected way of correlation	**Between**	**Correlation coefficient (*r*)[i]**	**Confirmed?**
Negative	H1. ACPPS-PVD[b] and FSFI[c] – Pain	−0.55	Yes
Positive	H2. ACPPS-IBS[d] and CRADI-8[e]	0.42	Yes
Positive	H3. ACPPS-LUTS[f] and UDI-6[g]	0.68	Yes
Positive	H4. ACPPS-Stress and Fibromyalgia 16 criteria	0.57	Yes
Negative	H5. ACPPS Total score and MHI-5[h]	−0.40	Yes

[a]Amsterdam Complex Pelvic Pain Symptom Scale.
[b]Provoked Vulvodynia.
[c]Female Sexual Function Index.
[d]Irritable Bowel Syndrome.
[e]Colorectal-Anal Distress Inventory-8.
[f]Lower Urinary Tract Symptoms.
[g]Urinary Distress Inventory-6.
[h]Mental Health Inventory-5.
[i]$P < 0.01$.

items contributing to a factor.[20,27] Next, we needed to address the misspecifications of the five-factor model. To achieve a good model fit, some items had to be removed one by one, namely items 13, 27, 7, 2, 25, 4 and 26. In our study, these items either explained very little of the variance in the respective construct, subsequently providing a low factor loading, or the item's wording was too similar, so that respondents seemed to assume that some of the items sought to assess roughly the same thing as other items, which generated a substantial correlated error variance. Hence, these seven items hampered the model fit and were removed. These seven items, together with the two items from the removed Urinary tract infection factor and five items already removed from the original Dutch prespecified factor structure, gave a total of 14 unscored items. Reliability and structural validity both relate to the sufficiency of a scale's items. Good indicators of a factor show highly significant factor loadings, accompanied by strong squared multiple correlations (R^2), which represent how much variation in an item is explained by the latent construct.[30] In our modified model all loadings were significant at the 1% level. The factor loadings were reliable: seven were excellent (>0.70), three were good (>0.55–0.70) and six were good to fair (0.55–0.45)[31] (Table 4). The model had factors with a good composite reliability (ρc), demonstrating a good internal consistency ($\rho c > 0.7$). Accordingly, the solution had good reliability and clear dimensionality. In our investigation, the prespecified six-factor structure with 25 items did not fit well with the data. However, our Model-8, including five of the six factors and only 16 items, was the most parsimonious model with a good statistical fit.

Construct validity measures whether indicators reflect the theoretical latent construct the items are designed to measure.[32] In the present study, all hypotheses regarding a correlation (H1–H5) were supported. No hypothesis was formulated for the Rectal construct due to a lack of suitable scales to compare it with. The association between the Lower urinary tract symptoms factor and Urinary Distress Inventory-6 was high (0.68), as expected since the four items in the Lower urinary tract symptoms factor cover to a large extent the Urinary Distress Inventory-6 construct of urinary distress. The other correlations were all moderate (0.40–0.57) since all these scales measure similar, but not equivalent, constructs.

When tested, the new constructs in the ACPPS-16 demonstrated good reliability in terms of internal consistency, with composite reliability >0.7 for all five constructs. Previous studies of the AHPFS questionnaire have not reported composite reliability, only Cronbach's alphas with acceptable values.[17] The stability over time was measured by a test–retest for the ACPPS-16 and demonstrated good to excellent reliability for all five subscales. Furthermore, the ACPPS-30 demonstrated almost perfect stability (weighted kappa >0.80) for seven items, substantial stability (0.6–0.79) for 21 items, moderate stability for one item (0.56), and a fair weighted kappa of 0.34 for one item (item 9).[20] However, this latter item (item 9) concerning menstrual pain and the timing of the natural menstrual cycle could influence and bias the participant's response to such a question.

In this study we have translated a Dutch questionnaire, in English called "The Amsterdam Hyperactive Pelvic Floor Scale", to a Norwegian version, "Amsterdam Complex Pelvic Pain Symptom Scale-30 (ACPPS-30)" and assessed its psychometric properties. In our investigation, a modification of the factor structure including five of the six factors and only 16 items, was the most parsimonious model with a good statistical fit. The factors Lower urinary tract symptoms and Rectal symptoms ended up identical to the original

TABLE 6 Weighted Kappa for single items in ACPPS[a]-30 and ICC for subscales in ACPPS[a]-16, $n = 117$.

	Weighted Kappa[b]	95% CI[c]	ICC[d]	95% CI
PVD[e] symptoms				
Q1	0.75	0.62–0.87	0.91	0.86–0.93
Q5	0.74	0.62–0.87		
Q6	0.85	0.76–0.94		
IBS[f] symptoms				
Q10	0.68	0.56–0.80	0.89	0.85–0.93
Q11	0.69	0.59–0.79		
Q12	0.75	0.67–0.83		
LUTS[g] symptoms				
Q14	0.77	0.69–0.85	0.85	0.78–0.90
Q15	0.70	0.60–0.81		
Q16	0.69	0.58–0.81		
Q17	0.60	0.46–0.74		
Rectal symptoms				
Q21	0.72	0.60–0.83	0.92	0.89–0.95
Q22	0.90	0.86–0.94		
Q23	0.76	0.67–0.86		
Stress symptoms				
Q28	0.84	0.79–0.89	0.93	0.90–0.95
Q29	0.83	0.76–0.90		
Q30	0.75	0.67–0.84		
Single unscored items – outside the factor structure				
Q2	0.82	0.74–0.89	N/A	
Q3	0.76	0.63–0.88		
Q4	0.56	0.39–0.73		
Q7	0.63	0.46–0.80		
Q8	0.80	0.72–0.88		
Q9	0.34	0.18–0.50		
Q13	0.70	0.60–0.80		
Q18	0.66	0.49–0.83		
Q19	0.75	0.64–0.86		
Q20	0.61	0.46–0.75		
Q24	0.79	0.68–0.90		
Q25	0.79	0.73–0.85		
Q26	0.73	0.61–0.86		
Q27	0.86	0.79–0.93		

[a]Amsterdam Complex Pelvic Pain Symptom Scale.

[b]Quadratic.

[c]Confidence Intervall.

[d]Intraclass Correlation Coefficient, two-way mixed, absolute agreement, average measure.

[e]Provoked Vulvodynia.

[f]Irritable Bowel Symptoms.

[g]Lower Urinary Tract Symptoms.

version. In research, a statistically well-functioning measurement model is required. Hence, to be used in clinical studies using, for example, structural equation modeling or regression analysis, this five-factor model is superior (ACPPS-16). This 16-item version is also valid in a clinical context. The questionnaire may be a useful short instrument in multidisciplinary teams for patients with complex pelvic floor dysfunctions related to an abnormally high pelvic floor muscle tone. The 14 unscored items removed from the ACPPS-16 may, however, support the clinical relevance and be used in a clinical context to provide more information about the patient's condition. Hence, the 30-item ACPPS-30 should be a preferred version in clinical practice.

We have established a full 30-item version of the Dutch AHPFS with minor modifications in wording of response categories and instructions. ACPPS-30 has been assessed for face/content validity and test–retest stability. We have also established a shorter version with only 16 of the translated items distributed among five factors associated with complex pelvic pain symptoms. These constructs have been assessed for construct validity, internal consistency and test–retest reliability.

A strength of this study was the rigorous methodology used to translate and validate a Norwegian version of the Dutch AHPFS in accordance with the COSMIN guidelines.[18] In addition, we had a large sample size and no missing data. However, some limitations need to be addressed. Women with an ACPPS-30 total score of ≥11 according to the 25-item prespecified factor structure in the Dutch version of AHPFS were included in this study to ensure a population of symptomatic women. However, so far, the cutoff scores have not been validated, and whether this is a reasonable clinical cutoff score to identify patients with symptoms associated with high pelvic floor muscle tone is still uncertain. Future studies to develop a clinical cutoff score are therefore needed. Further, this study has not investigated the association between ACPPS-16 and a clinical presentation of pelvic floor muscle tone or the sensitivity of ACPPS-16 to change.

The quantitative survey allowed only an electronic questionnaire response, which may exclude many elderly women from participation. A large group of respondents were recruited among students and employees at the NTNU, which may be a bias in the representation of societal and educational level. Finally, it is worth noting that although a good model fit was achieved in our study, alternative models could have fitted the data equally well, as the model found.[27]

5 | CONCLUSION

The Dutch AHPFS questionnaire was translated into Norwegian through rigorous methodology and presented good content and face validity. However, an assessment of the factor structure concluded with a modified five-factor structure similar to the prespecified factor structure from the Dutch instrument but without the Urinary

tract infection factor and seven other items that were removed. This five-factor model, named ACPPS-16, demonstrated the best fit, covering most of the original scales but with only 16 items. That left 14 of the 30 items unscored, but these may be useful in a clinical context where more information about the patient's condition is needed (ACPPS-30). The translated and modified Norwegian versions, the ACPPS-30 and the shorter version ACPPS-16, are not intended to be used as diagnostic tools, but rather to measure the complexity in patient-reported symptoms. The questionnaires have only been validated for women with symptoms. The questionnaires may be used as both paper-based and digital versions in clinical studies and in clinical practice. They are suitable for women with complex pelvic pain probably due to an abnormally high pelvic floor muscle tone in different home-dwelling settings and age groups. The ACPPS-30 and the ACPPS-16 both proved to be valid, stable and reliable tools and could be useful and adequate instruments for measuring complex pelvic pain symptoms in research and in clinical practice.

AUTHOR CONTRIBUTION

Conception of the study: SS, SNS and CH. Design of the study: SS, SNS and CH. Data acquisition: SS, SNS, VD, IB, SiSp, and CH. Data curation and analyses: SS, VD, IB and CH. Interpretation of results: SS, SNS, VD, IB, and CH. Writing: SS, SNS, VD, IB, and CH. Revising the paper critically and approval of the final version: SS, SNS, VD, IB, SiSp and CH.

ACKNOWLEDGMENTS

We are indebted to the late Professor Ellen T.M. Laan at the University of Amsterdam for her encouragement and discussions during the Norwegian translation and modification of the AHPFS. Since we last were in contact with Ellen, she sadly passed away. Thanks to the clinical nurse Elisabeth Hellem participating in the cognitive interviews and to Anne Caroline Wiik at the Clinical Research Unit, NTNU/St. Olav's Hospital for facilitating the delivery of the electronic questionnaires.

CONFLICT OF INTEREST STATEMENT

The authors have stated explicitly that there are no conflicts of interest in connection with this article.

ORCID

Susan Saga https://orcid.org/0000-0003-0094-1102
Signe Nilssen Stafne https://orcid.org/0000-0003-4922-477X
Cecilie Therese Hagemann https://orcid.org/0000-0001-9565-5824

REFERENCES

1. Thibault-Gagnon S. Definitions and basic etiology of the overactive pelvic floor. In: Padoa A, Rosenbaum TY, eds. *The Overactive Pelvic Floor*. Springer; 2016:1-16.
2. Messelink B, Benson TB, Berghmans B, et al. Standardization of terminology of pelvic floor muscle function and dysfunction: report from the pelvic floor clinical assessment Group of the International Continence Society. *Neurourol Urodynam*. 2005;24:374-380.
3. Bø K. Physiotherapy management of urinary incontinence in females. *J Physiother*. 2020;66:147-154.
4. Frawley H, Shelly B, Morin M, et al. An international continence society (ICS) report on the terminology for pelvic floor muscle assessment. *Neurourol Urodynam*. 2021;40:1217-1260.
5. Bump RC, Norton PA. Epidemiology and natural history of pelvic floor dysfunction. *Obstet Gynecol Clin North Am*. 1998;25:723-746.
6. Kadah S, Soh S, Morin M, Schneider M, Heron E, Frawley H. Is there a difference in pelvic floor muscle tone between women with and without pelvic pain? *A Systematic Review and Meta-Analysis, J Sex Med*. 2023;20:65-96.
7. Nickel JC, Shoskes DA, Wagenlehner FH. Management of chonic prostatitis/chronic pelvic pain syndrome (CP/CPPS): the studies, the evidence, and the impact. *World J Urol*. 2013;31:747-753.
8. Doggweiler R, Whitmore KE, Meijlink JM, et al. A standard for terminology in chronic pelvic pain syndromes: a report from the chronic pelvic pain working group of the international continence society. *Neurourol Urodynam*. 2017;36:984-1008.
9. Rognlid M, Lindsetmo RO. Overaktiv bekkenbunn-syndrom. *Over Active Pelvic Floor Syndrome*] (in Norwegian. No abstract available.) *Tidsskr Nor Laegeforen*. 2010;130:2016-2020.
10. Kennedy VA, Leiserowitz GS. Preserving sexual function in women and girls with cancer: survivorship is about more than just surviving. *Am J Obstet Gynecol*. 2015;213:119-120.
11. Klovning A, Avery K, Sandvik H, Hunskaar S. Comparison of two questionnaires for assessing the severity of urinary incontinence: the ICIQ-UI SF versus the incontinence severity index. *Neurourol Urodynam*. 2009;28:411-415.
12. Teig CJ, Grotle M, Bond MJ, et al. Norwegian translation, and validation, of the pelvic floor distress inventory (PFDI-20) and the pelvic floor impact questionnaire (PFIQ-7). *Int Urogynecol J*. 2017;28:1005-1017.
13. Vaizey CJ, Carapeti E, Cahill JA, Kamm MA. Prospective comparison of faecal incontinence grading systems. *Gut*. 1999;44:77-80.
14. Renzi A, Brillantino A, Di Sarno G, d'Aniello F. Five-item score for obstructed defecation syndrome: study of validation. *Surg Innov*. 2013;20:119-125.
15. Rosen R, Brown C, Heiman J, et al. The female sexual function index (FSFI): a multidimensional self-report instrument for the assessment of female sexual function. *J Sex Marital Ther*. 2000;26:191-208.
16. Rogers RG, Rockwood TH, Constantine ML, et al. A new measure of sexual function in women with pelvic floor disorders (PFD): the pelvic organ prolapse/incontinence sexual questionnaire, IUGA-revised (PISQ-IR). *Int Urogynecol J*. 2013;24:1091-1103.
17. Postma R, Bicanic I, van der Vaart H, Laan E. Pelvic floor muscle problems mediate sexual problems in young adult rape victims. *J Sex Med*. 2013;10:1978-1987.
18. Mokkink LB, Prinsen CAC, Patrick DL, et al. COSMIN study design checklist for patient-reported outcome measurement instruments. July 2019. Accessed March 9, 2023. https://www.cosmin.nl/wp-content/uploads/COSMIN-study-designing-checklist_final.pdf
19. Gagnier JJ, Lai J, Mokkink LB, Terwee CB. COSMIN reporting guideline for studies on measurement properties of patient-reported outcome measures. *Qual Life Res*. 2021;30:2197-2218.
20. deVet HCW, Terwee CB, Mokkink LB, Knol DL. *Measurement in Medicine: A Practical Guide*. Cambridge University Press; 2011.
21. Rivera-Riquelme M, Piqueras JA, Cuijpers P. The revised mental health Inventory-5 (MHI-5) as an ultra-brief screening measure of bidimensional mental health in children and adolescents. *Psychiatry Res*. 2019;274:247-253.
22. Fors EA, Wensaas KA, Eide H, et al. Fibromyalgia 2016 criteria and assessments: comprehensive validation in a Norwegian population. *Scand J Pain*. 2020;20:663-672.

23. StataCorp. *Stata Statistical Software Release 17*. StataCorp LLC; 2021.
24. Mokkink LB, Terwee CB, Patrick DL, et al. 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*. 2010;19:539-549.
25. Jolliffe IT, Cadima J. Principal component analysis: a review and recent developments. *Philos Trans A Math Phys Eng Sci*. 2016;374:20150202.
26. Kline RB. *Principles and Practice of Structural Equation Modeling*. 4th ed. Guildford Press; 2016.
27. Brown TA. *Confirmatory Factor Analysis for Applied Research*. 2th ed. Guilford Press; 2015.
28. Mehmetoglu M, Jakobsen TG. *Applied Statistics Using STATA. A Guide for the Social Sciences*. SAGE Publications Ltd; 2022.
29. Tabachnick BG, Fidell LS. *Using Multivariate Statistics*. 6th ed. Pearson Education; 2013.
30. Raykov T. Estimation of congeneric scale reliability using covariance structure analysis with nonlinear constraints. *Br J Math Stat Psychol*. 2001;54:315-323.
31. Netemeyer RG, Bearden WO, Sharma S. *Scaling Procedures: Issues and Applications*. Sage Publications; 2003.
32. Fayers P, Machin D, Quality of Life. *The Assessment, Analysis and Interpretation of Patient-Reported Outcomes*. 2th ed. John Wiley & Sons Ltd; 2007.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Saga S, Stafne SN, Dons V, Bratlie I, Spetalen S, Hagemann CT. Amsterdam complex pelvic pain symptom scale with subscales: Based on a Norwegian translation, psychometric assessment and modification of the Amsterdam hyperactive pelvic floor scale. *Acta Obstet Gynecol Scand*. 2023;00:1-15. doi:10.1111/aogs.14665