

ORIGINAL ARTICLE



The Abbey Pain Scale: not sufficiently valid or reliable for assessing pain in patients with advanced cancer

Sussi Tegenborg^a , Per Fransson^b and Lisa Martinsson^a

^aDepartment of Radiation Sciences, Oncology, Umeå University, Umeå, Sweden; ^bDepartment of Nursing, Umeå University, Umeå, Sweden

ABSTRACT

Background: Patients with advanced cancer can be unable to verbalize their pain. The Abbey Pain Scale (APS), an observational tool, is used to assess pain in this setting, but has never been psychometrically tested for people with cancer. The aim of this study was to assess the validity, reliability, and the responsiveness of the APS to opioids for patients with advanced cancer in a palliative oncology care setting.

Material and Methods: Patients with advanced cancer and poor performance status, drowsiness, unconsciousness, or delirium, were assessed for pain using a Swedish translation of the APS (APS-SE) and, if possible, the Numeric Rating Scale (NRS). The assessments using APS were conducted simultaneously, but independently, by the same raters on two separate occasions, approximately one hour apart. Criterion validity was assessed by comparing the APS and NRS values using Cohen's kappa (κ). Inter-rater reliability was determined using the intraclass correlation coefficient (ICC), internal consistency using Cronbach's α , and responsiveness to opioids using the Wilcoxon signed-rank test.

Results: Seventy-two patients were included, of whom $n = 45$ could rate their pain using the NRS. The APS did not detect any of the $n = 22$ cases of moderate or severe pain self-reported using the NRS. The APS at first assessment had a κ of 0.08 (CI: -0.06 to 0.22) for criterion validity, an ICC of 0.64 (CI: 0.43 – 0.78) for inter-rater reliability, and a Cronbach's α of 0.01 for internal consistency. The responsiveness to opioids was $z = -2.53$ ($p = 0.01$).

Conclusion: The APS was responsive to opioids but displayed insufficient validity and reliability and did not detect moderate or severe pain as indicated by the NRS. The study showed a very limited clinical use of the APS in patients with advanced cancer.

KEY MESSAGE

The Abbey Pain Scale (APS), an observational pain assessment tool, is used in patients with cancer who cannot verbalize their pain. However, when psychometrically tested, the APS did not display sufficient validity or reliability, so it cannot be recommended for clinical use in patients with advanced cancer.

ARTICLE HISTORY

Received 26 April 2023

Accepted 19 June 2023

KEYWORDS

Pain assessment; cancer; Abbey Pain Scale; validity; reliability; responsiveness

Background

Patients with cancer suffer from many different symptoms, some of the most common symptoms being lack of energy, fatigue, and pain [1–3]. Although the prevalence of pain is high in patients undergoing curative cancer treatment, being found in 40% [4] of them, it is even higher in patients with advanced, metastatic, or terminal cancer, among whom its prevalence is 58–74% [2,5,6]. Pain management is not always successful, and inadequate pain assessment has been described as a major obstacle when identifying physician-related barriers to cancer pain management [7–9]. When assessing pain, self-reporting is considered the gold standard [10,11], and the most widely used self-report pain instruments are the Numerical Rating Scale (NRS), the Visual Analogue Scale (VAS), and the Verbal Rating/Descriptor Scale (VRS/VDS). All these are considered valid and reliable for

clinical use [10], although the NRS is recommended for patients with cancer according to patients' preference and in order to try to standardize which scale to use [11–13]. In the NRS, the patient verbally rates the pain intensity on a scale of 0–10, in which 0 represents 'no pain' and 10 is 'the worst pain imaginable' [11]. Many different wordings of anchor labels exist [11], and efforts are being made to create an international standard for anchor labels in the future [12].

As cancer progresses, it becomes increasingly difficult for some patients to express their pain due to delirium, sedation, or imminent death [14]. When patients can no longer verbally communicate their pain, observational pain assessment by staff members is needed instead. The Abbey Pain Scale (APS) is an observational pain assessment scale recommended by the Australian Pain Society [15] and the British Geriatrics Society [16] for patients with cognitive impairment.

CONTACT Sussi Tegenborg  sussi.tegenborg@umu.se  Department of Radiation Sciences, Oncology, Umeå University, 901 87 Umeå, Sweden

© 2023 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

It is also mentioned in the Swedish National Clinical Practice Guidelines for Palliative Care [17] as a tool for patients with dementia. The APS was developed in Australia in 2004 to assess pain in people with end-stage dementia in residential aged-care homes [18]. It quickly became the most widely used pain assessment tool for this purpose in Australia [19] and has been translated and psychometrically tested in Spain [20], Italy [21], Japan, [22], and Denmark [23].

The APS was created to be an easy-to-use assessment scale and consists of six items: vocalization, facial expression, change in body language, behavioral change, physiological change, and physical change. Each item is illustrated by different examples, such as 'vocalization, e.g., whimpering, groaning, crying', that are rated from absent to severe (0–3 points) and are summed to a total pain score. The staff member completes the instrument by ticking boxes for the total pain score, ranging from 'no pain' (0–2 points) to 'severe pain' (14–18 points), and by classifying the pain as chronic, acute, or acute on chronic. The original study demonstrated a reasonable degree of validity between the APS and nurses' holistic pain assessments when calculated with gamma, 0.59 ($p \leq 0.001$), and a Cronbach's α for internal reliability of 0.74. The numbers for inter-rater reliability were not accounted for but declared as modest [18].

In Sweden, the APS is mainly distributed by the Swedish Register of Palliative Care (SRPC), a national quality register that collects data on end-of-life care [24]. The SRPC distributes the scale to almost 650 Swedish healthcare facilities, such as nursing homes, specialized palliative care units (stand-alone palliative care units, palliative hospital wards, and outpatient homecare), and, to some extent, hospitals. Although the APS was designed for pain assessment in patients with dementia, it is also widely used to assess pain in patients with other diagnoses in Sweden, such as cancer [Maria Andersson, SRPC register manager, personal communication, 25 September 2020].

The APS has been culturally adapted and translated into a Swedish version, the APS-SE [25]. A previous qualitative study of the APS in patients with advanced cancer revealed that it was not optimal in this context [26]. Some parts were not considered useful, and the pain score tended not to reflect the healthcare professionals' intuitive perception of patient suffering [26]. The aim of the present study was accordingly to assess the validity, reliability, and the responsiveness of the APS to opioids for patients with advanced cancer in a palliative oncology care setting.

Material and methods

Recruitment and setting

Recruitment was performed in care units that already had implemented the APS-SE: [25] four oncology wards at a university hospital and a stand-alone specialized palliative inpatient care unit. During weekly contacts, healthcare staff members were asked which patients potentially fulfilled the inclusion criteria. The charts for the suggested patients were then reviewed and the responsible nurses or physicians were consulted regarding the inclusion and exclusion criteria.

Patients 18 or more years old with advanced cancer were considered eligible if they exhibited poor performance [27], delirium [28], aphasia, and/or drowsiness/unconsciousness, [29]. The patient could not be included if diagnosed with dementia, had an intradural catheter, or if less than four hours had passed since the last given as needed dose of opioid. For details, see Table 1.

Patients were included regardless of staff members' pre-understanding of whether or not the patients were suffering from pain at the time of assessment. Patients able to communicate were given oral and written information about the study and signed an informed consent form. If patients were unconscious, suffering from delirium or fatigue, and unable to give informed consent, they were included without consent as permitted by the Swedish Ethical Review Authority. If any family members were present at the time of inclusion, they were informed and asked whether they believed that the patient wanted to be included. Although such proxy decisions do not exist in Swedish legislation, family member assumptions regarding patient opinions on participation were respected.

Data collection

The pain assessments were executed the day of inclusion. The assessments using APS were conducted simultaneously, but independently, by the same raters on two separate occasions, approximately one hour apart. Prior to the first assessment of each patient, a brief review of the APS was made with the staff member involved in the rating. Rater number one, either a nurse, physician, or researcher, varied between the patients, but the same staff member performed both the first and second assessments of each patient. Rater number two was the same researcher, a physician, during all assessments of all patients.

All patients able to communicate were asked whether they consented to physical examination including measuring the blood pressure (BP), as BP outside normal limits is an example of a physiological change given in the APS. If the patient was not communicative, the examination was limited to only checking the body temperature by lightly touching the patient. When doable, the patient was asked about any physical changes such as skin tears or pressure areas, otherwise the nurse in charge was consulted.

After completing the APS in the first and second assessments, all communicative patients were asked to rate their pain with NRS by using the standardized phrase 'On a scale from 0 to 10, where 0 corresponds to no pain and 10 corresponds to worst pain imaginable - can you describe your pain right now?'. If they did not answer in digits even after a repeated request, they were considered incapable of completing the NRS. The scores and information about any patient requests for pain medication were forwarded to the nurse in charge, who then decided whether to give the patient pain medication, and if so, what, as part of routine healthcare. If the patient was given any medication between the two assessments, the kind of drug, dose, and route of administration were recorded.

Table 1. Additional inclusion criteria besides advanced cancer and age ≥ 18 years and exclusion criteria.

Inclusion criteria^a	
Poor performance	Performance status 4 according to World Health Organization (WHO) or Eastern Cooperative Oncology Group (ECOG) [27]
Delirium	Patient totally confined to bed or chair. Using the Single Question in Delirium (SQiD) instrument [28] 'Do you feel that [patient's name] has been more confused lately?'
Drowsiness or unconsciousness	Reaction Level Scale (RLS85): [29] ≥ 2 RLS 2; drowsy, responsive to light stimulation verbally or by touch. RLS 3; very drowsy, responsive to strong stimulation with repeated loud verbal approach or shaking. RLS 4-8; unconscious. No pain stimulation was executed to grade the level of drowsiness or unconsciousness. If considered unconscious, i.e., not responding to strong verbal stimulation or shaking, no additional classification has been made to separate level 4 to 8.
Exclusion criteria^b	
Aphasia	If previously diagnosed according to chart
Dementia	If previously diagnosed according to chart
Intradural catheter for pain	Regardless of administrated drug
Less than four hours since last given as needed opioid before first assessment	The first assessment could be made as soon as more than four hours passed since the last given as need opioid had been administrated.

^aAt least one inclusion criterion had to be present for patients to be eligible for inclusion.

^bPatients meeting either of these criteria were excluded.

Time of death (if the patient was deceased by the end of data collection) was recorded from the patients' charts. Data were collected between February 2022 and March 2023.

Analysis

The statistical analysis was performed using SPSS 28.0.1.1 software.

Validity testing

Criterion validity, i.e., the association between the APS score and the gold standard, i.e., self-reported NRS, was determined using Cohen's kappa (κ). Criterion validity was only calculated for the first pain assessment. The mean APS score between raters one and two at the first assessment was calculated for each patient, and then translated to pain categories: no, mild, moderate, or severe pain. The APS score followed the original study's categorization: no pain = 0–2 points, mild pain = 3–7, moderate pain = 8–13, and severe pain = 14–18 [18]. The self-rated NRS score was translated to: no pain = 0 points, mild pain = 1–4, moderate pain = 5–6, and severe pain = 7–10 [30]. Criterion validity was also calculated for no pain versus pain as assessed with the APS (0–2 vs. ≥ 3 points) compared with the NRS (0 vs. ≥ 1 point).

Reliability testing

Inter-rater reliability, i.e., whether two independent observers agreed in their APS ratings, was determined using the intra-class correlation coefficient (ICC) based on a single-rating, absolute-agreement, two-way, random-effects model [31]. The ICC was calculated at both the first and second assessments separately at $n = 50$ and at $n = 72$ patients.

Internal consistency reliability, i.e., how the individual items in the APS relate to one another and the APS score as a whole at the first assessment, was determined using Cronbach's α [32]. A separate α value was calculated for the subgroup of patients who consented to BP assessments.

Responsiveness to opioids

Responsiveness, i.e., the ability to detect a change in the mean APS score and NRS score (if possible) between the first and second assessments, was determined using the Wilcoxon signed-rank test. The responsiveness was calculated separately for the group receiving opioids and, as a control, the group not receiving opioids.

Sample size

When planning this study, a statistician active in the field of cancer research was consulted. The statistician concluded that there was no certain way to calculate power due to the uncertainty of the how the different APS scores would range within the target population. In previous studies validating APS, 50–171 patients have been included [18,20–23], so we aimed for at least 50 patients with the possibility of including up to 100. When analyzing the data from 50 patients, the inter-rater reliability test resulted in an ICC with a wide confidence interval, even though it has been suggested that 30–50 is a reasonable number of patients for estimating reliability [31,33]. We accordingly decided to include additional patients to a total of 72, but this resulted in only a slightly narrower confidence compared to the confidence interval at 50 patients.

Results

The participants

A total of 72 patients were included in the study (see Table 2).

The patients were bedridden during $n = 134$ (93%) of the assessments and sitting in a chair during the rest, $n = 10$ (7%). The median time between the assessments and death was 11 days (range 0.5–201). At the time of analysis, $n = 70$ (97%) of the patients were deceased.

Table 2. Patient characteristics, $n = 72$.

Demographics		
Female	$n = 41$	57%
Male	$n = 31$	43%
Mean age (range), years	76	(47–95)
Type of cancer		
Gastrointestinal	$n = 23$	32%
Gynaecological	$n = 9$	13%
Urologic	$n = 9$	13%
Breast	$n = 6$	8%
Lung	$n = 6$	8%
Central nervous system	$n = 6$	8%
Head and neck	$n = 5$	7%
Hematological, including lymphoma	$n = 5$	7%
Cancer of unknown primary (CUP)	$n = 3$	4%
Reason(s) for inclusion		
Poor performance status	$n = 36$	50%
Poor performance status and delirium	$n = 15$	21%
Delirium	$n = 6$	8%
Drowsiness/unconsciousness	$n = 14$	19%
Aphasia	$n = 1$	1%
Ability to complete NRS		
Yes, at both assessments	$n = 41$	57%
Yes, at one assessment	$n = 4$	6%
No	$n = 27$	37%
Received opioids between assessments		
Yes	$n = 26$	36%
No	$n = 46$	64%
Measured blood pressure		
Yes, with consent from patient	$n = 56$	78%
No, patient declined	$n = 0$	0%
No, patient incapable of giving consent	$n = 16$	22%

Most participants, $n = 51$ (71%), received ongoing regular opioid medication with a median equivalent dose of 80 mg of morphine p.o. per day (10–935 mg p.o.). The rest, $n = 21$ (29%), had a prescription for opioids only as needed, with or without ongoing regular paracetamol and/or non-steroidal anti-inflammatory drugs (NSAIDs). Drugs for neuropathic pain, gabapentin or pregabalin, were given regularly to $n = 11$ (15%) of the patients, to complement the opioids. In $n = 26$ (36%) of the cases, the patient was given opioids as needed between assessments 1 and 2, with a median equivalent dose of 12.5 mg morphine p.o. (10–108 mg p.o.).

The staff members involved in the assessment were nurses ($n = 32$), physicians ($n = 6$), and researchers ($n = 2$).

Validity

The mean APS score for all items at the first assessment was 2.35, with each item contributing as follows: vocalization, 0.28; facial expression, 0.66; change in body language, 0.13; behavioral change, 0.42; physiological change, 0.38; and physical change, 0.48.

The criterion validity, i.e., the association between pain categories assessed using the APS versus the self-reported NRS ($n = 45$) at the first assessment was $\kappa = 0.08$ (95% CI: -0.06 to 0.22) (for the cross-tabulation table, see Table 3). The corresponding κ value for the subgroup that answered the NRS question and did not suffer from delirium ($n = 38$) was 0.15 (95% CI: 0.01 to 0.31). The association between the APS and NRS ratings regarding pain versus no pain for all patients was $\kappa = 0.04$ (95% CI: -0.19 to 0.26). Regardless of removing the patients with delirium or not, the κ values were ≤ 0.2 , which is considered slight [34]. None of the patients had more than mild pain according to the APS

scores, although when using the NRS 13 patients self-reported moderate and nine severe pain at the first assessment.

Reliability

Inter-rater reliability, there was agreement between the APS ratings of the two observers, with an ICC at first assessment of 0.65 (95% CI: 0.38 – 0.80) at 50 patients and of 0.64 (95% CI: 0.43 – 0.78) at 72 patients. At the second assessment the ICC was 0.64 (95% CI: 0.36 – 0.79) at 50 patients and 0.61 (95% CI: 0.38 – 0.76) at 72 patients. An ICC of 0.5 – 0.75 is considered moderate [31].

Internal consistency reliability, i.e., how the individual items in the APS relate to one another and to the APS score as a whole at first assessment, was indicated by a Cronbach's α of 0.01. When each item was separately deleted, the Cronbach's α ranged from -0.22 to 0.28 (see Table 4).

For the group of completed BP assessments ($n = 56$), the Cronbach's α was 0.05, and when each item was separately deleted, the Cronbach's α ranged from -0.19 to 0.32 (see Table 4). An α value < 0.5 is considered unacceptable [32].

Responsiveness to opioids

For the group ($n = 26$) receiving opioids between the first and second assessments, there were statistically significant decreases in both the APS ($p = 0.011$) and NRS ($p = 0.003$) scores. For the group not receiving opioids ($n = 46$), there was no statistically significant change for either of the instruments (see Table 5). A total of $n = 45$ patients self-reported pain with the NRS. Almost all, $n = 41$, were able to use the NRS in both assessments and thus responsiveness could be calculated. However, in $n = 4$ patients, only one assessment with the NRS was doable, making it impossible to calculate responsiveness for these patients.

Discussion

The APS is not sufficiently valid or reliable when used for patients with advanced cancer suffering from poor performance status, delirium, drowsiness, or unconsciousness. The APS displayed responsiveness to opioids and can therefore be useful to evaluate the effect of medication administered for pain, at least when assessed by the same staff member before and after administering medication. However, the APS did not detect moderate or severe pain as indicated by the NRS and the correlation with the patients' own ratings was only slight. If used as a screening instrument, there is an imminent risk of the assessment indicating no pain or a lower intensity of pain than the patient is actually suffering. This problem was previously addressed in a qualitative study of healthcare staff members' experience of using the APS in patients with advanced cancer [26]. To conclude, this study showed a very limited clinical use of the APS in patients with advanced cancer.

To maintain and improve healthcare quality and research, appropriate instruments and tools are needed. An instrument is usually tested for validity when originally developed, but

Table 3. Pain categories for NRS and APS scores at first assessment for those who could rate their pain using NRS, $n = 45$.

		NRS pain category ^a (n)				Total
		No pain	Mild pain	Moderate pain	Severe pain	
APS pain category ^b (n)	No pain	6	5	8	6	25
	Mild pain	4	8	5	3	20
	Moderate pain	0	0	0	0	0
	Severe pain	0	0	0	0	0
	Total	10	13	13	9	45

^aNRS score: no pain = 0, mild pain = 1–4, moderate pain = 5–6, and severe pain = 7–10.

^bAPS score: no pain = 0–2, mild pain = 3–7, moderate pain = 8–13, and severe pain = 14–18.

Table 4. Cronbach's α if each item is deleted separately.

	Total group ($n = 72$)	Subgroup with completed blood pressure assessments ($n = 56$)
Vocalization	–0.16	–0.14
Facial expression	–0.22	–0.19
Change in body language	–0.14	–0.01
Behavioral change	0.11	0.22
Physiological change	0.07	–0.04
Physical change	0.28	0.32

Table 5. APS and NRS Responsiveness to opioids.

		Score at first assessment	Score at second assessment	Score difference	z-value ^a	p-value ^b
APS score $n = 72$	Opioid, $n = 26$	2.96	2.25	0.71	–2.53	0.011
	No opioid, $n = 46$	2.00	1.92	0.08	–0.46	0.646
NRS score $n = 41$	Opioid, $n = 15$	5.00	3.33	1.67	–3.02	0.003
	No opioid, $n = 26$	2.93	2.54	0.39	–0.18	0.860

^aWilcoxon signed-rank test. ^b $p \leq 0.05$ is considered statistically significant.

this is unfortunately not always the case when instruments are adopted in different countries and/or settings. To ensure validity, guidelines exist to ensure that instruments are not only well translated linguistically but also culturally adapted to maintain their content validity across different settings and cultures before being psychometrically tested [35,36]. The APS has previously been scientifically translated and culturally adapted for patients with dementia in Sweden [25], but is being used in patients with cancer without any previous psychometric testing in this different population. In fact, although the APS is widely used to assess pain in patients with advanced cancer, no literature was found regarding its use in cancer populations worldwide apart from a single case report. In this case report the APS was considered successfully for pain assessment for a patient with terminal cancer with understanding difficulties [37]. To our knowledge, this is the first study psychometrically testing the APS in patients with advanced cancer.

When testing the criterion validity, i.e., the APS scores compared with the NRS scores as the gold standard, it was only considered slight [34]. This slight correlation prevailed when comparing the ability to detect any pain between the APS and NRS. Since the applied gold standard varies between studies, the criterion validities are difficult to compare; nevertheless, our study shows less correlation than do studies of patients with dementia [18,20–23]. When APS was compared with nurses' overall holistic pain assessment, the validity of the APS was considered reasonable [18] and highly significant compared with observer estimation of pain

[20]. The validity was moderate compared with that of the self-reporting VRS/VDS [22,23].

Previous reports indicate that patients with cancer rate their pain higher than do their caregivers [11,38–40], but there is nonetheless a striking difference between the self-reported pain and the observed pain intensity rated using the APS: although 49% of the able patients self-reported moderate or severe pain, indicating a need for pain relief, there was *not a single such rating* in these categories when using the APS. This indicates that the scale is too insensitive.

The inter-rater reliability was considered moderate: the ICC was 0.64 at first assessment but with a confidence interval ranging from poor to good [31]. The ICC for the inter-rater reliability in the original study was never disclosed but was declared modest at best [18]. Other studies reported ICC values that indicated good reliability, 0.82 in the Japanese study [22] and 0.84 in the Danish study [23]. In both these studies the ICC was calculated between only two raters. When exploring the feasibility and clinical utility of the Japanese version of the APS, the ICC was tested for several different raters ($n = 23$), so the inter-rater reliability decreased to 0.68–0.76 [41].

The Cronbach's α for internal consistency in this study was unacceptable [32] at $\alpha = 0.01$. In contrast, the Cronbach's α varied between 0.52 and 0.71 [20,22,23] when calculated for patients with dementia. When looking separately at each item, there was a substantial correlation among all the items except for 'physiological change' in the Japanese study [22], whereas the Danish study [23] showed fairly good correlation

between 'vocalization', 'facial expression', and 'change in body language', but no or poor correlation for the rest of the items. When tested in patients with advanced cancer, there was no substantial correlation for any of the items.

The APS displayed responsiveness to the administration of pain medication to patients, as shown in the original study as well as in the Spanish and Danish study [18,20,23]. As a measure of quality, such responsiveness was also calculated for the NRS, and the result was coherent with the decrease in post-intervention scores found in other studies [42,43]. A two-point change in the NRS has been suggested to be a clinically meaningful change in pain [44,45], but no data have been reported for a clinically meaningful change in APS scores.

Overall, the psychometric properties of the APS are better in patients suffering from dementia than in patients with advanced cancer. This seems reasonable since the instrument was developed for patients with dementia and, accordingly, some of the suggested examples in the APS are not always appropriate for patients with cancer. Patients with dementia may also still be in relatively good physical health, compared with patients with advanced cancer, when an observational pain assessment scale is needed. Patients with dementia are more likely to be bodily active, for example to display withdrawal or rocking back- and forth while sitting in a chair, even if the patients' impaired cognitive abilities make it difficult to express their pain verbally. Patients with cancer are often confined to their bed in end-of-life, and do not present changes in body language as rocking or fidgeting. In this study, 62% of the patients self-reported pain, although the median time between the assessments and death was only 11 days. The difference in overall health between the populations, and the fact that cancer per se might yield invalid cancer-generated measurements [26], could explain why the APS is more appropriate for people with dementia than those with advanced cancer.

According to the World Health Organization (WHO), cancer is ranked 6th and dementia 7th among the leading causes of death in the world, and thus affect millions of people worldwide [46]. The prevalence of pain is high and about the same in both populations, being experienced by 35–79% of patients with dementia [47–49] and 40–74% of patients with cancer [2,4–6]. In both cases, inadequate pain assessment is considered a problem in managing pain effectively [8,50,51]. While this problem has been embraced in the clinical setting of dementia, leading to the development of at least eight pain assessment tools with strong or moderate psychometric evidence [51], no comparable observational pain assessment tool has been developed especially for patients with advanced cancer. It is time to develop an observational pain assessment tool for the clinical oncology setting.

Strengths and limitations

To our knowledge, this is the first study examining the validity and reliability of an observational pain assessment tool for patients with advanced cancer.

The inclusion criteria 'delirium' was determined using the SQiD instrument, which was previously found feasible for patients hospitalized with cancer [28]. The SQiD was originally designed to be answered by a relative or friend, but in this study it was answered by staff members in charge of the patients. It is possible that this shift caused us to under- or overestimate the number of patients with delirium compared to relatives. Friends and relatives presumably have a greater familiarity with the patient and thus more accurately can determine if the patient is suffering from delirium or not.

Participating in this study probably generated some staff discussion of how to rate different items, likely making the inter-rater reliability higher than it would have been without the ongoing study. With fewer raters involved, the ICC would probably have been higher.

Conclusion

The APS displayed responsiveness to opioids but did not display sufficient validity or reliability and did not detect moderate or severe pain as indicated by the NRS in patients with advanced cancer. The study showed a very limited clinical use of the APS in this population. Further research is needed to develop a valid and reliable observational pain assessment scale for patients with advanced cancer.

Ethical approval

This study was approved by the Regional Ethical Review Board in Umeå, Sweden [registration number 2017/504-31] and an amendment was approved by the Swedish Ethical Review Authority, Sweden [registration number 2021-06206-02].

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by: the Department of Oncology, Norrlands University Hospital, Umeå, Sweden [grant number 52903/7002228]; Cancerforskningsfonden i Norrland, Umeå, Sweden [grant number AMP 18-922]; and Västerbottens läns landstings ettåriga forskningsanslag, Umeå, Sweden [grant number 7003533]. The funding sources were not involved in the study.

ORCID

Sussi Tegenborg  <http://orcid.org/0000-0002-0125-629X>

Data availability statement

The data that support the findings of this study are available from the corresponding author, [ST], upon reasonable request.

References

- [1] Van Lancker A, Beeckman D, Van Den Noortgate N, et al. Frequency and intensity of symptoms and treatment

- interventions in hospitalized older palliative cancer patients: a multicentre cross-sectional study. *J Adv Nurs*. 2017;73(6):1455–1466. doi: [10.1111/jan.13230](https://doi.org/10.1111/jan.13230).
- [2] Van Lancker A, Velghe A, Van Hecke A, et al. Prevalence of symptoms in older cancer patients receiving palliative care: a systematic review and meta-analysis. *J Pain Symptom Manage*. 2014; 47(1):90–104. doi: [10.1016/j.jpainsymman.2013.02.016](https://doi.org/10.1016/j.jpainsymman.2013.02.016).
- [3] Potter J, Hami F, Bryan T, et al. Symptoms in 400 patients referred to palliative care services: prevalence and patterns. *Palliat Med*. 2003;17(4):310–314. doi: [10.1191/0269216303pm7600a](https://doi.org/10.1191/0269216303pm7600a).
- [4] Evenepoel M, Haenen V, De Baerdemaeker T, et al. Pain prevalence during cancer treatment: a systematic review and Meta-Analysis. *J Pain Symptom Manage*. 2022;63(3):e317–e335. doi: [10.1016/j.jpainsymman.2021.09.011](https://doi.org/10.1016/j.jpainsymman.2021.09.011).
- [5] van den Beuken-van Everdingen MH, Hochstenbach LM, Joosten EA, et al. Update on prevalence of pain in patients with cancer: systematic review and Meta-Analysis. *J Pain Symptom Manage*. 2016;51(6):1070–1090.e9. doi: [10.1016/j.jpainsymman.2015.12.340](https://doi.org/10.1016/j.jpainsymman.2015.12.340).
- [6] Fallon M, Giusti R, Aielli F, et al. Management of cancer pain in adult patients: ESMO clinical practice guidelines. *Ann Oncol*. 2018;29(Suppl 4):iv166–iv191. doi: [10.1093/annonc/mdy152](https://doi.org/10.1093/annonc/mdy152).
- [7] Von Roenn JH, Cleeland CS, Gonin R, et al. Physician attitudes and practice in cancer pain management. A survey from the Eastern cooperative oncology group. *Ann Intern Med*. 1993; 119(2):121–126. doi: [10.7326/0003-4819-119-2-199307150-00005](https://doi.org/10.7326/0003-4819-119-2-199307150-00005).
- [8] Stewart J. The challenges of cancer pain assessment. *Ulster Med J*. 2014;83(1):44–46.
- [9] Jacobsen R, Sjøgren P, Møldrup C, et al. Physician-related barriers to cancer pain management with opioid analgesics: a systematic review. *J Opioid Manag*. 2007;3(4):207–214.
- [10] Karcioğlu O, Topacoglu H, Dikme O, et al. A systematic review of the pain scales in adults: which to use? *Am J Emerg Med*. 2018; 36(4):707–714. doi: [10.1016/j.ajem.2018.01.008](https://doi.org/10.1016/j.ajem.2018.01.008).
- [11] Hjermstad MJ, Fayers PM, Haugen DF, et al. Studies comparing numerical rating scales, verbal rating scales, and visual analogue scales for assessment of pain intensity in adults: a systematic literature review. *J Pain Symptom Manage*. 2011;41(6):1073–1093. doi: [10.1016/j.jpainsymman.2010.08.016](https://doi.org/10.1016/j.jpainsymman.2010.08.016).
- [12] Kaasa S, Apolone G, Klepstad P, et al. Expert conference on cancer pain assessment and classification—the need for international consensus: working proposals on international standards. *BMJ Support Palliat Care*. 2011;1(3):281–287. doi: [10.1136/bmjspcare-2011-000078](https://doi.org/10.1136/bmjspcare-2011-000078).
- [13] Firdous S, Mehta Z, Fernandez C, et al. A comparison of numeric pain rating scale (NPRS) and the visual analog scale (VAS) in patients with chronic cancer-associated pain. *J Clin Oncol*. 2017; 35(31_suppl):217–217. doi: [10.1200/JCO.2017.35.31_suppl.217](https://doi.org/10.1200/JCO.2017.35.31_suppl.217).
- [14] Tapp D, Chenacher S, Gérard NPA, et al. Observational pain assessment instruments for use with nonverbal patients at the end-of-life: a systematic review. *J Palliat Care*. 2019;34(4):255–266. doi: [10.1177/0825859718816073](https://doi.org/10.1177/0825859718816073).
- [15] The Australian Pain Society. Pain in Residential Aged Care Facilities Management Strategies Sydney 2005. 2019. https://www.apsoc.org.au/PDF/Publications/Pain_in_Residential_Aged_Care_Facilities_Management_Strategies.pdf
- [16] British Geriatrics Society. Silver Book II Quality urgent care for older people. 2021. <https://www.bgs.org.uk/resources/resource-series/silver-book-ii>
- [17] Regionala cancercentrum i samverkan. Nationellt vårdprogram palliativ vård. [National Clinical Practice Guideline for Palliative Care]. 2021. <https://kunskapsbanken.cancercentrum.se/diagnoser/palliativ-varld/>.
- [18] Abbey J, Piller N, De Bellis A, et al. The abbey pain scale: a 1-minute numerical indicator for people with end-stage dementia. *Int J Palliat Nurs*. 2004;10(1):6–13. doi: [10.12968/ijpn.2004.10.1.12013](https://doi.org/10.12968/ijpn.2004.10.1.12013).
- [19] Abbey J. Putting pain scales to the test. *Aust Nurs J*. 2007;14(7): 43.
- [20] Chamorro P, Puche E. Traducción al castellano y validación de la escala abbey Para la detección del dolor en pacientes no comunicativo. [Translation into Spanish and validation of the Abbey scale for the detection of pain in non-communicative patients]. *Rev Soc Esp Dolor*. 2013;20(1):3–7. doi: [10.4321/S1134-80462013000100002](https://doi.org/10.4321/S1134-80462013000100002).
- [21] Storti M. The validation of a pain assessment scale for patients with cognitive impairment: the italian version of abbey's scale. *Recenti Prog Med*. 2009;100(9):405–409.
- [22] Takai Y, Yamamoto-Mitani N, Chiba Y, et al. Abbey pain scale: development and validation of the japanese version. *Geriatr Gerontol Int*. 2010;10(2):145–153. doi: [10.1111/j.1447-0594.2009.00568.x](https://doi.org/10.1111/j.1447-0594.2009.00568.x).
- [23] Gregersen M, Melin AS, Nygaard IS, et al. Reliability of the danish abbey pain scale in severely demented and non-communicative older patients. *Int J Palliat Nurs*. 2016;22(10):482–488. doi: [10.12968/ijpn.2016.22.10.482](https://doi.org/10.12968/ijpn.2016.22.10.482).
- [24] Lundström S, Axelsson B, Heedman PA, et al. Developing a national quality register in end-of-life care: the swedish experience. *Palliat Med*. 2012;26(4):313–321. doi: [10.1177/0269216311414758](https://doi.org/10.1177/0269216311414758).
- [25] Tegenborg S, Fransson P, Martinsson L. Translation, cultural adaptation and recommendations for clinical implementation of the abbey pain scale to a Swedish dementia care context. *Nurs Open*. 2022;10(3):1367–1374.
- [26] Tegenborg S, Fransson P, Martinsson L. Physicians' and nurses' experience of using the abbey pain scale (APS) in people with advanced cancer: a qualitative content analysis. *BMC Nurs*. 2023; 22(1):95. doi: [10.1186/s12912-023-01227-7](https://doi.org/10.1186/s12912-023-01227-7).
- [27] International Association for Hospice and Palliative Care. ECOG score. 2022. <https://pallipedia.org/ecog-score/>
- [28] Sands MB, Sharma S, Carpenter L, et al. SQiD, the single question in delirium; a single question help clinicians to detect delirium in hospitalised cancer patients?" running heading single question in delirium" (Bcan-D-20-01665). *BMC Cancer*. 2021;21(1):75. doi: [10.1186/s12885-020-07504-x](https://doi.org/10.1186/s12885-020-07504-x).
- [29] Starmark JE, Stålhammar D, Holmgren E. The reaction level scale (RLS85). manual and guidelines. *Acta Neurochir*. 1988;91(1-2):12–20. doi: [10.1007/BF01400521](https://doi.org/10.1007/BF01400521).
- [30] Oldenmenger WH, de Raaf PJ, de Klerk C, et al. Cut points on 0–10 numeric rating scales for symptoms included in the edmonton symptom assessment scale in cancer patients: a systematic review. *J Pain Symptom Manage*. 2013;45(6):1083–1093. doi: [10.1016/j.jpainsymman.2012.06.007](https://doi.org/10.1016/j.jpainsymman.2012.06.007).
- [31] Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J Chiropr Med*. 2016;15(2):155–163. doi: [10.1016/j.jcm.2016.02.012](https://doi.org/10.1016/j.jcm.2016.02.012).
- [32] Gliem JG. Calculating, interpreting, and reporting cronbach's alpha reliability coefficient for likert-type scales. 2003.
- [33] Hopkins WG. Measures of reliability in sports medicine and science. *Sports Med*. 2000;30(1):1–15. doi: [10.2165/00007256-200030010-00001](https://doi.org/10.2165/00007256-200030010-00001).
- [34] Kassambara A. Kappa Coefficient Interpretation 2019. Available from: <https://www.datanova.com/en/blog/kappa-coefficient-interpretation/>.
- [35] Beaton DE, Bombardier C, Guillemin F, et al. Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine (Phila Pa 1976)*. 2000;25(24):3186–3191. doi: [10.1097/00007632-200012150-00014](https://doi.org/10.1097/00007632-200012150-00014).
- [36] Antunes B, Daveson B, Ramsenthaler C, et al. The palliative care outcome scale (POS) manual for cross-cultural adaptation and psychometric validation. London: Cicely Saunders Institute. 2012.
- [37] Okimasa S, Saito Y, Okuda H, et al. Assessment of cancer pain in a patient with communication difficulties: a case report. *J Med Case Rep*. 2016;10(1):148.
- [38] Grossman SA, Sheidler VR, Swedeen K, et al. Correlation of patient and caregiver ratings of cancer pain. *J Pain Symptom Manage*. 1991;6(2):53–57. doi: [10.1016/0885-3924\(91\)90518-9](https://doi.org/10.1016/0885-3924(91)90518-9).
- [39] Yun YH, Heo DS, Lee IG, et al. Multicenter study of pain and its management in patients with advanced cancer in korea. *J Pain*

- Symptom Manage. 2003;25(5):430–437. doi: [10.1016/s0885-3924\(03\)00103-9](https://doi.org/10.1016/s0885-3924(03)00103-9).
- [40] Hovi SL, Lauri S. Patients' and nurses' assessment of cancer pain. *Eur J Cancer Care (Engl)*. 1999;8(4):213–219. doi: [10.1046/j.1365-2354.1999.00171.x](https://doi.org/10.1046/j.1365-2354.1999.00171.x).
- [41] Takai Y, Yamamoto-Mitani N, Chiba Y, et al. Feasibility and clinical utility of the Japanese version of the Abbey pain scale in Japanese aged care. *Pain Manag Nurs*. 2014;15(2):439–448. doi: [10.1016/j.pmn.2012.02.003](https://doi.org/10.1016/j.pmn.2012.02.003).
- [42] Chien CW, Bagraith KS, Khan A, et al. Comparative responsiveness of verbal and numerical rating scales to measure pain intensity in patients with chronic pain. *J Pain*. 2013;14(12):1653–1662. doi: [10.1016/j.jpain.2013.08.006](https://doi.org/10.1016/j.jpain.2013.08.006).
- [43] Castarlenas E, Jensen MP, von Baeyer CL, et al. Psychometric properties of the numerical rating scale to assess Self-Reported pain intensity in children and adolescents: a systematic review. *Clin J Pain*. 2017;33(4):376–383. doi: [10.1097/AJP.0000000000000406](https://doi.org/10.1097/AJP.0000000000000406).
- [44] Farrar JT, Portenoy RK, Berlin JA, et al. Defining the clinically important difference in pain outcome measures. *Pain*. 2000;88(3):287–294. doi: [10.1016/S0304-3959\(00\)00339-0](https://doi.org/10.1016/S0304-3959(00)00339-0).
- [45] Childs JD, Piva SR, Fritz JM. Responsiveness of the numeric pain rating scale in patients with low back pain. *Spine (Phila Pa 1976)*. 2005;30(11):1331–1334. doi: [10.1097/01.brs.0000164099.92112.29](https://doi.org/10.1097/01.brs.0000164099.92112.29).
- [46] World Health Organization. The top 10 causes of death. World Health Organization, Geneva. 2020. <https://www.who.int/news-room/fact-sheets/detail/the-top-10-causes-of-death>
- [47] Atee M, Morris T, Macfarlane S, et al. Pain in dementia: prevalence and association with neuropsychiatric behaviors. *J Pain Symptom Manage*. 2021;61(6):1215–1226. doi: [10.1016/j.jpainsymman.2020.10.011](https://doi.org/10.1016/j.jpainsymman.2020.10.011).
- [48] Rajkumar AP, Ballard C, Fossey J, et al. Epidemiology of pain in people with dementia living in care homes: longitudinal course, prevalence, and treatment implications. *J Am Med Dir Assoc*. 2017;18(5):453.e1–453–e6. doi: [10.1016/j.jamda.2017.01.024](https://doi.org/10.1016/j.jamda.2017.01.024).
- [49] Feast AR, White N, Lord K, et al. Pain and delirium in people with dementia in the acute general hospital setting. *Age Ageing*. 2018;47(6):841–846. doi: [10.1093/ageing/afy112](https://doi.org/10.1093/ageing/afy112).
- [50] Fink RM, Gallagher E. Cancer pain assessment and measurement. *Semin Oncol Nurs*. 2019;35(3):229–234. doi: [10.1016/j.soncn.2019.04.003](https://doi.org/10.1016/j.soncn.2019.04.003).
- [51] Smith TO, Harvey K. Psychometric properties of pain measurements for people living with dementia: a COSMIN systematic review. *Eur Geriatr Med*. 2022;13(5):1029–1045. doi: [10.1007/s41999-022-00655-z](https://doi.org/10.1007/s41999-022-00655-z).