UMEÅ UNIVERSITY

PAIN OR NO PAIN? THAT IS THE QUESTION

An evaluation of the observational pain assessment instrument Abbey Pain Scale in patients with cancer

Sussi Tegenborg

Department of Diagnostics and Intervention, Oncology.

Umeå 2024

Responsible publisher under Swedish law:
the Dean of the Medical Faculty

Dissertation for PhD
ISBN: 978-91-8070-348-2 (print)
ISBN: 978-91-8070-349-9 (pdf)
ISSN: 0346-6612
New Series Number 2296
Information about cover design: by Sussi Tegenborg created in Leonardo.Ai
Electronic version available at: http://umu.diva-portal.org/
Printed by: Cityprint i Norr AB
Umeå, Sweden 2024

In loving memory of my mother, my best friend who taught me everything important in life. I still miss you every day…

Table of Contents

Abstract

Background: The Abbey Pain Scale (APS) is an observational pain assessment instrument that was created for use among patients with dementia. It is sometimes used in Sweden to assess pain in patients with advanced cancer who are unable to vocalize their pain, but it has never been evaluated in this population.

Aim: To evaluate observational pain assessment for people with advanced cancer in a palliative care setting, focusing on the APS.

Methods: The APS was translated and adapted into a Swedish dementia context through interviews with health care professionals (n=11). The resulting APS-SE was then used in a qualitative content analysis exploring health care professionals' experience of using the instrument in patients with advanced cancer (n=12). The psychometrics of the APS-SE when used in patients with advanced cancer (n=72) were determined through test-retest and inter-rater reliability, internal consistency reliability, and responsiveness to opioids. Criterion validity was determined against the patients' self-reported pain (n=45).

Results: The APS-SE was comprehensible by users within dementia care, regardless of their educational and linguistic background. The qualitative analysis revealed that when used in patients with advanced cancer, the APS-SE fulfilled the need for an observational pain assessment instrument, but was not always on target and did not fully suit the clinical situation. The psychometric analysis showed slight criterion validity (κ=0.08) and unacceptable internal consistency reliability (Cronbach's α=0.01). The test-retest reliability was good (ICC=0.82) and the inter-rater reliability moderate (ICC=0.64), but the latter had a confidence interval ranging from poor to good. Responsiveness to opioids was demonstrated (p=0.01).

Conclusions: This thesis underscores the need for a specialized observational pain assessment instrument explicitly tailored for patients with advanced cancer. The current lack of a recognized alternative emphasizes the importance of developing such an instrument to address the critical gap in observational pain assessment in the palliative oncology setting.

Abbreviations

α	Cronbach's alpha
APS	Abbey Pain Scale
APS-SE	A Swedish version of the Abbey Pain Scale
AI	Artificial intelligence
ASCO	American Society of Clinical Oncology
BPI	Brief Pain Inventory
CNPI	Checklist of Nonverbal Pain Indicators
COSMIN	COnsensus-based Standards for the selection of health Measurement INstruments
ECOG	Eastern Cooperative Oncology Group
ECS-CP	Edmonton Classification System for Cancer Pain
EAPC	European Association for Palliative Care
ESAS-r	Edmonton Symptom Assessment System-revised or Edmonton Symptom Assessment Scale-revised
ESMO	European Society for Medical Oncology
FACS	Facial Action Coding System
HCP	Health care professional
IASP	International Association for the Study of Pain
IAHPC	International Association for Hospice and Palliative Care
ICC	Intraclass correlation coefficient
ICD-11	International Classification of Diseases-11
IPOS	Integrated Palliative Care Outcome Scale
κ	Cohen's kappa
MPQ	McGill Pain Questionnaire
MDASI	MD Anderson Symptom Inventory
MOBID	Mobilization-Observation-Behaviour-Intensity-Dementia pain scale
MOBID-2	Mobilization-Observation-Behaviour-Intensity-Dementia pain scale-2
NBHW	National Board of Health and Welfare
NSAIDs	Non-Steroidal Anti-Inflammatory Drugs
NRS	Numerical rating scale
PACSLAC	Pain Assessment Checklist for Seniors with Limited Ability to Communicate
PACSLAC-II	Pain Assessment Checklist for Seniors with Limited Ability to Communicate-II
PAINAD	Pain Assessment in Advanced Dementia

RLS	Reaction Level Scale
SF-MPQ-2	Short form McGill Pain Questionnaire-2
SRPC	Swedish Register of Palliative Care
SQiD	Single Question in Delirium
VAS	Visual analogue scale
VDS	Verbal descriptor scale
VRS	Verbal rating scale
WHO	World Health Organization
WHYMPI	West Haven-Yale Multidimensional Pain Inventory

Sammanfattning på svenska

Globalt lider uppskattningsvis omkring två miljoner människor varje år av underbehandlad cancersmärta i livets slutskede. Motsvarande siffra för Sverige ligger kring 5000 personer. En stor bakomliggande faktor till underbehandlingen är att sjukvården inte är medveten om att patienten har ont.

När patienten inte själv kan förmedla sin smärta används idag så kallade observationssmärtskattningsinstrument. Vårdpersonalen observerar patienten, skattar de delar som ingår i instrumentet, och avgör sedan om patienten behöver någon smärtlindring. Eftersom det inte finns något specifikt instrument utvecklat för patienter med avancerad cancer, används i stället ibland instrument framtagna för patienter med demens. Ett av dessa instrument är Abbey Pain Scale, APS. Trots att det australiensiska APS är framtaget för att användas på demensboenden, används det runt om i Sverige även för att bedöma smärta hos patienter med avancerad cancer. Det är tidigare aldrig undersökt om instrumentet fungerar även för denna patientgrupp.

APS består av sex olika delar där vårdpersonalen skattar röstuttryck, ansiktsuttryck, förändrat kroppsspråk, förändrat beteende, fysiologiska förändringar (puls, blodtryck med mera) och kroppsliga förändringar. Varje del skattas från 0=inga besvär till 3=mycket besvär. Poängen summeras sedan ihop och omvandlas till olika nivåer av smärta, till exempel 0–2 poäng=ingen smärta medan 14 poäng och högre bedöms som svår smärta.

I den här avhandlingen som består av tre studier, har jag översatt och anpassat APS till en svensk vårdmiljö, efterforskat vårdpersonals erfarenhet av att använda APS på patienter med avancerad cancer, samt studerat om APS mäter smärta eller inte hos dessa patienter och hur träffsäkra resultaten är.

Studie I resulterade i en ny svensk version, APS-SE, som nu är redo att användas i olika typer av vårdmiljöer där patienter med demens vårdas. APS-SE är anpassad för att fungera oavsett språkmässig eller utbildningsmässig bakgrund hos den vårdpersonal som använder instrumentet.

Även om APS-SE inte är testat fullt ut på patienter med demens i en svensk vårdmiljö, tyder internationella studier på att instrumentet har en relativt god träffsäkerhet och förmåga att känna igen smärta hos patienter med demens. Om instrumentet ska användas är det dock viktigt att vårdpersonalen får en ordentlig introduktion och att man är noga med att de förstår alla ord och begrepp som ingår i APS-SE. Det kan finnas ett behov av att utbilda vårdpersonalen i hur man bedömer de ingående delarna 'blekhet' och 'rodnad' hos de patienter som har en mörkare hudton. APS-SE bör enbart användas i en vårdmiljö där vårdpersonalen känner patienten väl.

I studie II intervjuades läkare och sjuksköterskor om deras erfarenheter av att använda APS-SE vid smärtskattning av patienter med avancerad cancer. Det framkom att det finns ett behov av ett smärtskattningsinstrument i livets slut till de patienter med cancer som inte själva kan förmedla sin smärta. APS-SE fyller ett behov och fungerade både som stöd och påminnelse om vad vårdpersonalen ska bedöma när de smärtskattar patienterna.

Instrumentet upplevdes dock delvis missa målet, eftersom skattingsvärdena bedömdes som oprecisa och eftersom APS-SE visade på för låga smärtnivåer jämfört med vårdpersonalens egen uppfattning av patientens smärtnivå. Vissa av de exempel som enligt APS-SE tyder på smärta uppfattade vårdpersonalen snarare berodde på cancern i sig. De var också osäkra på om instrumentet enbart påvisade smärta, eller om det även kunde ge utslag för annat lidande, till exempel ångest eller oro.

Vårdpersonalen rapporterade också att APS-SE inte alltid passade den kliniska situationen. Vissa delar av instrumentet, som att till exempel mäta blodtryck, upplevdes inte tillföra smärtbedömningen något. Att mäta blodtryck på en döende patient uppfattades dessutom som oetiskt. Några deltagare upplevde också att de kände sig oempatiska när de använde instrumentet och befarade att patient eller närstående kände likadant när vårdpersonalen använde APS-SE. Sammantaget upplevdes APS-SE inte som optimalt för patienter med avancerad cancer, men användes i brist på något bättre. Problemet uppfattades bero på instrumentet i sig och inte på själva översättningen från engelska till svenska.

I studie III inkluderades patienter med avancerad cancer som antingen led av stor trötthet, förvirring eller medvetandesänkning. Alla patienter smärtskattades vid två olika tillfällen med en timmes mellanrum. Skattningarna gjordes av två olika bedömare samtidigt, men oberoende av varandra. Om möjligt skattade även patienten själv muntligt sin egen smärta vid båda mättillfällena. Information om skattningen förmedlades till ansvarig sjuksköterska som enligt klinisk rutin avgjorde om patienten skulle få smärtlindring eller ej.

När mätresultaten analyserats framkom att APS-SE har svårt att påvisa smärta hos patienter med avancerad cancer. Av de patienter som muntligt självskattade sin smärta hade hälften måttligt eller svår smärta – men APS-SE påvisade aldrig mer än lindring smärta hos någon av de inkluderade patienterna.

Det fanns en måttlig samstämmighet mellan de två olika bedömarnas totalpoäng. Instrumentet verkar vara stabilt över tid, eftersom APS-SE-poängen inte förändrades mellan första och andra mättillfället för de patienter som inte fick någon smärtlindring. Däremot sågs ingen samstämmighet mellan de olika delfrågorna, det vill säga frågorna verkar inte mäta samma sak. Slutligen visade analysen att APS-SE-poängen sjönk något för de patienter som fick smärtlindring, men även om minskningen var statistiskt säkerställd, är det oklart vilken klinisk betydelse detta har.

Sammanfattningsvis visade både studie II och III att APS-SE inte fungerar tillfredsställande vid smärtskattning av patienter med avancerad cancer. Instrumentet rekommenderas därför inte för smärtskattning av denna patientgrupp. Behovet att kunna smärtskatta patienter med avancerad cancer kvarstår dock och det är viktigt att ett observationssmärtskattnings-instrument för patienter med cancer utvecklas i framtiden.

Original papers

Study I

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. 2023;10(3):1367-1374. doi:10.1002/nop2.1386

Study II

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

Study III

Tegenborg S, Fransson P, Martinsson L. The Abbey Pain Scale: not sufficiently valid or reliable for assessing pain in patients with advanced cancer. Acta Oncol. 2023;62(8):953-960. doi:10.1080/0284186X.2023.2228992

Introduction

Pain and cancer

Pain is unique and personal experience. The International Association for the Study of Pain (IASP) defines pain as shown in Figure 1[1].

> 'An unpleasant sensory and emotional experience associated with, or resembling that associated with, actual or potential tissue damage.'

Figure 1. The International Association for the Study of Pain definition of pain.

Pain is multidimensional and does not consist solely of nociceptive and/or neuropathic pain caused by tissue or nerve damage. The great pioneer of palliative care, Dame Cicely Saunders, established the concept of 'total pain', which has had a major impact on the holistic paradigm in palliative care. The concept of total pain not only incorporates the physical dimension of pain, but also acknowledges the emotional, social, spiritual, and psychological elements of experiencing pain[2].

Nociceptive pain arises when certain receptors, the nociceptors, respond to stimuli from damaged tissue in the body. Neuropathic pain is the result of nerve damage. Patients with advanced cancer often suffer from both nociceptive and neuropathic pain[3]. It is important to distinguish between the different types of pain mechanisms since the treatment is partly different; for example, nociceptive pain responds better to opioids than neuropathic pain.

Within the concept of total pain, the emotional pain caused by cancer can manifest itself as worry or distress over the prognosis, pain management, or the consequences for loved ones. Pain can also lead to social withdrawal, creating isolation from the social network of friends and family. Spiritual pain is concerned with finding meaning and purpose in the suffering. Finding inner peace of mind, sometimes through faith, can help the patient to cope with pain by finding a sense of purpose while connecting to something greater than themself. The psychological element is concerned with the patient's sense of self-efficacy and coping mechanisms to manage the pain[4-6].

The prevalence of pain in patients with cancer varies during the disease trajectory. During treatment it is around 40–50%[7,8], but this proportion increases to 50–66% in patients with advanced disease or near the end of life[8-11]. Although not all of these patients will suffer from pain, as many will be on an effective pain management, it is estimated that 32–40% of all patients are undertreated[12-14]. Globally, this means that approximately a staggering two million people each year lack effective pain management. In Sweden, the number is roughly 5000 patients.

There are several obstacles to effective pain management. Poor pain assessment is recognized as a major barrier, while others include fear related to analgesic use and opioid side effects, and patients' reluctance to discuss pain[3,8,15-19]. In a global perspective, opioids may not even be available to relieve pain when approaching the end of life[20].

Pain is preferably self-reported[15,21,22]. Nevertheless, many patients are unable to express their pain, and in this scenario health care professionals (HCPs) have to rely on observational pain assessment instruments to detect pain. When using an observational instrument, the HCPs observe the patient and rate the included items in the instrument to assess the pain. However, no observational pain assessment instrument based on nonverbal pain behaviours exists for the oncology setting[15].

There can be many reasons why a patient may be unable to self-report pain, including cognitive impairment, delirium, sedation, unconsciousness, or imminent death[23-25]. Many of these patients suffer from intractable pain, and are in need of palliative care.

Palliative care

Definitions of palliative care

The concept of palliative care has changed and evolved since the first definition was published by the World Health Organization (WHO) in 1990. Moreover, the view on palliative care differs between different parts of the world. All of this makes it difficult for policy makers, the public, and even HCPs to fully grasp the concept[26,27]. The updated WHO definition from 2002 describes palliative care as an improvement of quality of life in patients with a life-threatening illness (Figure 2)[28].

> 'Palliative care is an approach that improves the quality of life of patients and their families facing the problem associated with life-threatening illness, through the prevention and relief of suffering by means of early identification and impeccable assessment and treatment of pain and other problems, physical, psychosocial and spiritual.'

Figure 2. The World Health Organization definition of palliative care.

Palliative care according to the WHO focuses on helping the patient to live their life to the fullest extent possible until death, with the help of a team approach that is integrated in the health care service. The health care should be person-centred and responsive to the patient's own preferences. It should include taking care of problems beyond only the physical issues[29], such as the bereavement process in the patient and family.

However, palliative care keeps evolving. Almost two decades later, another definition of palliative care was suggested by the International Association for Hospice and Palliative Care (IAHPC) in 2020[30]. The IAHPC recognizes that palliative care should be given to all ages, that it should respect the cultural beliefs of the patient and family, and that any intervention must be evidence-based whenever possible and be delivered in conjunction with disease-modifying therapies[30].

Unlike the WHO definition, which speaks of life-threatening illness, the IAHPC instead uses the term 'serious health-related suffering due to severe illness'[30]. This statement was chosen in order to better encompass all levels of health care systems including those in low- and middle-income countries[31]. It offers a way to move from a disease-centred approach to an increasingly person-centred framework, emphasizing that palliative care should be based on need and not prognosis[31]. Many of the patients in

need of palliative care suffer from illnesses that can be acute, but also can be temporary if treated, such as malnutrition or haemorrhagic fever[32].

The updated definition from the European Association for Palliative Care (EAPC), published in 2022, once more focuses on disease; in this case, disease that does not respond to curative treatment[33]. Similarly, the Swedish National Board of Health and Welfare (NBHW; Socialstyrelsen) focuses on progressive incurable disease, but also mentions injury[34,35].

Although the definitions from the WHO, the EAPC, and the Swedish NBHW are more restricted regarding which patients should receive palliative care compared to the definition from the IAHPC, all definitions state that palliative care not only includes treatment of the disease or injury but also involves addressing social, psychological, and spiritual problems[28,30,33,34]. This more holistic approach is something that partly sets palliative care apart from many of the other major medical disciplines, along with palliative care's approach on actively embracing the surrounding family members.

Palliative care in a global setting

A 2019 WHO survey revealed that funding for palliative care was available in 132 out of 194 countries worldwide[36]. It is estimated that the need for palliative care among people near the end of life ranges from 38% to 82% in different countries[37-40]. Many of these people live in low- and middle-income countries with little or no access to basic palliative care or proper pain management[41].

Cancer is one of the most common diagnoses for patients in palliative care worldwide[29,32,42,43] and according to the WHO approximately 10 million people die from cancer each year[44]. Many patients with cancer suffer from pain, and pain is also one of the most common symptoms in palliative care. Unfortunately, there is a great inequality in access to pain relief in the global setting. As opioids are essential for managing pain in many palliative patients, good palliative care is hindered by unnecessarily restrictive regulations or lack of access to morphine or other palliative medications[29]. The poorest 50% of the world population receive only 1% of the available opioid analgesics[20].

According to the WHO, receiving palliative care is part of the human right to health. In reality, only 14% of all people worldwide in need of palliative care obtain it[29].

Palliative care in Sweden

Each year approximately 90 000 people die in Sweden, comprising roughly 1% of the population. It is estimated that about 80% of these patients have a foreseeable death, and hence should be able to receive palliative care[45]. According to a Swedish Government Official Report from 1995, the care of end-of-life patients should be prioritized as equal to acute life-saving procedures[46].

Palliative care in Sweden is integrated into the broader healthcare system, and can be divided into basic and specialized palliative care[47]. Basic palliative care is available throughout the healthcare system; at home, in nursing homes, and in hospitals. These healthcare units can provide elementary palliative care to patients with less complicated palliative needs. In many cases, general practitioners and district nurses play an essential role in providing basic palliative care[45,47,48].

However, there will always be some patients who require specialized palliative care, due to the complexity of their situation or symptoms such as intractable pain. A specialized palliative care unit often consists of either a specialized palliative home care unit that provides health care in the patient's own home, or a standalone specialized palliative care unit (i.e., 'hospice' in Swedish). Specialized palliative care can also be provided in palliative in-patient care units in the hospitals[45].

If a specialized palliative care unit is not available, for example due to geographical distances, specialized palliative care can be offered through a collaboration between a specialized palliative consultant team and the primary care provider[45,48].Specialized care units often consist of specialist physicians and nurses in palliative medicine, together with other professions such as occupational therapists, physiotherapists, dieticians, social workers, and spiritual counsellors[49]. Although there is no official record of the exact number of specialized palliative care units in Sweden, it is estimated to be around 130 [M Andersson, Director of the Swedish Register of Palliative Care (SRPC), personal communication].

According to a report from the SRPC, 11% of the patients who died in Sweden in 2022 received care from specialized palliative care units; if specialized consultant teams are included, this proportion rises to 18%[48]. The majority of patients who receive specialized palliative care (over 80%) are patients with cancer[45]. Cancer is the second most common cause of death in Sweden, surpassed only by cardiovascular disease[50]; overall, around 23 000 people die of cancer in Sweden each year[51].

Pain and pain assessment instruments

An investigation of the priorities and wishes among people with advanced disease in palliative care in Sweden found that patients number one priority was to be free from pain[52]. Still, many patients with cancer pain are undertreated[12-14], and one of the factors hindering beneficial pain management is poor pain assessment[3,15-19]. The use of pain assessment and reassessment after any intervention is a key factor in successfully managing cancer pain[15].

Self-report instruments

Since pain is inherently subjective, self-reported pain assessment is the gold standard[15,21,22]. Self-reporting can be performed by the patients themselves, or with the help of an assisting HCP or family member. One way to assess pain is by using a pain assessment instrument. An instrument is something that can be used when gathering data or measurements that can be analysed with regard to the chosen outcome(s)[53,54].

Pain can be assessed either as the only entity, or as part of an instrument that covers multiple symptoms simultaneously in the palliative patient. Two common multiple symptoms scales are the Integrated Palliative care Outcome Scale (IPOS)[55] and the Edmonton Symptom Assessment System-revised/Edmonton Symptom Assessment Scale-revised (ESAS-r)[56]. The IPOS and the ESAS integrate several common symptoms such as pain, nausea, depression, and anxiety, to be evaluated by the patient all at the same time. For cancer-related symptoms and their impact on daily living, the MD Anderson Symptom Inventory (MDASI), is an option[57].

When estimating pain as a single entity, two different types of instruments can be used: unidimensional instruments for pain intensity, or multidimensional instruments that examine multiple aspects of pain. The McGill Pain Questionnaire (MQP)[58] is an example of the latter that focuses on both the quality and intensity of the pain. Another multidimensional instrument is the West Haven-Yale Multidimensional Pain Inventory (WHYMPI)[59], which incorporates several dimensions of the impact of chronic pain on the patient's life and daily activity.

There also exist multidimensional instruments developed specifically for patients with cancer; for example, the Edmonton Classification System for Cancer Pain (ECS-CP)[60]. This instrument integrates psychological distress, addictive behaviour, and cognitive function along with incident pain and the mechanism of pain, with the aim of guiding HCPs in how to

optimize pain control. Another widely used instrument is the Brief Pain Inventory (BPI), which considers the severity of pain along with the impact of pain on daily functions[61]. The BPI has also been validated for non-cancer patients[62].

Among the unidimensional instruments assessing pain intensity, the most commonly used are the visual analogue scale (VAS)[21] and the numerical rating scale (NRS)[21]. There are several different versions of the VAS, but it usually consists of a straight horizontal line measuring 100 mm from end to end. The scale is often divided into 11 different numbers, with 0 at the left representing 'no pain' and 10 at the right representing 'worst pain imaginable'.

First, the patient places a mark on the scale corresponding to the intensity of the pain, from no pain to worst pain imaginable[21] (Figure 3).

Figure 3. The patient's view of the visual analogue scale.

Next, the HPC converts the mark to a number representing the pain level on the other side of the instrument (Figure 4).

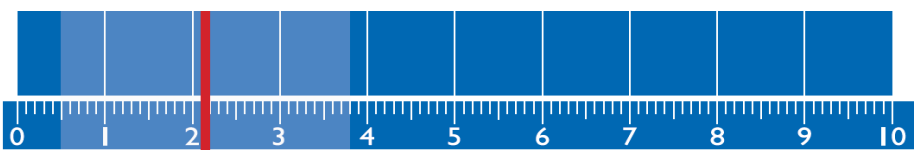


Figure 4. The health care professional's view of the visual analogue scale, converting the patient's mark on the instrument to a number.

The NRS uses the same 11 number scale, but instead of marking the line, the patient verbally rates the pain intensity directly in numbers from 0 to 10; that is, from 'no pain' to 'worst pain imaginable'[21]. However, many different descriptors are used to anchor the extremes of both the VAS and the NRS, especially for the anchor corresponding to the number 10[22]. This causes a problem both when trying to compare different studies and in attempts to harmonize and create an international standard for the anchor labels[63].

Both the VAS and the NRS have been found to be valid and reliable[21,22,64], although the NRS is often recommended for both patients with cancer and for patients in general, on the basis of patients' preference and the effort to standardize which pain assessment instrument to use[22,63,65-67].

Another commonly used unidimensional pain assessment instrument is the verbal rating scale, also known as the verbal descriptor scale [21], which uses descriptors rather than numbers to measure pain intensity. As with the VAS and NRS, there are several versions available; the number of descriptors ranges from 4 (none, mild, moderate, or severe pain) to as many as 15[21]. There seems to be a good correlation between the verbal rating scale, the VAS, and the NRS. The verbal rating scale is considered reliable, valid, and appropriate to use in a clinical environment, and has been shown to be preferred by patients with less education or with mild to moderate cognitive impairment[21,22].Examples of different types of self-report pain assessment instruments are given in Table 1.

Table 1. Different types of self-report pain assessment instruments.

Abbreviated name	**Name**
	Instruments assessing multiple symptoms
IPOS[55]	Integrated Palliative care Outcome Scale *Available in Swedish*[68]
ESAS-r[56,69]	Edmonton Symptom Assessment System-revised/ Edmonton Symptom Assessment Scale-revised *Available in Swedish*[70]
MDASI[57]	MD Anderson Symptom Inventory *Not available in Swedish*

Table 1. *Continued.*

Abbreviated name	Name
	Multidimensional pain assessment instruments
MPQ[58]	McGill Pain Questionnaire *Available in Swedish*[71]
WHYMPI[59]	West Haven-Yale Multidimensional Pain Inventory *Available in Swedish*[72]
ECS-CP[60]	Edmonton Classification System for Cancer Pain *Not available in Swedish*
BPI[62]	Brief Pain Inventory *Available in Swedish, translation not published in a peer-reviewed journal*
	Unidimensional instruments assessing pain intensity
VAS[21]	Visual analogue scale *Available in Swedish, translation not published in a peer-reviewed journal*
NRS[21]	Numerical rating scale *Available in Swedish, translation not published in a peer-reviewed journal*
VRS/VDS[21]	Verbal rating scale/ Verbal descriptor scale *Available in Swedish, translation not published in a peer-reviewed journal*

Observational instruments

When patients are not able to verbalize their pain, there is a high risk of both under-assessment and under-treatment of the pain[73,74]. If self-reporting is not possible, observational pain assessment instruments must be used to assess the pain.

When using an observational instrument, the assessment is made by a proxy. The proxy is usually an HCP, but these assessments are sometimes performed in conjunction with a friend or a family member[75]. Observer-rated pain assessment seems to be capable of recognizing if pain is present, as well as indicating the intensity of the pain[24]. Many of the existing observational pain assessment instruments were developed for patients with dementia.

Observational instruments for people with dementia

In 2022, a systematic review by Smith and Harvey[74] used the COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN) checklist[76] to evaluate different observational pain assessment instruments. COSMIN was developed to make recommendations for choosing health measurement instruments in both clinical and research settings[76]. Following this checklist, the authors suggested eight different instruments (Table 2) but concluded that the instruments Abbey Pain Scale (APS)[77] and the Pain Assessment in Advanced Dementia Scale (PAINAD)[78] could not be recommended due to limited evidence.

Table 2. Observational pain assessment instruments for people with dementia as recommended by Smith and Harvey[74].

Abbreviated name	Name
ALGOPLUS[79]	ALGOPLUS (same as abbreviated) *Not available in Swedish*
CNPI[80]	Checklist of Nonverbal Pain Indicators *Not available in Swedish*
DOLOPLUS-2[81]	DOLOPLUS-2 (same as abbreviated) *Available in Swedish*[82]

Table 2. *Continued.*

Abbreviated name	Name
FACS[83]	Facial Action Coding System *Not available in Swedish*
MOBID[84]	Mobilization-Observation-Behaviour-Intensity-Dementia Pain Scale *Not available in Swedish*
MOBID-2[85]	Mobilization-Observation-Behaviour-Intensity-Dementia Pain Scale-2 *Not available in Swedish*
PACSLAC[86]	Pain Assessment Checklist for Seniors with Limited Ability to Communicate *Available in Swedish*[87]
PACSLAC-II[88]	Pain Assessment Checklist for Seniors with Limited Ability to Communicate-II *Not available in Swedish*

Two of the instruments suggested by the abovementioned review have been translated into Swedish; DOLOPLUS-2[82] and the Pain Assessment Checklist for Seniors with Limited Ability to Communicate (PACSLAC)[87]. No specific instrument for assessing pain is recommended in the Swedish National Guidelines for Care in Dementia[89], but the Swedish National Clinical Practice Guideline for Palliative Care[90] recommends DOLOPLUS-2, SÖS-stickan, and the APS.

DOLOPLUS-2 is a multidimensional instrument that incorporates somatic reactions as well as psychomotor and psychosocial reactions[81,82], while both SÖS-stickan and the APS are unidimensional instruments focusing on the intensity of pain. SÖS-stickan, which was developed at Södersjukhuset in Stockholm, Sweden, is a behavioural scale for patients with impaired cognitive function in emergency care; according to its developers, the pain assessment is preferably made during activity[91]. The APS is described in greater detail on pages 12–15 of this thesis.

Observational instruments for people with cancer

The majority of people with cancer will experience pain when near the end of life[8,10], but as the trajectory of the cancer proceeds, their ability to vocalize pain will often be compromised due to reduced consciousness, pronounced delirium, or being near death[25]. In this scenario, there is a need for an observational pain assessment instrument, but no such instrument is currently available for the palliative oncology setting[15].

The WHO guidelines for management of cancer pain in adults do not address the issue of pain assessment in end-of-life[92], while the guideline on the use of opioids for adults with pain from cancer from the American Society of Clinical Oncology (ASCO) does not mention pain assessment at all[93]. The clinical practise guideline for management of cancer pain in adult patients from the European Society for Medical Oncology (ESMO), stipulates that initial and continuous pain assessment is a fundamental part of cancer care, but only recommends 'observations of pain-related behaviours' if the patient with cancer has a cognitive impairment[94].

There is no specific mention or recommendation for patients with cancer and cognitive impairment in the Swedish National Clinical Practice Guideline for Palliative Care[90]. The 2016 version of this guideline (which was the current version when the present project began) recommended three different pain assessment instruments for adult patients with cognitive impairment, including the APS[95]. This recommendation was amended in the updated version from 2023, and the guideline now recommends the APS for use in people with dementia[90].

The Abbey Pain Scale

The APS was originally developed by Abbey et al. in Australia during 1997–2002[77], with the aim of creating a reliable and efficient instrument for assessing pain among late- or end-stage dementia patients in residential aged care facilities. The scale was intended to be easy to use, and to be used by diverse HCPs.

The first draft of the instrument was based on previous work by Hurley et al.[96] and Simons and Malabar[97], and then modified by pain experts and gerontological specialists through an international Delphi study. The Delphi process is a method for gaining consensus through feedback from experts/specialists in an area that usually has limited or conflicting evidence[98-101]. Additional modifications were made after focus group interviews with nurses and medical practitioners[77].

Further development took place in two stages, as the instrument was refined by reducing the number of pain indicators from 12 to 6 and by creating a summative scale between the first and second stage. The first stage involved 52 residents with end- or late-stage dementia who were unable to describe their pain consistently or coherently, while the second stage included 61 residents with the same characteristics[77].

During the original development of the APS, the scale was administered whenever the HCPs perceived that a resident was experiencing pain[77]. As well as administering the APS, the HCPs were asked to classify the pain as acute, chronic, or acute on chronic, and to rate the observed pain on a scale from no pain to severe pain; this was done to provide an overall holistic subjective assessment of the observed pain that would serve as a baseline[77]. The assessment was then repeated 45 minutes after the pain-reliving intervention, by the same HCPs. During the second stage, whenever possible, the assessments were made by two separate HCPs independently from each other[77].

The finished version of the APS consists of six items, five of which are rated using observation. In one item, in addition to observation, the HCPs are asked to measure temperature, pulse, and blood pressure. Pulse and blood pressure are to be rated against normal limits as shown in Table 3.

Table 3. The Abbey Pain Scale, with suggested examples for each item[77].

Items	Suggested examples
Q1. Vocalization	whimpering, groaning, crying
Q2. Facial expression	looking tense, frowning, grimacing, looking frightened
Q3. Change in body language	fidgeting, rocking, guarding part of body, withdrawn
Q4. Behavioural change	increased confusion, refusing to eat, alteration in usual patterns
Q5. Physiological change	temperature, pulse or blood pressure outside normal limits, perspiring, flushing or pallor
Q6. Physical changes	skin tears, pressure areas, arthritis, contractures, previous injuries

Each item has four grades of severity: 0=absent, 1=mild, 2=moderate, and 3=severe, giving a total score of 0–18. The total pain score is then converted into different categories of pain: 0–2 points=no pain, 3–7 points=mild pain, 8–13 points= moderate pain, and 14–18 points=severe pain[77].

Psychometric testing of the APS has shown a reasonable degree of agreement between the APS pain scores and the HCPs' overall holistic pain assessments. Moreover, the mean APS scores decreased when measured before and after a pain-relieving intervention. However, only a modest match in scoring was seen between HCPs assessing the same patient[77]. More information on the testing of instruments is given on pages 15–23 of this thesis, and the psychometric attributes of the APS are discussed in more depth in the Discussion (see pages 56–60).

After its publication in 2004, the instrument soon became the most universally used pain assessment instrument in Australia for people with dementia in nursing homes[102]. It has also been translated into several other languages including Danish[103], Japanese[104], Spanish[105], and Italian[106]. Besides being psychometrically tested in conjunction with these translations, it has been tested in Australia in comparison with other assessment instruments such as DOLOPLUS-2 and the Checklist of Nonverbal Pain Indicators (CNPI)[107]. A non-validated APS translation to Dutch was compared to the PAINAD in palliative patients with cognitive impairment in nursing homes[108]. The APS has also been psychometrically tested for a different population; elderly people with osteoarthritic pain[109].

During 2010–2020, a non-scientifically-translated Swedish version of the APS[110] was distributed by the Swedish Register of Palliative Care (SRPC) [M Andersson, Director of the SRPC, personal communication]. The SRPC is a national quality register that was developed to improve end-of-life care[111].

In 2015, a scientific translation of the APS to Swedish, the APS-SWE, was created as part of a doctoral thesis[112]. However, when approached during the start of this project in 2017, the researcher did not give permission to use the APS-SWE, since it had not yet been published in a peer-reviewed journal. The APS-SWE has been distributed via the SRPC since late 2020.

During the years since 2010, Swedish translations of the APS have been distributed to more than 850 different health care units throughout Sweden, and the instrument is now used in end-of-life care in nursing homes, in

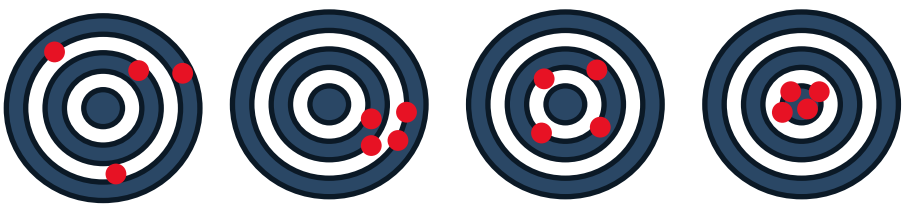
specialized palliative care units, and in some hospitals [M Andersson, Director of the SRPC, personal communication].

The use of validated and reliable instruments is imperative for maintaining and improving the quality of healthcare and research. Today, different Swedish versions of the APS are continuously used to assess pain in patients with cancer in Sweden, but until the present project there has been no knowledge of their psychometric values in this population.

The evaluation of instruments

The psychometric testing of an instrument is the science of evaluating whether the instrument's measurement properties produce valid and reliable results when applied in a specific population. Psychometric testing usually consists of different validity and reliability tests, and sometimes also tests for responsiveness[113].

Validity describes how well the instrument measures what it intends to measure, while reliability describes the extent to which the outcome is repeated if the instrument is used under the same conditions[113]. When determining whether an instrument is useful when applied to a population, the scores generated by the instrument must be both accurate and repeatable[113]. In the case of observational pain assessment, not only must the instrument correctly answer the question of how much pain the patient has, but different raters should also agree on and arrive at the same estimated pain score. The instrument is appropriate to use when the validity and reliability superimpose into a bull's-eye (Figure 5).



Low validity and reliability.	Reliability is good, but validity is not.	Validity is good, but reliability is not.	Validity and reliability superimpose.

Figure 5. Schematic illustration of validity and reliability.

Validity

Validity describes how accurately the instrument measures the construct it is intended to measure[113]. A unidimensional instrument such as the NRS is easier to validate than multidimensional instruments such as the ESAS or the IPOS. When validating a multidimensional instrument, the scales for each dimension must be validated separately[113].

If the instrument is being used in a context other than the one it was originally validated for, it must be re-validated. Thus, a new validation is required if using the instrument in a different population, setting, or language. In the latter case, a cross-culture validation is usually performed[113], as in Study I in this thesis. Since the initial framework for an instrument evolves over time, and instruments are often applied in different situations, the need for validation is a continuous process[113].

Several different types of validity exist, with the most commonly-used being face validity, content validity, criterion validity, and construct validity[113,114].

Face validity assesses how suitable the instrument seems on the surface; that is, on the face of it. It is similar to content validity (see below) but is more informal. This view of the instrument is considered the weakest form of validity, but it can provide an insight into how future raters will respond to the different items[113,114].

Content validity assesses whether the instrument is representative of all aspect of the construct being measured. The instrument must be both relevant and comprehensive. For example, does the instrument cover all aspects of pain? Is some item missing or irrelevant when assessing pain in a specific population, such as patients with advanced cancer?[113]. The process of developing an instrument begins with a literature review to identify potential items for the instrument, and continues with all the potential items being reviewed by a panel of experts. The most important outcome measures for the instrument can then be selected through consensus among the experts. As with face validation, there is no statistical testing; instead, the validation is based on the judgment of relevant experts[113,114].

Criterion validity assesses how well the instrument correlates with an established standard of comparison (i.e., a criterion variable). The criterion should be considered valid, and is sometimes referred to as the ‘gold standard’[113]. The gold standard in pain assessment is self-reported pain

assessed using the NRS[15,21,22]. Criterion validity is determined by calculating the agreement between the outcome from the instrument and the gold standard. If there is a high correlation according to the statistical calculation, this suggests that the instrument measures what it is intended to assess[113,114].

Criterion validity itself can be divided into *concurrent validity* and *predictive validity*. The difference lies in the timing of the evaluation. Concurrent validity is the agreement between the instrument and the criterion when the assessments are made at the same time. Predictive validity is the agreement between the score of the instrument and some future criterion, determining whether the instrument's score can predict a future outcome[113,114].

Construct validity assesses whether the instrument's items and the provided scores seem to be related to an existing framework and concept of the construct[113,114]. This type of validity can be useful if there is no accepted gold standard for the assessed construct[113].

After formulating a hypothesis of how the measurements will compare to other instruments measuring related constructs, two subgroups of construct validity can be created: *convergent* and *discriminant validity*. These are related but opposite concepts. Convergent validity explores positive correlations between similar constructs, while discriminant validity is based on the hypothesis of there being little or no correlation between the instruments. In *discriminative validity*, mean scores for different subgroups could be analysed to differentiate between, for example, mild, moderate, or severe pain[113].

While content, criterion, and construct validity are concerned with the accuracy of the test, questionnaire, or measurement method itself, other dimensions of validity also exist that can be used in contexts such as clinical medical studies. For example, validity can be divided into *internal* and *external validity[115].*

Internal validity is concerned with the validity of the results within the test. It is often used when evaluating experimental research. Internal validity is concerned with the research design and the strength of conclusions that can be drawn from it; that is, how accurately it establishes cause-and-effect between the investigated variables[115,116].

To achieve internal validity, any intervention must predate the response variable, and both variables must change together. The result should not be explained by confounding or external factors. Thus, it is easier to achieve a high internal validity with designs such as an experimental laboratory study, where external factors can be kept under better control, than when conducting research in, for example, a health care environment[115,116].

Ways of improving internal validity include the use of standardized procedures, valid and reliable measures, randomization, and/or control groups. Another way is to increase the sample size, thus improving the statistical power as described below. A large sample size also enhances the external validity, since it improves the probability that the sample is representative of the larger population[115,116].

External validity or *generalizability* is concerned with the validity of applying the result of a study to a broader context, such as a different population or setting[115,117]. There can sometimes be conflict between internal and external validity. For example, a randomized controlled trial with very strict inclusion and exclusion criteria might produce results that have excellent internal validity, but that are hard to generalize to clinical reality with a diverse population of patients[118]. Investigators generally want to demonstrate a good effect of the intervention without the risk of any complications, and so patients with poor performance, any form of affected organ function, or with medical comorbidities are likely to be excluded by strict exclusion criteria in studies[119,120]. This becomes a problem in palliative oncology care, due to uncertainty over whether the results from these types of strictly regulated studies are applicable in a palliative care setting.

Reliability

Reliability describes how consistently an instrument performs when used repeatedly in a population. The reliability of an instrument is assessed by investigating the results across time, between different raters, and across the different items in the instrument itself. Some common types of reliability are test-retest reliability, inter-rater reliability, and internal consistency reliability[113,121].

Test-retest reliability assesses the consistency and reproducibility across time when the instrument is applied on separate occasions in the same population. It tests the instrument's stability and ability to resist external factors that can influence the results over time. There is no consensus on what the time span between the first and second test should be, and

different studies have used spans varying from hour(s) to months[114]. If measuring a stable construct, two to four weeks is a generally accepted time span; any shorter, and there is a risk of the raters remembering the ratings they gave at baseline. If longer time is allowed to elapse then the scores might change for other reasons, such as increased pain due to worsening of the underlying disease[114].

Inter-rater reliability, which is also known as *inter-observer reliability*, assesses how well two or more independent raters agree when measuring the same thing. Since rating is subjective, there will be an anticipated difference between the raters' perception of the observed items[113,121].

Intra-rater reliability assesses how the same rater assesses the items on different occasions or in different studies. It measures the single-rater consistency during repeated administrations of a test[113,121].

Internal consistency reliability assesses how well the different items in an instrument relate to each other, in order to see whether they measure the same construct. If the ratings of the different items in the instrument will end up in a combined score, for example a pain score, then each item must be essential for the overall score. Internal consistency measures how well these different items come together conceptually. A higher reliability can be achieved by adding more items if they are relevant. If adding an item that does not really measure the construct of interest, it will decrease the reliability. There is no real consensus how high the reliability should be, but if it is very high, it can be argued that the instrument could be shortened[114].

Parallel forms reliability, which is also known as *alternative forms reliability,* assesses reliability by applying two or more different versions of an instrument to measure the same construct. After making a collection of different items evaluating the same concept, the items are usually spilt into different versions. However, generating enough equivalent items for several instruments may be difficult. One advantage of using parallel instruments is that it minimizes the risk of the rater remembering their previous response[114].

Responsiveness

Responsiveness is the ability of an instrument to accurately detect change in the construct of interest[113]. For example, one might ask whether an instrument can detect a change in pain intensity if the patient is given as-needed opioids.

The literature contains many different definitions of responsiveness, as well as many different suggestions for how to measure it[113]. One ongoing discussion is whether responsiveness should be considered as an aspect of validity (the validity of a change score) or if it is an entity of its own. If responsiveness is considered to be one facet of validity, then it can be determined through a criterion or construct approach[113].

Criterion approach responsiveness can be used when a gold standard (i.e. a criterion) is available and the gold standard also demonstrates responsiveness to the construct of interest. The assessments with the instrument and the gold standard should be executed independently but at the same time, before and after any intervention. The agreement between the calculated changes is then determined in the same way as when assessing criterion validity[113].

Content approach responsiveness can be used if no gold standard is available, but it is much less commonly utilized. In this case, it is important to hypothesize the anticipated outcome a priori of the assessments; after an intervention, should the outcome measure increase or decrease, and by how much? Without this kind of specification of the predicted differences or correlations, it is difficult to state whether the hypothesis has been confirmed or not[113].

Trustworthiness

Qualitative research can be evaluated in terms of *trustworthiness*[122-125], which addresses the question of whether a study's results can be trusted. Trustworthiness is an 'overarching concept' consisting of different approaches to describe the concept within different qualitative research designs[123]. It can be evaluated through credibility, authenticity, dependability, confirmability, and transferability, as suggested by Lincoln and Guba[126], and is a commonly-used approach in qualitative content analysis[124].

Credibility assesses whether the research result is believable; for example, whether the information is in line with the data gathered from the participants and whether the authors' interpretation is representative of the participants' views[122-126]. It is considered the most important part of trustworthiness. Credibility in qualitative research corresponds to internal validity in quantitative research[125].

To achieve credibility in qualitative content analysis, it is important to ensure that enough data are collected. The necessary amount of data depends on the richness of the assembled data, including the participants' various experiences of the studied phenomena, as well as the overall number of participants, as described by Graneheim et al (2017)[123]. Furthermore, all analysis includes interpretation of the data. Providing examples of how the qualitative content analysis was performed, together with the use of representative quotations, helps the reader to evaluate the credibility. Using quotations also enhances the participants' voices over the voices of the researchers in the presentation of the results[123].

Additional ways to enhance credibility in qualitative research in general have been described in a methodological article on trustworthiness[125]. *Prolonged engagement* involves making sure that the interviews are long enough for the researcher to establish trust and become familiar with the context and setting. *Triangulation* is conducted by employing different approaches such as using different investigators, different data sources, or different methods when collecting data. To further strengthen the credibility, a *member check* can be used to gain feedback on the results from the participants themselves[125].

Authenticity assesses the degree to which the researchers have managed to capture and show the range of the different perspectives among the participants. According to Graneheim et al. (2017), the higher the level of interpretation and abstraction, the more difficult it is to demonstrate credibility and authenticity. However, giving examples of the interpretation and abstraction process through representative quotations allows the reader to more easily evaluate the authenticity of the results[123].

Dependability assesses the stability of the research results over time and conditions. It evaluates the quality of the analysis process and whether the standards for the chosen research design have been followed[122-126]. Conducting an extensive data collection over a long period of time carries the risk of inconsistency when the gathering of data. Another challenge is that the researcher's pre-understanding of the studied phenomena will change as new insights are gained during the data collection. This may

influence the choice of which questions or follow-up questions to ask, as well as the interpretation of the collected data. One way to address dependability is by involving several researchers in the analysis process[122,123].

Confirmability assesses the degree to which research results could be confirmed by other researchers[123-125]. Graneheim and Lundman (2004) do not discuss confirmability other than by noting that the participants' recognition of the results can be an aspect of confirmability[122]. According to the abovementioned methodological overview of trustworthiness in qualitative research in general, both dependability and confirmability can be assessed with the help of an audit. An external expert should be able to follow how the data collection, analysis, and interpretation were performed, and there should be transparency regarding how the research process developed throughout the project. In theory, another researcher should be able to confirm the results by being able to repeat the study[125].

Transferability assesses whether the results are transferable to other settings and populations. Graneheim et al. (2017) emphasize the importance of the included participants, as transferability can be enhanced if the participants provide a variety of perceptions[123]. Moreover, as it is the reader who ultimately decides whether the results are applicable in their own clinical setting, it is important for researchers to provide a 'thick description' of the context in which the research was performed. Without a thorough description, the relevance of the result is difficult to evaluate[123]. Transferability in qualitative research corresponds to external validity or generalizability in quantitative research[127].

Sample size

The sample size must be large enough to ensure a valid research result. In quantitative research, the sample size can be estimated with the help of a power analysis.

Power is concerned with the question 'How many participants are needed in order to ensure a certain probability of detecting a clinically relevant difference at a defined significance level?'[128]. The researcher has to start by deciding what is to be considered a clinically relevant outcome. For example, if the outcome is pain relief, one might ask how large a decrease would be needed in the NRS score for the outcome to be considered relevant. The researcher must also decide the probable standard deviation of the outcome measure, along with the probabilities of type I and type II errors[128].

A *type I error* is when the null hypothesis is wrongfully rejected; that is, the null hypothesis is true, but the results suggest that it is false. A null hypothesis is the hypothesis of there being *no difference* between the outcome from the experimental intervention and the outcome from the control intervention. The likelihood of a type I error is called the significance level; it is usually set at 0.05, meaning that the risk of the null hypothesis being wrongfully rejected is 5% or less.

Conversely, a *type II error* occurs when a null hypothesis is wrongfully accepted; that is, the null hypothesis is false, but the results suggest that it is true. The probability of a type II error is usually set at 0.10 or 0.20[128].

As an example, consider a study that measures the outcome 'pain relief' in order to decide whether a new substance X (the experimental intervention) is better than an opioid (the control intervention). Thus, the null hypothesis is that there is no difference between the pain-relieving ability of substance X and of the opioid. A type I error means that the study finds that the substance X is better than the control at relieving pain, even though it actually is not. A type II error means that the study finds no difference between the interventions, even though a difference actually exists.

The statistical power of a trial is calculated as 1 minus the probability of a type II error. If the probability of a type II error is set at 0.2, then the power is 1 - 0.20=0.80, meaning that a true difference will be found in 8 out of 10 tests. There are several ways to enhance the power; for example, by increasing the significance level or by increasing the sample size[128].

In qualitative research, the sample size is instead determined on the basis of being able to gather sufficiently varied data[123]. Data collection should continue until the point is reached where additional data do not contribute any new insights. According to Graneheim et al. (2017), the necessary number of individual interviews cannot be predicted, as it depends on the research question(s) and the richness of the data. However, several attempts have been made to estimate this number. According to a systematic review from 2022 focusing on sample size in qualitative research in general, the number of individual interviews needed ranges from 9 to 17, with a mean of 12–13 interviews[129].

Aims

The overall aim of this thesis was to evaluate observational pain assessment for people with advanced cancer in a palliative care setting, focusing on the APS.

Study I
The aim was to translate and culturally adapt the APS for people with end-stage dementia in various care settings in Sweden, and to investigate factors important for clinical implementation.

Study II
The aim was to explore physicians' and nurses' experiences of using a Swedish translation of the APS (the APS-SE) in people with advanced cancer.

Study III
The aim was to assess the APS-SE regarding validity, reliability, and responsiveness to opioids for patients with advanced cancer in a palliative oncology care setting.

Materials and Methods

This thesis used various study designs to evaluate an observational pain assessment instrument, the APS, for people with cancer in a palliative care setting. A summary of the different methods is given in Table 4.

Table 4. Overview of the studies.

Study	Design	Participants	Setting	Method
I	Qualitative research with semi-structured interviews	Interviewees: n=6 physicians n=3 nurses n=2 nursing assistants	Geriatrics, emergency care, nursing homes, specialized palliative home care unit	Translation with forward and backward techniques and thematic analysis of interviews
II	Qualitative research with semi-structured interviews	Interviewees: n=6 physicians n= 6 nurses	Oncology hospital wards, specialized palliative care units	Qualitative content analysis of interviews
III	Quantitative research with prospectively collected data	Patients: n=72	Oncology hospital wards, standalone specialized palliative inpatient care unit	Criterion validity, test-retest, inter-rater, and internal consistency reliability, and responsiveness to opioids were analysed using different statistical methods

Study I

Design

A qualitative research design with translation using forward and backward techniques and a thematic analysis of semi-structured interviews were used when developing a Swedish version of the APS, the APS-SE, for people with dementia.

Participants and data collection

The process of translation and cultural adaptation of the APS consisted of two stages: a translation with forward and backward techniques, and then adaptation to a Swedish care context through a series of interviews. A cross-cultural adaptation includes both the process of translation and cultural adjustments. This technique is used when preparing an instrument for use in another context (in this case, a Swedish care context), instead of just a literal translation of the words[130].

The process of translating and adjusting the APS-SE was primarily based on two different guidelines. The first guideline, by Beaton et al.[130], is a well-established general guideline for healthcare instrument translations, while the second is an established guideline in the palliative care context from Antunes et al.[131]. The suggested psychometric testing in the latter guideline was addressed in Study III. The APS-SE was developed in six steps (Figure 6.)

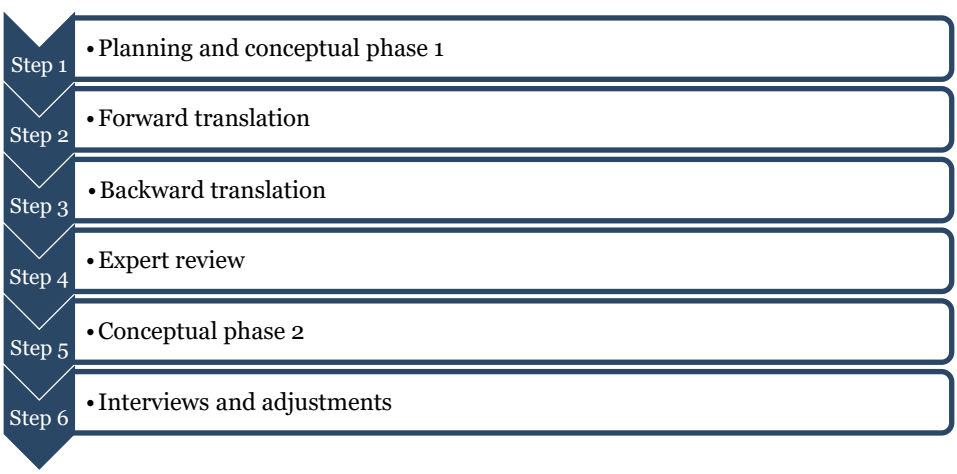


Figure 6. This Swedish version of Abbey Pain Scale was developed in six steps.

Step 1: Planning and conceptual phase 1

A literature review of translation guidelines and pain assessment instruments was performed as part of planning the study. This was followed by the first conceptual phase, which included an informal group discussion on the different APS items together with HCPs: three nurses in palliative care, a senior lecturer in nursing, and a manager of a palliative care unit. All of them had previous experience of patients with dementia.

Step 2: Forward translation

The forward translation (i.e., the translation of the original English version of the APS into Swedish) was performed independently by two different native Swedish speakers. One was a professional translator and the second was a physician specializing in both primary care and palliative medicine. Consensus on the different translations was then reached during a meeting with the translators and a mediator, resulting in a preliminary Swedish version of the APS.

Step 3: Backward translation

The mediated Swedish version was translated backward from Swedish to English, independently by two different native English speakers. One was a professional translator and the second was a nurse working in a palliative care setting. To minimize bias, the translators were asked to avoid studying the original English version of the APS ahead of the translation. Consensus on the different translations was then reached during a meeting with the translators and a mediator.

The purpose of this backward translation was to help ensure that the translation accurately reflected the item content in the original version and to uncover any conceptual errors in the translation.

Step 4: Expert review

The translations were reviewed and discussed in comparison with the original APS by an expert group consisting of the professional translators, the HCPs involved in the translation, and the researchers. Consensus was reached on the translations of all items except two.

Step 5: Conceptual phase 2

A renewed literature review concerning the remaining two items was conducted, along with consultation with experts and colleagues. Next, the new alternatives for the items were proposed to the expert group. When consensus was reached, this Swedish version of the APS (the APS-SE) was ready for testing.

Step 6: Interviews and adjustments

HCPs were interviewed to uncover potential difficulties with the initial APS-SE. When recruiting the HCPs, the aim was to achieve variation in occupation (nurses n=3, nursing assistants n=2, physicians n=6), age (range 20–65 years), and gender (women n=6, men n=5). All of them had experience in caring for and interacting with patients with dementia either in a geriatric department (n=2), in a nursing home (n=4), through specialized palliative home care (n=1), or in an emergency department (n=4). The nursing assistants had the shortest duration of working experience with people with dementia (range: 1–8 years) while the physicians had the longest (range: 5–25 years). The majority had Swedish as a first language (n=7) versus non-Swedish (n=4).

The HCPs were recruited through snowball sampling, a technique where recruitment continues based on referrals from participants who are already involved[132]. When approached by email, by phone, or face-to-face, all HCPs accepted the invitation. Before each interview, the researcher(s) presented herself/themselves, gave oral and written information concerning the study, and collected oral and written informed consent.

The interviews were conducted in Swedish, in the participants' own workplace, with no one else present except the participant and the researcher(s). A semi-structured interview guide was used, with questions such as 'What do you think of question 1?'. The guide had been tested earlier during a pilot interview. The participants were instructed to think of a patient with dementia they had taken care of, and then think out loud while completing the translated APS-SE.

The eleven interviews were audio-recorded, and field notes were taken during all interviews. When each interview had finished, the interview was directly transcribed verbatim. After a first readthrough of the transcript of the interview, ambiguous items were identified and revised. The process was repeated between each interview so that each new version of the APS-SE could be directly tested in the next interview. The process continued until nothing essentially new emerged during the interviews.

After five drafts, the final version of the APS-SE was considered ready for use in different settings to assess pain in people with dementia.

Thematic analysis

Once the APS-SE had been finalized, the interviews were also analysed to investigate factors important for clinical implementation. Thematic analysis was used, inspired by Braun and Clark[133]. This is a qualitative research method widely employed to explore and understand patterns of meaning within a text. It offers a systematic and flexible approach to identifying, analysing, and reporting meaningful themes within a dataset[133].

Thematic analysis can be applied to various theoretical frameworks and does not require a specific theoretical orientation. This flexibility means that it is applicable across various research designs, as well as being suitable for both novice and experienced researchers, and so it has a widespread use. However, this flexibility has sometimes been seen as problematic, since the absence of clear and concise guidelines has led to a criticism that 'anything goes'[133]. To address this issue, Braun and Clark published a guideline for thematic analysis in 2006[133].

A theme is a recognized pattern in the data set that captures the essence of the data regarding the overall research question. This is not a question of prevalence, as the theme can be present in just one interview or across the entire data set. Most importantly, themes are not 'discovered', but are a creation of the researchers' active choices: whether to do an inductive or theoretical thematic analysis and whether the analysis should be semantic or latent[133].

In an inductive analysis, the themes are data-driven and strongly linked to the data itself. The research question(s) are formulated during the read through of the data, and continue to evolve during the coding process itself, which makes it possible to code the data without having to make them fit into a pre-existing coding frame. When using a more theoretical thematic approach, the coding is often performed in line with a very specific research question linked to previous research on the topic[133].

Another decision is whether to choose a semantic or latent analysis. In semantic analysis, the analysis does not seek anything beyond what the participants have explicitly said or written. In contrast, latent analysis strives to go beyond the semantic level — the surface level of the content — and to explore and uncover the underlying meanings and assumptions that shape the participants' perspectives[133].

Analysis

After the cultural adaption and translation of the APS-SE was finished, the material from the 11 transcribed interviews was revisited and analysed with an inductive semantic analysis.

To allow refamiliarization with the data, the analysis began by repeatedly reading through the transcriptions while noting initial ideas that emerged from the content. The parts of the dataset regarding clinical implementation of the APS-SE were then coded, and the codes were merged into potential themes. These suggested themes were then checked against the data set and in relation to the codes, until finally naming the different semantically derived themes.

A schematic figure of the analytical process based on the work of Braun and Clark[133] is given in Figure 7.

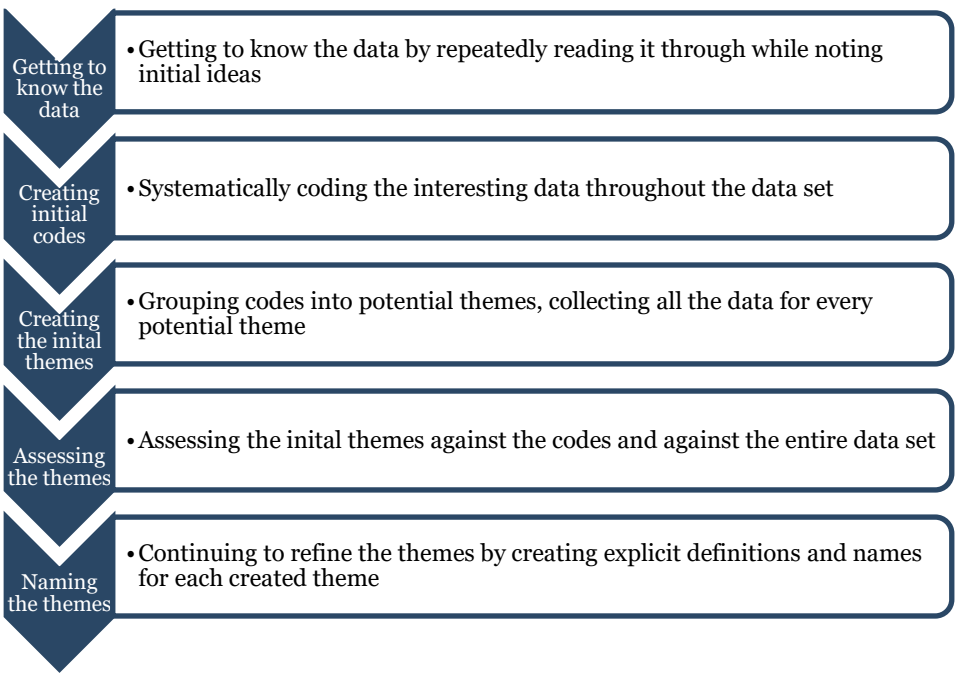


Figure 7. A schematic figure of the analytical process.

Although the summary in Figure 7 suggests a one-way progression, in reality the process involves going back and forth between the different steps before the final result is reached. The analysis resulted in three themes: comprehension, usefulness, and implementation.

Sample size

In the planning phase, it was decided that data collection would continue until no new essential data were found during three consecutive interviews.

Study II

Design

A qualitative research design including semi-structured interviews was used to explore physicians' and nurses' experience of using the APS-SE in people with advanced cancer.

Participants and data collection

HCPs were interviewed about their experiences of using the APS-SE in patients with advanced cancer. When recruiting HCPs, the aim was to achieve variation in occupation (nurses n=6, physicians n=6), age (range 29–63 years), and gender (women n=7, men n=5). All participants had experience of working with patients with advanced cancer near end of life in specialized palliative care units (n=8) or in inpatient oncology hospital wards (n=4). Two participants had less than one year of work experience with patients with advanced cancer, and the rest had six or more years of experience.

The HCPs were recruited through snowball sampling[132], which is one of the sampling methods recommended for qualitative content analysis[134]. All HCPs were approached face-to-face, and all agreed to participate. Any inexperienced users of the APS-SE (n=4) were obliged to use the instrument in at least two different assessments before the actual interview.

The interviews were conducted in Swedish, in the participants' workplace, with no one else present except the participant and the researcher(s). Before the interview, the researcher(s) presented herself/themselves, gave oral and written information concerning the study, and collected oral and written informed consent.

A semi-structured interview guide was used, with questions such as 'What is your experience of using the APS-SE?'. The guide had been previously tested during a pilot interview that was subsequently also included in the analysis. The participants were instructed to think of a patient with advanced cancer they had taken care of, and then think out loud while finishing the APS-SE. They were allowed to choose any patient with advanced cancer, regardless of cognitive status.

Qualitative content analysis

The interviews were analysed with qualitative content analysis[122]. Like thematic analysis, this is a qualitative research method used to explore and understand patterns of meaning within a text. It is also similar to thematic analysis in that it can be used both when there is limited pre-existing knowledge in an area, and when a rich theoretical framework exists[122,135].

Depending on the aim and the available data, either categories or themes are created during analysis. The difference between a category and a theme is the depth of interpretation and degree of abstraction of the data. The higher the degree of abstraction, the further away it is from the original wording used by the participants[122].

Although all analysis includes a degree of interpretation, a more concrete interpretation ends up in a manifest analysis. If interpretation goes beyond what the participants explicitly said or wrote, the analysis is latent. Thus, a category describes the data on a descriptive manifest level, while a theme is on a more interpretative latent level[122].

Analysis

A descriptive manifest qualitative content analysis was performed as described by Graneheim and Lundman (2004)[122]. The twelve interviews were audio-recorded and then transcribed verbatim. Following this, additional contact was made with one of the participants for clarification of some of the answers, but no further contact with the rest was needed.

The analysis started with a repeated reading through of the interviews to recapitulate and summarize the first impression of the content. Next, meaning units (i.e., excerpts answering the aim of the study) were identified, condensed (shortened), and coded. Similar codes were merged into subcategories, and similar subcategories were then sorted into three categories. The analysis was performed using version 12 of the NVivo software package (QSR International).

An example of the analytical process using a quotation from one of the interviews in Study II is given in Figure 8.

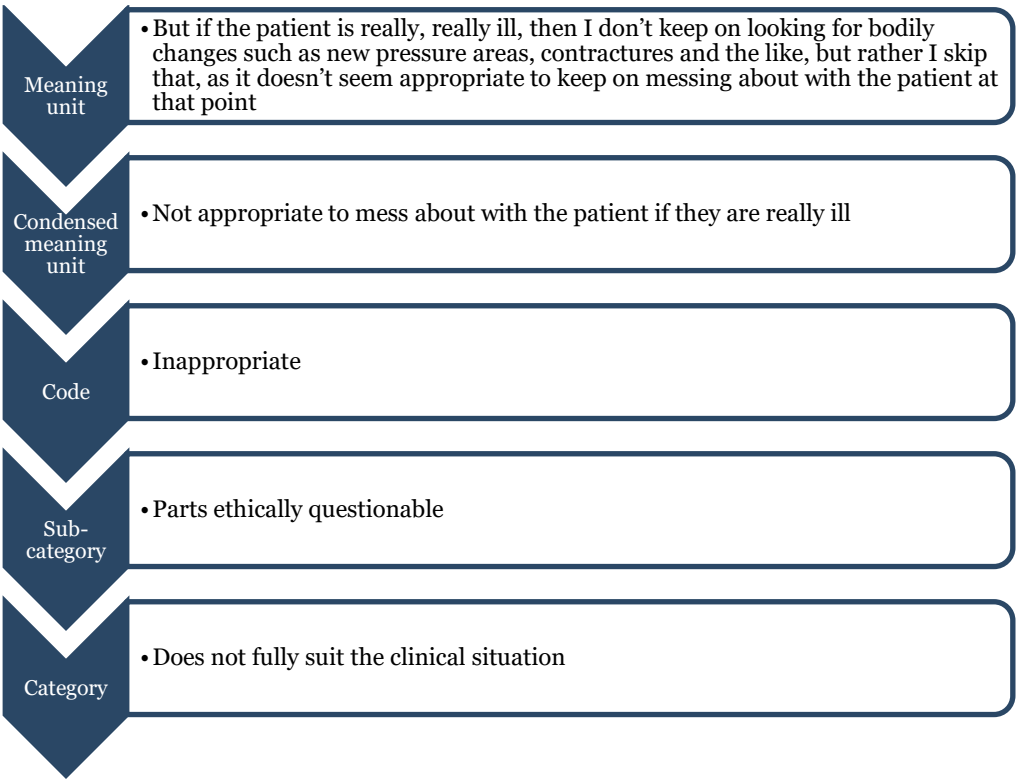


Figure 8. Example of the qualitative content analysis process.

Sample size

In the planning phase, it was decided that data collection would continue until no new essential data were found during three consecutive interviews.

Study III

Design

A quantitative research design with prospective data collected from February 2022 to March 2023 was used to assess the APS-SE regarding validity, reliability, and responsiveness to opioids for patients with advanced cancer in a palliative oncology care setting.

Participants and data collection

Patients with advanced cancer experience many different symptoms, some of which, such as fatigue and delirium, are a hindrance when the patient tries to self-report pain. Not only are fatigue and delirium among the most common symptoms in palliative patients with cancer[11,137,138], but the intensity of these symptoms also often increases when approaching death[137].

Fatigue is a subjective feeling of tiredness, weakness, or lack of energy[139,140]. Its prevalence is around 60% in advanced cancer[141]. Severe fatigue influences the patient's ability to perform daily tasks. The ability to carry out normal activities is often categorized with the help of the WHO or Eastern Cooperative Oncology Group (ECOG) Performance Status[142].

Delirium is a medical condition characterized by changes in awareness, attention, and cognition that fluctuate over time[143]. Its prevalence in cancer patients is 26–67%[144-146]. A systematic review and meta-analysis from 2023 reported eleven risk factors for developing delirium; some were associated with high risk, such as sleep disturbance, cachexia, or infection, while others were more low-risk factors such as old age or opioid use[144]. The screening instrument Single Question in Delirium (SQiD) offers one way to determine delirium in patients with cancer[147,148].

The inclusion and exclusion criteria
All patients over 18 years of age with advanced cancer were eligible if they had least one of the following: poor performance, delirium, drowsiness/unconsciousness, and/or aphasia.

Poor performance was defined as Performance Status 4 according to the WHO or ECOG; that is, the patient is bed bound or totally confined to a bed or chair, cannot carry out any self-care or is completely disabled [142].

Delirium was determined with the help of the SQiD question: 'Do you feel that [patient's name] has been more confused lately?'[147]. If the nurse or physician in charge of the patient answered 'yes' to the SQiD question, the patient was eligible.

Drowsiness or unconsciousness was determined with the Reaction Level Scale (RLS-85)[149], a hierarchically ordered instrument that describes the extent of impaired consciousness via eight 'reaction levels' ranging from 1 (alert, no delay in response) to 8 (unconscious, no response to pain stimulation). Patients were eligible for Study III if they were perceived as at least RLS 2; that is, drowsy or confused but responsive to light verbal or touch stimulation. If the patient was considered unconscious, no additional classification was made to separate the different categories in the RLS instrument.

Aphasia was defined as a previous diagnosis of aphasia according to the patient's chart.

Patients were excluded if they had already been diagnosed with dementia at the time of screening, if they had an intra-dural catheter against pain, or if less than four hours had passed since they were last given an as-needed dose of opioids. When four hours had passed, the patient was eligible again. Screening was made independent of the HCPs' understanding of whether or not the patient suffered from pain.

Screening, recruitment, and consent
Screening and recruitment of patients took place in collaboration with HCPs (nurses, nursing assistants, physicians) in a standalone specialized palliative inpatient care unit and in university inpatient hospital oncology wards. These care units had previously implemented the APS-SE[150]. When a feasible patient was suggested, the chart was reviewed and the physician or nurse in charge was consulted concerning the inclusion and exclusion criteria.

If the patients were able to communicate, they were given oral and written information concerning the study, and written and/or verbal informed consent was collected. They were also asked for an additional oral consent to assess temperature, pulse, and blood pressure as part of the APS-SE assessment. If the patients were unable to give informed consent due to delirium, fatigue, or unconsciousness, they were included without consent as permitted by the Regional Ethical Review Board.

The pain assessment
All the participating HCPs were given a brief overview of the study and the ASP-SE prior to the first assessment of each patient, to address the possibility of their having different amounts of experience in using the instrument.

On the day of inclusion, the APS-SE pain assessment was executed simultaneously, but independently, by two raters. However, only one rater measured pulse and blood pressure, then informed the other rater of the results. The same raters then repeated the assessment a second time for each specific patient, approximately an hour later. The first rater, a researcher, executed the assessments for all patients. The second rater, either a physician, nurse, or researcher, alternated between the patients.

After each assessment with the APS-SE, every communicative patient was asked to self-report their pain using the NRS. A standardized phrase was used: '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 the patient did not answer with a number from 0 to 10, the standardized phrase was repeated, and if the patient still did not give a numerical answer, then they were considered incapable of completing the NRS.

The APS-SE score, along with any request from the patient for pain medication, was then given to the attending nurse (if this nurse was not one of the raters). As part of routine health care, the attending nurse decided if the patient should receive pain-relieving medication or not. Data were collected on whether the patient received pain-relieving medication, and if so, the type of drug, dose, and route of administration. At the end of the study, data were also collected on time of death if the patient was deceased. Data collection was performed between February 2022 and March 2023.

Analysis

The collected data were analysed to assess the APS-SE regarding validity, reliability, and responsiveness to opioids for patients with cancer in a palliative oncology care setting. The statistical analysis was executed using version 28 of the IBM SPSS software package.

Validity testing
A mean APS-SE score was calculated for each patient by averaging the first and second raters' APS-SE scores. The scores were translated to different pain categories using the original APS classification: no pain = 0–2 points,

mild pain = 3–7, moderate pain = 8–13, and severe pain = 14–18[77]. The self-reported NRS was similarly translated to the same pain categories: no pain = 0 points, mild pain = 1–4, moderate pain = 5–6, and severe pain = 7–10[151].

The criterion validity; that is, the association between the APS-SE pain categories and the self-reported NRS categories, was calculated using Cohen's kappa (κ). Criterion validity was also determined for the pain categories 'no pain' versus 'pain'; that is, an APS-SE score of 0–2 versus ≥3 was compared with an NRS score of 0 versus ≥1. The criterion validity was determined at the first pain assessment.

Reliability testing

The inter-rater reliability; that is, the agreement between the two independent observers' APS-SE ratings, was calculated using the intraclass correlation coefficient (ICC) based on a single-rating absolute-agreement two-way random-effects model[121]. The ICC was calculated separately at both the first and second assessments.

The internal consistency reliability; that is, how the separate items in the APS-SE related to each other and the APS-SE score as a whole, was calculated using Cronbach's α[152]. Cronbach's α was also determined for the subgroup who consented to the blood pressure assessments. The internal consistency reliability was calculated at the first assessment.

The test-retest reliability; that is, how the APS-SE score at first and second assessment related to each other for the group of patients who did not receive any opioids between the assessments, was calculated using a single-rating absolute-agreement two-way random-effects model[121].

Responsiveness to opioids

Responsiveness to opioids, reflected in the ability to capture a change in the mean APS-SE score when opioids were given between the first and second assessments, was calculated using the Wilcoxon signed-rank test. As a control, the same test was applied in the group *not* receiving opioids between the assessments. As an additional control, both measurements were also calculated for the NRS scores.

Sample size

A statistician was consulted to determine sample size during planning of the study. However, sample size could not be calculated because the range of the APS-SE scores within the cancer population could not be predicted. After a

literature search for studies validating the APS in patients with dementia and studies validating two other common assessment instruments, the IPOS[153-162] and the ESAS[69,163-179], the sample size was set at 50–100 patients.

Ethics and ethical considerations

Studies I and II

The Regional Ethical Review Board in Umeå, Sweden gave an advisory opinion that ethical approval was not needed for Studies I and II because neither of them gathered any sensitive personal data (Dnr 2017/504-31).

In both studies, the participants received oral and written information concerning the study before giving informed consent. They were informed about voluntary participation and the right to withdraw their consent to participate without giving an explanation at any time. The interviews were conducted with only the interviewee and the researcher(s) present to preserve anonymity. All data from the interviews, including the audio files and transcriptions, were securely stored according to Umeå University's guidelines on data management to preserve confidentiality. The results were purposefully presented in such a way that no participant could be identified.

When interviewing there is an imbalance in power in the relationship between researcher and participant, but the relationship ends after the interview. However, some of the participants in Studies I and II worked, or had worked, with the researcher(s). In this scenario, the relationship between researcher and participant continues even after the interview is over. There is a possibility that the participant might be afraid of expressing themself freely when discussing a controversial subject, or might later regret things said during the interview[180]. Although discussing an already implemented assessment instrument was not perceived as private or very controversial, it is important to be aware that the participants might have a different perspective. Thus, as has been suggested[180], all the participants were asked once more for consent to use the material at the end of their interview.

Study III

Study III received ethical approval from the Regional Ethical Review Board in Umeå, Sweden (Dnr: 2017/504-31). An amendment concerning additional assessment with the NRS at the second pain assessment was approved by the Swedish Ethical Review Authority (Dnr: 2021-06206-02).

The participants received oral and written information concerning the study. They were also informed that if they declined to participate, this would not affect their future care; and that if they agreed to participate, they would be free to withdraw their consent without giving an explanation at any time.

However, some patients were unable to give oral or written consent. This concern had previously been addressed in the research protocol submitted to the Regional Ethical Review Board, and permission was granted to include such patients without consent. If the patient was unable to communicate, any family member present at the time of screening or inclusion was informed about the study and then asked whether they thought the patient would like to participate or not. Proxy decisions do not exist in Swedish legislation, but the family members' beliefs about the patients' opinions on participation were respected.

All collected data were securely stored according to Umeå University's guidelines on data management to preserve confidentiality. The results were presented at group level in such a way that no participant could be identified.

Assessment of pain with the APS-SE includes measurement of pulse, blood pressure, and temperature. To avoid exposing the patients to unwanted interventions, all patients able to communicate were asked for their additional verbal consent to this physical examination, including two blood pressure measurements. If a patient was not able to communicate, the examination was limited to lightly touching the patient to check their temperature.

End-of life patients and dying patients constitute a vulnerable group, and as such they should receive specifically considered protection in the spirit of the Helsinki Declaration[181]. Nevertheless, it is important not to exclude these patients from participating in beneficial research that could result in more competent and efficient care. Furthermore, studies show that patients near the end of life often actively want to participate in research, due to a desire to help themselves and others[182,183]. In this case, the evaluated procedure and instrument were already implemented in everyday care in different health institutions in Sweden, and the information collected in this project could not have been obtained from a non-vulnerable group.

Results

Study I

Translation and adaptation of the APS

Development of the APS-SE for a Swedish care context took place in a series of six steps.

Step 1: Conceptual phase 1
Most of the key concepts and items in the APS were well understood within a Swedish care context. However, the word 'resident' as used in the original version of the APS could not be directly translated into the Swedish care context. In Sweden, patients with dementia are cared for in many different institutions, including nursing homes, in-home care services, hospitals, and standalone specialized palliative care units. These institutions use a diversity of different terms, such as 'patient' ['patient'], 'boende' [resident], or 'gäst' ['guest']. However, during the first conceptual phase the neutral designator 'person' ['person'] was suggested and accepted.

The concept of 'acute on chronic pain' was perceived as unclear and confusing, and no definition was provided in the original study[77]. Moreover, no definition could be found by consulting the literature, including in the International Classification of Diseases-11, which is a classification developed by the WHO as a global standard for diagnostic health information[184]. The concept was altered to 'kronisk och akut smärta' ['chronic and acute pain'] after consulting several colleagues.

Step 2: Forward translation
The two different forward translations were much alike, except for the terms 'fidgeting' (Q3) and 'skin tears' (Q6). Both translators had struggled to find a Swedish equivalent for 'fidgeting', but during mediation the Swedish term 'rör sig nervöst' ['moving anxiously'] was suggested. The professional translator without medical training misinterpreted the medical meaning of the term 'skin tears', but after some discussion, the Swedish term 'rivsår' ['laceration'] was proposed and accepted.

Step 3: Backward translation
The two backward translations were comparable without any major differences.

Step 4: Expert review
The expert group did not find any major conceptual errors when comparing the translations with the original APS, except for the abovementioned terms 'fidgeting' and 'skin tears'. The differences in semantics of the terms between the mediated backward translation and the original APS were recognized as being too far apart to be acceptable. An additional conceptual phase was recommended, and this was added as step 5 in order to rectify the terms.

Step 5: Conceptual phase 2
During the second conceptual phase, a renewed literature review was performed in combination with consulting colleagues and experts concerning 'fidgeting' and 'skin tears'. After dialogue with different HCPs, the term 'fidgeting' was translated as 'rastlös' ['restless']. The more targeted literature review on 'skin tears' identified a study that had previously translated and validated the International Skin Tear Advisory Panel's classification system for skin tears into Swedish. The study established the Swedish word 'hudfliksskada' as the best Swedish translation of the term 'skin tears' after a survey of wound care specialists in Sweden[185].

These new translations of the terms were presented to and accepted by the expert group, resulting in the first Swedish version of the APS-SE.

Step 6: Analysis and adjustments
The analysis of the interviews resulted in three themes: comprehension, usefulness, and implementation.

Comprehension

When adjusting the APS-SE to a Swedish care context, several revisions were necessary in relation to the participants' comprehension of the different items. Confusion or lack of comprehension seemed to mainly occur among participants who did not have Swedish as a first language and/or those who had less or no education in healthcare.

The first attempts at Swedish translations of the term 'whimpering' (Q1), 'kvider' and 'gnyr', created difficulties for some of the participants whose first language was not Swedish. Instead, the Swedish synonym 'jämrar sig' was suggested by a participant who regularly introduced another Swedish version of the APS to HCPs with native languages other than Swedish. The suggested word was recognized by the participants, irrespective of mother tongue.

The term ‘fidgeting’ (Q3) continued to require alteration, as the word ‘rastlös’ [restless] was interpreted primarily as inward restlessness. Other examples suggested by the participants, the Swedish words ‘plockar’ and ‘fipplar’, were recognized mainly by the native Swedish speakers, but not always by those with other native languages. Comprehension of the word ‘fidgeting’ did not increase until the term ‘move restlessly’ was used with additional explanatory examples; ‘rör sig oroligt (t.ex. pillar eller skakar benet)’ [‘move restlessly (e.g., to fiddle or shake one’s leg)’].

Level of education had an impact on comprehension. Participants with more extensive education had fewer linguistic difficulties, regardless of native language, and a better understanding of the medical terminology. The medical term ‘contractures’ was understood by all the physicians and nurses, but not by the nursing assistants, and so a supplementary explanation was added: ‘kan inte böja eller sträcka ut en led (kontraktur)’ [‘cannot bend or straighten a joint (contracture)’]. Although the participants were generally unfamiliar with the validated Swedish ‘hudfliksskada’, this word was also better recognized by participants with more extensive education.

The interviews revealed that the term ‘withdrawn’ (Q3) was interpreted as both a mental and/or a physical manifestation. This word has a dual meaning in Swedish as well as in English. The participants noticed both types of withdrawal among the patients with pain, and as this was also confirmed by experienced specialists in geriatrics outside the expert group, both interpretations were found acceptable.

During the interviews the participants were asked if any question or item in the APS-SE could be perceived as offensive or objectionable. The term ‘refusing to eat’ was initially translated as ‘matvägran’, but this translation was believed to have a negative connotation, as it was associated with the stubbornness or fussiness of a toddler rather than the behaviour of an adult. When changed to ‘vill inte äta’ [‘does not want to eat’], the term was perceived as neutral. Nothing else in the APS-SE was considered inappropriate.

Usefulness

In general, the HCPs considered the instrument to be clear and simple. It was perceived as straightforward to use, especially the items on vocalization (Q1) and facial expression (Q2), along with changes in body language (Q3) and behaviour (Q4). The last item, physical changes (Q6), reminded the HCPs to be aware of the patient's previous diseases and injuries when doing the pain assessment.

However, the utility of the instrument relies on prior knowledge of the patient, since the assessment in items Q3–Q6 depends on a potential change in the patient. This made the instrument particularly difficult to use for the HCPs working in the emergency department, who had usually never met the patient before. These HCPs generally had to rely on the accompanying family members and/or HCPs when making the assessment.

The Swedish terms 'rodnad' ['flushing'] and 'blekhet' ['pallor'] in Q5 were understood regardless of health care profession. Nevertheless, many of the HCPs found it difficult to apply the items in patients with darker skin tones. Without education, knowledge, and experience in assessing darker skin tones, they especially perceived their estimation of 'pallor' as dubious. The only participant expressing confidence in distinguishing pallor in dark skin tones had substantial experience of working in Africa. The item 'flushing' was considered somewhat easier to detect, since it was sometimes associated with elevated skin temperature.

Implementation

Several of the participants explicitly discussed the need for a meticulous introduction of the instrument to all new HCPs. Without the possibility of assessing each patient themselves, nurses and physicians had a great need for a well-functioning pain assessment instrument that could be used by the nursing assistants. There was concern that some HCPs did not understand and speak enough Swedish to grasp the different items in the APS. Thus, a thorough introduction of the instrument was needed, especially if the new HCP had less or no education in health care and/or did not have Swedish as a first language.

The APS-SE

After several adjustments, the final APS-SE was created (Figure 9). The original APS is shown in Figure 10.

Abbey Pain Scale-SE

För skattning av smärta hos personer med demens som inte kan uttrycka sig i ord

Använd skalan så här: *Skatta fråga 1 till 6 medan du observerar personen.*

Personens namn:_____

Namn och yrke på den som genomför skattningen:_____

Datum:_____ Tid:_____

Senast givna smärtlindring:_____ Klockan:_____

F1. Röstuttryck
t.ex. jämrar sig, stönar, gråter
Inget 0 Lite 1 Måttligt 2 Mycket 3 ☐

F2. Ansiktsuttryck
t.ex. ser spänd ut, rynkar pannan, grimaserar, ser rädd ut
Inget 0 Lite 1 Måttligt 2 Mycket 3 ☐

F3. Förändrat kroppsspråk
t.ex. rör sig oroligt (t.ex. pillar eller skakar benet), vaggar, skyddar någon kroppsdel, drar sig undan
Inget 0 Lite 1 Måttligt 2 Mycket 3 ☐

F4. Förändrat beteende
t.ex. ökad förvirring, vill inte äta, ändrat beteendemönster
Inget 0 Lite 1 Måttligt 2 Mycket 3 ☐

F5. Fysiologiska förändringar
t.ex. kroppstemperatur, puls eller blodtryck som avviker från det normala, svettning, rodnad eller blekhet
Inga 0 Lite 1 Måttliga 2 Mycket 3 ☐

F6. Kroppsliga förändringar
t.ex. hudfliksskador*, ledsmärta, kan inte böja eller sträcka ut en led (kontraktur), trycksår, tidigare skador
Inga 0 Lite 1 Måttliga 2 Mycket 3 ☐

Räkna ihop poängen för fråga 1–6 och för in summan här ⟶ Totalt antal smärtpoäng ☐

Kryssa i totalt antal smärtpoäng ➡

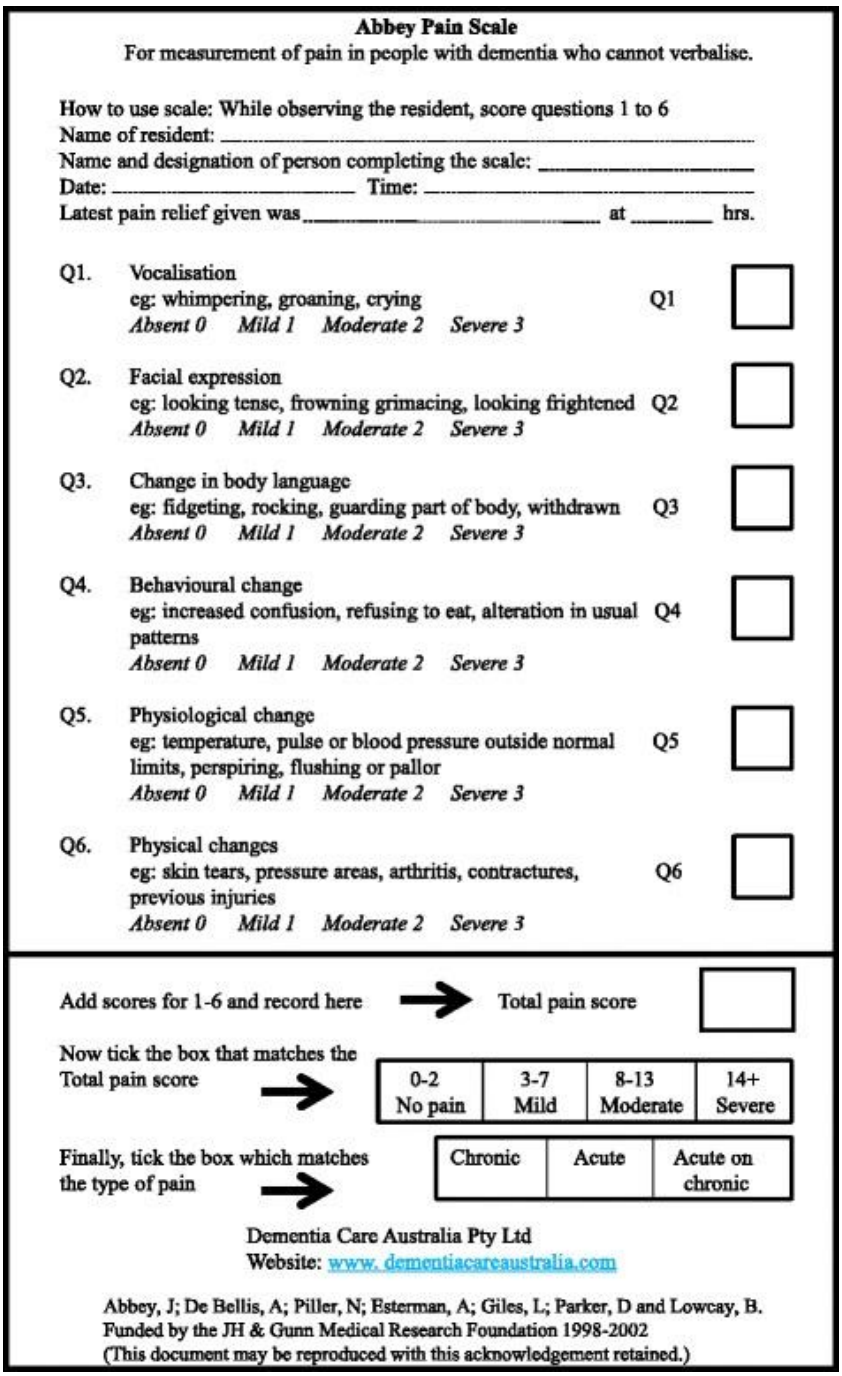
0–2	3–7	8–13	14+
Ingen smärta	Lindrig smärta	Måttlig smärta	Svår smärta

Bedöm och kryssa i typ av smärta ➡

Kronisk smärta	Akut smärta	Kronisk och akut smärta

* Källman U, Kimberly LB, Bååth C. Swedish translation and validation of the international skin tear advisory panel skin tear classification system. Int Wound J. 2019 Feb;16(1):13-18.

Figure 9. This Swedish version of the Abbey Pain Scale: the APS-SE.



Abbey Pain Scale

For measurement of pain in people with dementia who cannot verbalise.

How to use scale: While observing the resident, score questions 1 to 6

Name of resident: _____

Name and designation of person completing the scale: _____

Date: _____ Time: _____

Latest pain relief given was _____ at _____ hrs.

Q1. Vocalisation
eg: whimpering, groaning, crying Q1
Absent 0 Mild 1 Moderate 2 Severe 3

Q2. Facial expression
eg: looking tense, frowning grimacing, looking frightened Q2
Absent 0 Mild 1 Moderate 2 Severe 3

Q3. Change in body language
eg: fidgeting, rocking, guarding part of body, withdrawn Q3
Absent 0 Mild 1 Moderate 2 Severe 3

Q4. Behavioural change
eg: increased confusion, refusing to eat, alteration in usual patterns Q4
Absent 0 Mild 1 Moderate 2 Severe 3

Q5. Physiological change
eg: temperature, pulse or blood pressure outside normal limits, perspiring, flushing or pallor Q5
Absent 0 Mild 1 Moderate 2 Severe 3

Q6. Physical changes
eg: skin tears, pressure areas, arthritis, contractures, previous injuries Q6
Absent 0 Mild 1 Moderate 2 Severe 3

Add scores for 1-6 and record here → Total pain score

Now tick the box that matches the Total pain score →

0-2 No pain	3-7 Mild	8-13 Moderate	14+ Severe

Finally, tick the box which matches the type of pain →

Chronic	Acute	Acute on chronic

Dementia Care Australia Pty Ltd
Website: www. dementiacareaustralia.com

Abbey, J; De Bellis, A; Piller, N; Esterman, A; Giles, L; Parker, D and Lowcay, B.
Funded by the JH & Gunn Medical Research Foundation 1998-2002
(This document may be reproduced with this acknowledgement retained.)

Figure 10. The original Abbey Pain Scale.

Study II

The qualitative content analysis resulted in nine subcategories grouped into three categories: fills a need, not always on target, and does not fully suit the clinical situation (Figure 11).

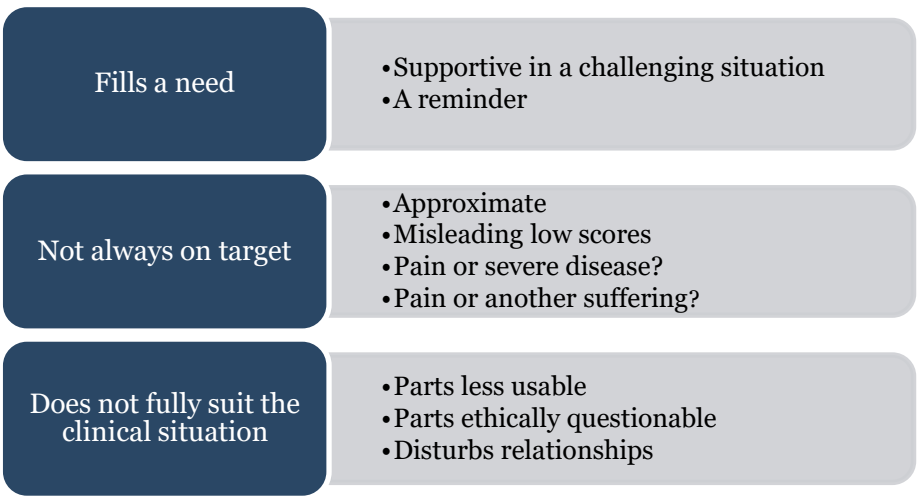
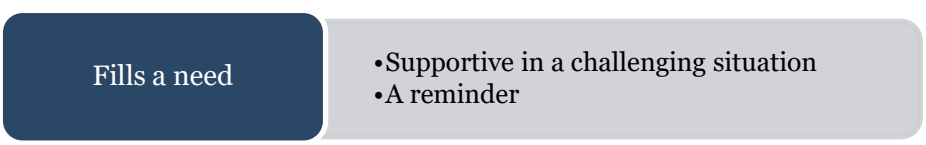


Figure 11. The categories and subcategories from the qualitative content analysis.

In general, HCPs with different professions did not differ in their experience of using the APS-SE in patients with advanced cancer. The only noticeable distinction was the physicians' emphasis on using the instrument systematically to assess pain both pre- *and* post-intervention.

Fills a need



Supportive in a challenging situation
Pain assessment in a palliative setting could be both demanding and clinically challenging, according to the HCPs. The APS-SE and their own intuitive estimation of the patient's pain intensity supplemented each other when they were gathering information to make a pain assessment.

The HCPs reported that although the APS-SE was mainly used in end-of-life situations, it was also applied when patients had difficulties expressing themselves verbally. There could be several reasons for these communication difficulties, such as delirium, severe fatigue, or drowsiness.

The APS-SE could also be used when the HCPs noticed a repeated dissimilarity between their own understanding of the patient's pain and the patient's self-reported pain. This disparity was mainly attributed to the patient's mild cognitive impairment, but occasionally to an addiction to pain-relieving medications.

If the patient's self-reported pain was consistently lower or higher than conceived by the HCPs, the APS-SE was applied as a 'rough guide'. This could help the HCPs to interpret the patient's self-reported NRS and to decide whether the self-reported pain intensity level seemed reasonable.

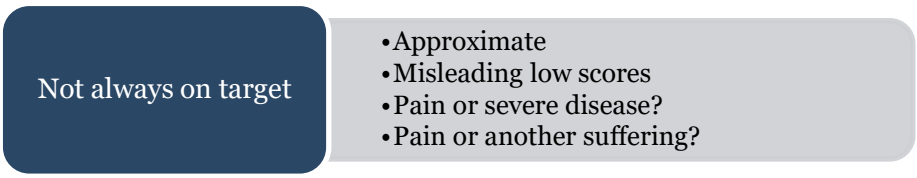
Some HCPs who worked in a clinically demanding palliative setting expressed that just knowing the APS-SE was available brought some emotional comfort and support in a potentially distressing situation.

A reminder

The APS-SE gave a structure to the pain assessments that was welcomed by HCPs with both longer and shorter work experience. It acted as a checklist that showed the inexperienced HCPs what to examine, while helping the more experienced HCPs not to forget anything.

However, when applying the APS-SE in patients with cancer, some of the proposed items such as 'fidgeting' or 'rocking' were not very useful, since the patients seldom expressed this behaviour. According to the HCPs, phenomena such as contractures were almost never seen, and although the APS-SE specifies that these are just examples of physical changes, the HCPs did not mentally replace the inappropriate examples with more suitable ones. This made the list less efficient as a reminder.

Not always on target



Approximate
The assessment for each question in the APS-SE is summarized in a total pain score. This numerical expression of pain was described as ambiguous, since it was expressed as a precise score while some HCPs had a feeling of the assessment being imprecise.

Inexperienced HCPs sometimes felt distressed by the thought that if they were more experienced, they might have rated a patient's pain differently and perhaps more accurately. Regardless of work experience, several of the HCPs commented on the difficulties of rating some of the items; for example, it was not clear how to differentiate between mild, moderate, and severe grimacing. This uncertainty regarding how to rate led to an understanding of the assessment as being approximate.

Misleading low scores
A recurring problem was the feeling of disparity between the resulting APS-SE pain scores and the HCPs' own intuitive perception of the patients' pain. The impression was that the APS-SE score often ended up indicating too low pain intensity levels.

Even if some of the HCPs persistently rated the higher score when they were uncertain which rating to choose, they nonetheless worried that the APS-SE score did not correspond to the patient's actual pain intensity. To avoid this disparity, some HCPs made their own interpretation of the pain scores. For example, if the instrument showed a total pain score of 7, this was believed to represent severe pain, and not mild as dictated by the APS-SE.

When the HCPs' 'gut feeling' differed from the APS-SE score, they usually administered pain-relieving medication even if the total pain score suggested mild or no pain.

Pain or severe disease?
Applying the APS-SE in patients with cancer created a dilemma over how to score physiological changes such as elevated temperature or perspiring. These changes were often interpreted as a manifestation of severe disease

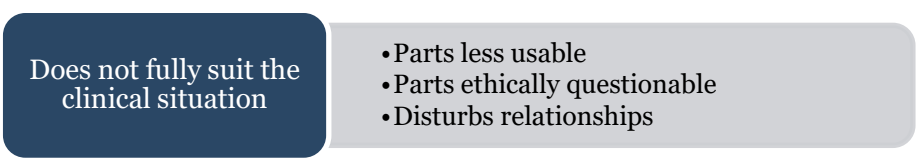
rather than pain, and thus the HCPs were confused about whether to score these changes or not.

According to the HCPs, patients with cancer in end-of-life were often fatigued and bed bound with minimal body language, making the score in the item 'change in body language' too low. Conversely, 'behavioural change' scored misleadingly high since many patients with cancer suffered from confusion and/or gradually lowered their intake of food.

These sometimes subtle changes in body language and behaviour as an expression of severe disease made it difficult when scoring. If they had no prior familiarity with the patient, the HCPs were compelled to ask family members for guidance to assure accurate scoring.

Pain or another suffering?
There was an understanding that a high APS-SE score was an expression of suffering, but not inevitably an expression of pain. End-of-life patients with cancer often display a complex symptomatology, making a high APS-SE score difficult to decipher. Many HCPs perceived that anxiety itself or in combination with pain, along with other symptoms such as breathing difficulties, might also generate high scores.

Does not fully suit the clinical situation



Parts less usable
While the items 'vocalization' and 'facial expression' were considered fundamental when assessing pain, the items 'physical changes' and 'physiological change' were often disregarded since they were considered less useful in this palliative setting. 'Physical changes' was arguably more of a diagnostic item, and the HCPs did not feel that any valuable information was gained by measuring blood pressure as part of deciding 'physiological change' in this end-of-life context.

Parts ethically questionable
Being aware of their patients' vulnerability near the end of life, the HCPs highlighted the importance of minimizing any procedures or examinations that could potentially be painful. The only exceptions were those that were absolutely necessary for symptom control.

Given this situation, the HCPs felt that some of the suggested examinations, such as measuring blood pressure or looking for skin tears, were ethically questionable, since these items were not considered useful to assess pain in a late palliative phase. There were ethical concerns that these procedures might be too intrusive in this setting, and this understanding also made the HCPs reluctant to use some of the items in the APS-SE.

Disturbs relationships
In a few cases, the APS-SE was seen as a barrier between the HCPs and the patient and/or family members. In a sensitive situation that requires trust, using a formal instrument in an end-of-life setting was seen as counterproductive instead of helpful. Not only did the use of an instrument trigger feelings of not really giving comfort within the HCPs themselves, it also made them fear that they seemed detached and lacking empathy towards the dying patient or the family members.

Study III

The participants
From February 2022 to March 2023, 72 patients were included, comprising 41 (57%) women and 31 (43%) men with a mean age of 76 years (range: 47–95) and various types of cancer (Table 5). Six additional patients were considered eligible during screening, but five of them died before inclusion, and the remaining case was not included because the family members present at the time of inclusion were hesitant.

Table 5. Types of cancer in the patients (n=72) included in Study III.

Type of cancer	**Number of patients** n (%)
Gastrointestinal	23 (32)
Urologic	9 (13)
Gynaecological	9 (13)
Lung	6 (8)
Breast	6 (8)
Central nervous system	6 (8)
Haematological, including lymphoma	5 (7)
Head and neck	5 (7)
Cancer of unknown primary (CUP)	3 (4)

The most common inclusion criterion was poor performance status (n=36, 50%), followed by a combination of poor performance status and delirium (n=15, 21%). In 14 (19%) cases the patient was either unconscious or drowsy, while in 6 (8%) cases the patient was suffering from delirium. One patient (1%) was diagnosed with aphasia. Four APS-SE assessments were performed per patient, giving a total of 288 assessments.

As poor performance was defined as being totally confined to bed or chair, 268 (93%) of the assessments were made while the patient was bedridden, and the rest (n=20, 7%) while the patient was sitting in a chair. The median time between assessment and death was 11 days (range: 0.5–201). Almost all patients had died at the time of analysis (n=70, 97%). The APS-SE assessments were performed by nurses (n=32), physicians (n=6), and researchers (n=2).

More than half of the patients (n=41, 57%) were able to complete two self-reported pain assessments with the NRS. An additional 4 (6%) managed to complete the NRS once, while the rest (n=27, 37%) were not able to use the NRS at all. When asked for additional verbal consent for measuring blood pressure, all of the communicative patients accepted, and consent was given by 56 (78%) patients. The rest of the patients (n=16, 22%) were not capable of expressing their opinion.

A majority of patients (n=51, 71%) received ongoing regular medication with opioids, while the rest (n=21, 29%) had opioids prescribed only as needed, with or without regular non-steroidal anti-inflammatory drugs (NSAIDs) and/or paracetamol. For the patients with ongoing opioids, the median equivalent morphine dose was 80 mg p.o. per day (range: 10–935). Complementary drugs against neuropathic pain, pregabalin or gabapentin, were prescribed to 11 (15%) of the patients. Approximately one-third of the patients (n=26, 36%) received opioids between their first and second assessment, with the median equivalent dose being 12.5 mg morphine p.o. (range: 10–108).

The data were analysed for validity, reliability, and responsiveness. An overview of the results is given in Table 6.

Table 6. Overview of the psychometric testing.

Test	Result	Interpretation
Criterion validity: *n=45* APS-SE vs. NRS divided into the pain categories: no, mild, moderate, severe pain	κ=0.08 (95% CI: -0.06 to 0.22)	Slight correlation
Criterion validity: *n=45* APS-SE vs. NRS divided into the pain categories: no pain, pain	κ=0.04 (95% CI: -0.19 to 0.26)	Slight correlation
Criterion validity without patients with delirium: *n=38* APS-SE vs. NRS divided into the pain categories: no, mild, moderate, severe pain	κ=0.15 (95% CI: 0.01 to 0.31)	Slight correlation
Inter-rater reliability: *n=72* Calculated at first assessment	ICC=0.64 (95% CI: 0.43 to 0.78)	Moderate correlation, ranging from poor to good
Inter-rater reliability: *n=72* Calculated at second assessment	ICC=0.61 (95% CI: 0.38 to 0.76)	Moderate correlation, ranging from poor to good

Table 6. *Continued*.

Test	Result	Interpretation
Internal consistency: *n=72* All patients	α=0.01	No correlation
Internal consistency: *n=56* Only patients with blood pressure assessments	α=0.05	No correlation
Test-retest: *n=46* Patients who did not receive opioids between assessments	ICC=0.82 (95% CI: 0.67-0.90)	Good short-term stability, ranging from moderate to good
Responsiveness to opioids: *n=26* Patients who received opioids between assessments	p=0.01	Statistically significant change
Responsiveness to opioids: *n=46* Patients who did not receive opioids between assessments, i.e., control group	p=0.65	No statistically significant change

κ=Cohen's kappa
CI=confidence interval
ICC=intraclass correlation coefficient
α=Cronbach's α

Validity

Criterion validity was calculated by comparing the pain categories reported by the HCPs using the APS-SE with the patient's self-report using the NRS at the first assessment. Among the 45 patients who were able to use the NRS, 38 were considered to be without delirium at the time of inclusion. Cohen's kappa was ≤0.20 both in the group of all 45 patients (κ=0.08) and among the patients without delirium (κ=0.15), which is considered a slight association[186]. When the APS-SE and NRS pain scores were categorized into 'pain' versus 'no pain', the association was still slight (κ=0.04).

None of the 288 assessments made with the APS-SE exceeded the pain category 'mild pain'. The APS-SE assessments were equally distributed between the categories 'no pain' (n=144, 50%) and 'mild pain' (n=144, 50%).

In contrast, the assessments made with the NRS ranged over all pain categories from 'no pain' to 'severe pain'. The distribution of pain categories for both the APS-SE and the NRS in the 45 patients who were able to use the NRS at their first assessment are given in Figure 12.

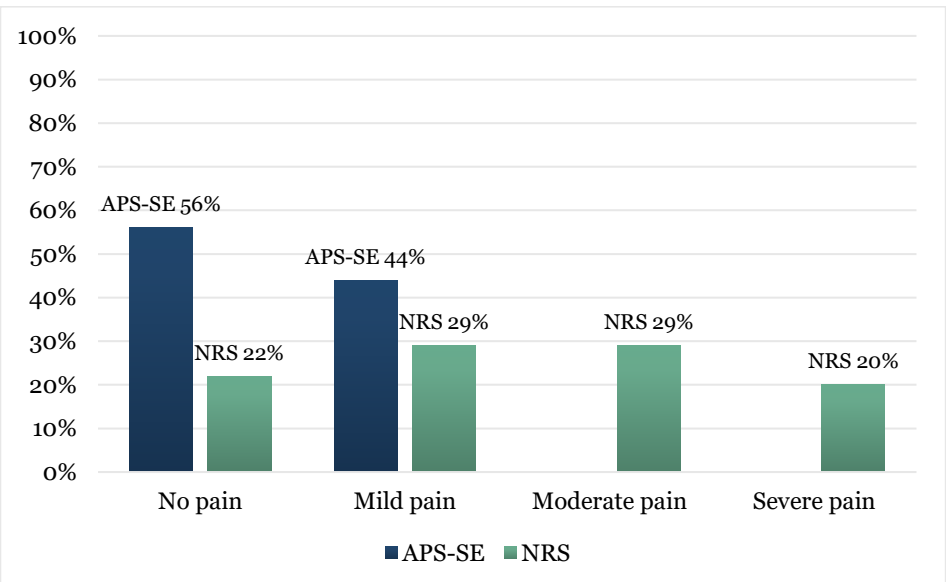


Figure 12. Percentage distribution of pain categories for this Swedish version of the Abbey Pain Scale (APS-SE) and the numeric rating scale (NRS) in 45 patients.

In descending order, the items contributing to the total mean APS-SE score of 2.35 points at the first assessment were: facial expression 0.66, physical change 0.48, behavioural change 0.42, physiological change 0.38, vocalization 0.28, and change in body language 0.13. The total mean APS-SE score for the patients who self-reported any pain (i.e., NRS ≥1) was almost identical (2.34), as was the score in the patients reporting NRS ≥5 (2.33).

Reliability

The ICC for inter-rater reliability at both the first assessment (ICC 0.64) and the second (0.61) was in the interval of 0.50–0.75 that is interpreted as moderate, while the confidence interval ranged from poor to good[121]. The ICC for test-retest among the patients who did not receive opioids (n=46) showed a high consistency between the APS-SE scores assessed an hour apart, with an ICC of 0.82 and a confidence interval ranging from moderate to good[121].

The internal consistency reliability as determined with Cronbach's α was <0.50 both when calculated for all patients (α=0.01) and when calculated only for those (n=56) with complete blood pressure measurements (α=0.05), which is to be interpreted as indicating unacceptable internal consistency[152].

To investigate whether removing an item would make the internal consistency higher, thus indicating that the item did not contribute to the overall pain assessment, Cronbach's α was also calculated with each item separately deleted. However, the α for both groups remained at <0.50, ranging from -0.22 to 0.28 among all patients and from -0.19 to 0.32 among patients with blood pressure measurements.

Responsiveness

The APS-SE responsiveness for opioids showed a statistically significant decrease of 0.71 points in the mean APS-SE score among the 26 patients who received opioids, from 2.96 points before the opioids to 2.25 points after ($p \leq 0.05$).

As a control, responsiveness for the 15 patients who received opioids and could self-report NRS at both assessments were also calculated. The mean NRS score was 5.00 points before and 3.33 points after the opioids ($p \leq 0.01$). For the patients who did not receive opioids, there was no statistically significant decrease in either the APS-SE or the NRS score.

Discussion

The overall aim of this thesis was to evaluate observational pain assessment for people with advanced cancer in a palliative care setting, focusing on the Abbey Pain Scale (APS). The APS is an observational pain assessment instrument that was developed for patients with end-stage dementia in nursing homes, but is also used to assess pain in patients with advanced cancer in Sweden. This thesis shows, using both qualitative and quantitative research methods, that a Swedish version of the APS (APS-SE) is not adequate to use in a palliative oncology care setting.

The APS-SE in patients with advanced cancer

Pain is both a highly subjective and a deeply personal sensation[187], and as such should preferably be self-reported whenever assessed[15,21,22]. Nevertheless, attempts have been made to objectively measure pain, for example through the use of biomarkers in saliva. An article published in 2021 reported a high correlation between salivary biomarkers and pain according to an observational pain assessment instrument in patients with advanced dementia[188]. An exploratory study from 2022 showed attenuated salivary cortisol patterns in cancer patients with chronic chemotherapy-induced peripheral neuropathy compared to healthy adults[189]. It remains to be seen whether measurement of biomarkers is the way to go, or even if it is clinically doable, but these results do suggest that other options for assessing pain more objectively may be available in the future. Nevertheless, until then, the use of observational pain assessment instruments will continue to be the imperative option for patients who cannot vocalize their pain.

Many different assessment instruments are used in end-of-life palliative care. A European survey found that 116 different instruments were used in clinical practice, but none of them were observational pain assessment instruments for patients with advanced cancer[190]. A lack of observational pain assessment instruments for this population means that HCPs are obliged to use observational instruments developed for other populations instead. Some of these instruments, such as the APS, were originally developed for patients with dementia.

There is a need for observational pain assessment instruments in palliative care. This need was confirmed in Study II when investigating HCPs' experiences of using a Swedish version of the APS, the APS-SE, in patients with advanced cancer. The instrument was seen as a complement to the HCPs' own pain assessment, as well as an emotional support in a clinically challenging situation. However, although the APS-SE did fill a need for

support, it was considered to be merely 'OK' when used in patients with advanced cancer. The HCPs continued to use it mainly since it was better than nothing, and no better option was available.

Psychometric data are essential in evaluating an instrument. The literature contains very little information on use of the APS in patients with cancer, with the exception of one case study. In that case study, the APS was successfully used to assess pain in a patient suffering from colon cancer and communication difficulties due to alcoholic encephalopathy[191]. Since no other literature on the subject is available, all comparison of psychometric data for the APS-SE in patients with cancer has to be made against its original population; that is, patients with dementia.

The relationship between the APS and other instruments

Criterion validity, which describes how well an instrument correlates with an established standard of comparison, was determined for the APS-SE in Study III. The NRS was chosen as the criterion for comparison, since it is considered to be the gold standard for pain assessment in patients with cancer[15,21,22]. Criterion validity was determined in the 45 patients who were able to self-report pain, and the result showed only slight agreement[186].

This slight criterion validity confirms the experiences of the HCPs in Study II, that the APS-SE produced misleadingly low scores. The HCPs expressed that although they intuitively felt that the patient was suffering from severe pain, the APS-SE only showed mild or moderate pain. A gap between the estimated pain intensity and the APS score was also noted when testing the clinical utility of the Japanese APS in a dementia population[192].

It is difficult to compare criterion validity within the studies that have investigated the APS in dementia populations, since different studies used different gold standards. However, regardless of the chosen gold standard, criterion validity was consistently higher in patients with dementia than in patients with cancer[77,103-106]. When nurses' overall holistic pain assessment was used as the gold standard, the correlation was deemed reasonable[77]; when observer estimation of pain was used as the gold standard, the correlation was deemed good[105]. When the self-reported verbal rating scale was used as gold standard, the correlation was moderate[103,104].

In a study comparing the observational instruments APS, DOLOPLUS-2, and CNPI in the same patients with dementia, the APS generally reported a lower level of pain than the other instruments[107]. Even so, the mean APS

score among the dementia patients in the original APS development study[77] was still almost four times as high as the mean score among the cancer patients in Study III (9.02 vs. 2.35 points, respectively). This suggests that the APS-SE is too insensitive when assessing pain in patients with cancer.

Moreover, this insensitivity was made apparent by the fact that the APS-SE did not report moderate or severe pain *even once* during 288 assessments in 72 patients, while almost half of the self-reporting patients in Study III indicated a need for better pain management by scoring moderate or severe pain in the NRS.

Do different HCPs get the same pain score when using the APS-SE?

The inter-rater reliability, which describes how well two or more independent raters agree when measuring the same thing, was moderate[121] in Study III at both first and second assessment when calculated for 40 different raters. However, the confidence interval ranged from poor to good[121] at both assessments, as further discussed below in the section on methodological considerations. In the previously mentioned studies examining the APS in dementia populations, the inter-rater reliability was considered good when using two raters[103,104], and moderate to borderline good in studies with more than 20 raters[107,192].

When asked about their experiences of using the APS-SE in patients with advanced cancer, the HCPs in Study II expressed concerns regarding how to rate the different items; for example, they felt it was not clear how to decide whether a patient's grimacing was moderate or severe. This gave them a feeling of the APS-SE scores as being approximate. The less experienced HCPs also wondered if their scores differed from other raters, but their more experienced counterparts did not consider this a problem, thinking that different raters would still end up in the same APS-SE pain category due to the categories' wide interval. Although only nurses and physicians were interviewed in Study II, previous research has found no apparent differences in APS scoring between nurses and nursing assistants when tested in patients with dementia[107].

Do the items measure the same concept, and is it stable?

The test-retest reliability, which describes how stable the scoring is over time, was good[121], in contrast to other studies which showed only moderate results when using the APS in patients with dementia[104,107]. This discrepancy is not surprising, since the test and retest were

administered one hour apart in the cancer population, while the retest in the dementia studies was sometimes administered up to 14 days after the baseline assessment[104,107].

Conversely, the internal consistency reliability, which assesses how well the different items in an instrument relate to each other, was considered unacceptable[152]. This result from Study III is in line with the qualitative findings in in Study II, where the HCPs expressed that they did not consider all parts of the ASP-SE to be usable. Both physiological and physical changes were often ignored by these HCPs.

Previous studies have shown a higher internal consistency in a dementia population, ranging from questionable to acceptable[77,103-105,107]. Even so, some of the APS items contributed less in the dementia population: behavioural change[103,108], physical changes[103,107], and physiological change[103,104,108]. When the psychometric qualities of the APS were investigated in an entirely different population, elderly patients with osteoarthritic pain, removal of the item 'physiological change' was recommended[109].

According to the original APS development study, the APS was intended to be applied whenever a patient seems to be suffering from pain[77]. However, a screening instrument designed for detecting and identifying pain regardless of any preunderstanding would probably be of more clinical value. The APS-SE appears to be unsuitable for use as a screening instrument due to its lack of sensitivity in patients with advanced cancer. This insensitivity poses a risk of failing to identify the patient's pain. Furthermore, the instrument seems incapable of detecting moderate or severe pain in this population. Without an understanding of the severity of the pain, it becomes challenging to ascertain whether the patient requires pain-relieving medication, and if so, what type of medication would be appropriate based on the intensity of the pain.

If the APS-SE cannot be used as a screening instrument, it could still potentially be useful as a reassessment instrument. The use of the APS-SE as a structured reassessment instrument after intervention was the only noted difference between nurses and physicians in Study II, as it was almost exclusively mentioned by the physicians.

Can the APS-SE be used for reassessment after an intervention?

The APS seems to show responsiveness in a dementia context, as its scores decrease after an intervention[77,103,105]. However, the clinical use of the APS-SE as a reassessment instrument in a palliative cancer context is not as clear. The responsiveness to opioids in Study III was statistically significant in the 26 patients who received opioids, but the actual score difference was 0.71 points. No literature has been found describing the level of a clinically meaningful change in APS score after any intervention, but for the NRS it is believed to be at least a two-point change[193,194].

With a mean APS-SE score of 2.34 in the patients who self-reported pain, a decrease of 0.71 points is equivalent to a 30% reduction. When defining a clinically significant of change in quality-of-life scores in health care, a change over 20% is classified as 'very much change'[195]. Nonetheless, for a specific patient's response to an intervention to be noticeable, the APS-SE score must decrease by at least one point. The clinical utility of the APS-SE as a reassessment instrument is therefore questionable, as it may fail to detect a reduction in pain.

Does the APS-SE differentiate between pain and other symptoms?

One fundamental question is whether the APS-SE is actually measuring the concept 'pain' in patients with advanced cancer, or whether it might be measuring something else. The participants in Study II agreed that a high APS-SE score indicated suffering, but it is not clear that the suffering was caused by pain rather than, for example, anxiety. This uncertainty has been recognized by others when using the APS in patients with cognitive impairment[108].

However, perhaps this uncertainty is something that cannot be solved. It is not always possible for the patients themselves to distinguish between different symptoms such as pain, anxiety, or a combination of both[196]. If a cognitively intact patient has trouble differentiating between different symptoms, perhaps this is an uncertainty all HCPs have to live with, regardless of whether a self-report or an observational pain assessment instrument is used.

Overall, the psychometric data show that the APS-SE is inadequate when used in patients with advanced cancer, even when used as meticulously as in Study III. In reality, the gap between the patient's self-reported pain and the APS-SE score might be even greater in the clinical context, as Study II

revealed that the APS-SE was not always used as intended. This finding, that the APS-SE was not being used as proposed, suggests problems with the implementation.

How to implement an instrument

When looking at facilitators and barriers for implementing a new measurement in a clinical palliative care practice, the timeline of the implementation can be divided into three different stages: the preparation, the implementation, and the assessment and improvement[197].

The first step during the preparation stage is to identify if there is a need for improvement within the clinical palliative care setting. This step is essential not only to get the needed support for implementing a new instrument, but also to sustain the change after the implementation[198]. During preparation, it is important that everyone involved becomes comfortable with the upcoming implementation of the chosen instrument. However, several barriers may exist, both on an organizational and an individual level[197].

The organization may have concerns regarding the costs of the necessary education before implementing the instrument, or any costs involved in using the instrument. There can also be a resistance on both the organizational and/or the individual level based on the fear of inadequate time management due to additional work and insufficient staff time[197].

On the individual level, there can be the fear of not receiving enough education on how to use the instrument, but the resentment can also be on a more personal level, with fear of any change. There may also be a belief that the instrument could become too burdensome for the patient or the HCPs[197].

It is important to address the incentives for implementing the instrument in the specific setting while also maintaining a realistic expectation of the perceived benefits[197]. Both time and resources are needed for proper education and training[198]. HCPs who will not directly use the instrument may also need to be informed, since they might be asked by patients and family members to explain any related changes in the clinical routine[198].

A decision must also be made on *how* to educate the HCPs on the application, interpretation, and documentation of the instrument. A thorough understanding will promote ownership of the instrument among those who will use it. The key factor to accomplish this seems to be the

appointment of a coordinator or facilitator[197], with support from identified 'unofficial leaders' within the setting[198].

During the preparation stage it is likewise important to contemplate *who* is going to administer the chosen instrument. Many of the nursing assistants in Sweden lack proficiency in the Swedish language[199] and/or do not have an acceptable standard of education[200]. Moreover, as in many other parts of the world, the turnover of HCPs is high[198], meaning that many HCPs are inexperienced, especially in managing end-of-life pain[201]. Hence it is important that the chosen instrument is adjusted for both the educational and linguistic background of the HCPs who will administer the instrument; this is something that has now been done for the APS-SE (Study I).

The actual implementation should preferably take place in several steps, both for better acceptability among the HCPs and to be made aware of any limitations for the implementation within the current context[197,198]. A practical challenge is how to structure the education around the different shifts, and how to address the ongoing need for introduction of the instrument due to high turnover of HCPs[198]. During the implementation, the coordinator must keep reminding the HCPs why it is important to use the instrument. The coordinator must also address and resolve any complaints promptly, while maintaining the strategies for incorporating the instrument into clinical routine[197].

It is essential to make sure there is enough time for staff reflection and debriefing. Incentives such as certification when attending the different educational steps can also be helpful to further promote the implementation[198]. External support may benefit both managers and HCPs, and could be provided, for example, by building networks with other facilities that are geographically near, or on a national level, to exchange information on the pros and cons of implementing the instrument[198]. Since the APS is used throughout Sweden, it is reasonable to believe that such a network could be created around the APS.

In the next stage, assessment and improvement, the instrument has to be evaluated to determine if it continues to be clinically relevant. It is also critical that the use is not too burdensome for the HCPs when applying it among their patients[197]. In general, implementing a new instrument into clinical practice takes both time and effort; it has been suggested that a full implementation may take at least a year[197].

A study evaluating the implementation of an earlier version of the APS in nursing homes in Sweden reported that the HCPs felt that the APS made them feel more secure in their assessment of pain, as they went from assuming to knowing[202]. However, the evaluation further revealed a lack of knowledge regarding why the APS had been introduced, since not all HCPs had been informed about the decision to implement the instrument. In general, the implementation was considered to be time consuming, and there was still a lack of routines for how often the assessments should be made, and how to report, follow-up, and document the results of the assessments[202]. This study showed the importance of the last stage of the implementation strategy, comprising assessment and improvement, after the introduction of a new instrument.

Why does the APS-SE not work in patients with cancer?

The lack of validity and reliability when using the APS-SE in patients with cancer seems to be an inherent problem within the original instrument itself, and not related to the actual translation from English to Swedish, as explained below. The instrument works much better in a dementia population than in a cancer population. This makes sense, since it was originally intended for patients with end-stage dementia[77].

Patients with cancer may need an observational pain assessment instrument later in their disease trajectory compared to patients suffering from dementia. If too fatigued or somnolent to vocalize their pain, patients with cancer are probably more physically affected and more likely to be bedridden. In this scenario, the suggested items in the APS-SE such as fidgeting and rocking are inapplicable. This is in contrast to patients with dementia, where these body movements are mostly seen in patients with severe dementia[203]. Even though it is recommended by the IASP that observational pain assessment should take place during physical activity[204], this may not always be possible in a palliative oncology setting.

The suggested behavioural changes in the APS-SE, such as increased confusion or refusal to eat, are seen as problematic to evaluate in patients with cancer. The HCPs interviewed in Study II found it hard to know how to assess these items, since they were often perceived as being more associated with the cancer itself than with pain.

This seems plausible, since the reported prevalence of delirium in dying patients with cancer is over 50%[205] and the prevalence of cancer cachexia is even higher, around 65%[206].

Cancer pain can be divided into acute pain, chronic pain, and a combination of both. Acute pain is often related to various diagnostic procedures and therapeutic interventions, and generally vanishes within three months, while chronic pain is usually more related to the cancer itself and can last for longer than three months[15,73]. Vital parameters such as temperature, pulse, respiration rate, and blood pressure are reliable measures for acute pain[73,104,203,207]. By the time an observational pain assessment instrument is needed due to fatigue or unconsciousness near the end of life, patients with advanced cancer are probably more likely to be suffering from persistent chronic pain than from acute pain.

The APS-SE item 'physical changes' was also disputed by some of the HCPs interviewed in Study II. They argued that physical changes were to be seen as a question of diagnostics, and not as a proxy for pain. In any case, some of the examples given for this item in the APS-SE are not very suitable for patients with cancer. No literature has been found describing the prevalence of contractures in patients with cancer, but my own clinical experience indicates that contractures are very seldom seen in this population.

If using the item 'physical changes' in patients with cancer, it would be more reasonable to register potential pain related to cancer, such as pain from bone metastasis, cancerous wounds, or tumour pressure on various organs. Other possible sources of pain could be from the cancer treatment, such as neuropathic pain after chemotherapy, osteoporotic compression fractures due to the use of corticosteroids[73], or the existence of catheters for procedures such as paracentesis or thoracentesis.

The APS-SE was perceived to not fully suit the clinical situation, partly because vocalization and facial expression were the only items considered fully useful by all participants in Study II. Vocalization consists of non-verbal expressions of pain such as groaning, whimpering, or crying, as well as different kinds of verbal expressions of pain. Verbal expressions can include the mentioning of pain or swearing, but do not include any deliberately or consciously self-reported pain[208].

A systematic review from 2020 investigating the validity of using vocalization as a pain indicator reported that an association was seen between pain and vocalization in most, but not all, of the studies found in the literature about the subject. Vocalization seemed to be more clearly

associated with acute pain than with chronic pain[208]. The review also concluded that both HCPs and family members firmly believed that vocalization served as an important proxy for pain[208]. This opinion was shared by the HCPs in Study II. The review included studies with a variety of different populations, such as children, patients with brain injury, elderly people with cognitive impairment, and patients in intensive care, but none of the study populations consisted of patients with cancer[208].

Several studies have found an association between facial expression and pain[209-211]. Although observational pain assessment instruments usually include an item for facial expression, few instruments use objectively described examples, as in the examples 'grimacing' versus 'closing eyes'[212]. Anatomically orientated expressions in particular, such as 'raising upper lip' or 'opened mouth', seem to help distinguish whether a patient is in acute pain[210], and facial expressions during pain seem to be unaffected by the patient's cognitive status[211].

Furthermore, a systematic review showed that even if not applicable to every patient in every situation, there seems to be a fairly consistent sub-set of facial expressions in patients experiencing pain. The systematic review included studies with various populations, such as people with cognitive impairment, intubated patients, and healthy patients after knee replacement, but again none of the study populations consisted of patients with cancer[211].

Pain assessment and artificial intelligence

Several attempts have been made in recent years to incorporate artificial intelligence (AI) in the assessment of facial expressions as a pain indicator. In a study from 2022 comparing manually coded facial expressions as gold standard against automatic coding with the program FaceReader7, the association was poor to moderate, indicating a need for further development of the automatic coding program[213]. PainChek® (formerly known as ePAT) is a commercially developed pain assessment tool with a mobile application that combines automated facial recognition with observed non-facial pain behaviours to estimate pain intensity in patients with dementia. A test of its validity showed a high correlation between PainChek® and the APS[214].

The use of technology-based pain assessments such as PainChek® allows 'big data' research; that is, the digital collection and analysis of large amounts of data[212]. A study involving more than 22 000 pain assessments concluded that eye-related expressions of pain were more common than mouth-related

expressions among dementia patients suffering from pain[212]. More such automatic coding or facial recognition programs will undoubtedly be created in the near future, but it remains to be seen whether all of them will be commercially developed or whether any non-profit versions will be available.

Clinical considerations if using the APS-SE

This thesis concludes that the APS-SE is inadequate when used in patients with advanced cancer. Today, different versions of the APS are also used in many Swedish nursing homes despite not being fully evaluated in a dementia population in Sweden. The APS-SE has been scientifically translated and adapted to a Swedish dementia care setting, but its validity and reliability in this context remain unclear.

Nevertheless, the instrument was considered easy to use according to the HCPs in Study I, and this has been confirmed by others [77,103,105,107,108,192,202]. When introducing the APS-SE it is important to allow sufficient time for its implementation, especially if the HCPs lack Swedish-language fluency and/or health care education, as described in Study I.

Study I also highlighted that the utility of the APS-SE depends on the assessor having some pre-existing knowledge of the patient, since several items are assessed against the patient's usual parameters, behaviour, and body language[77]. This explicit need for prior knowledge of the patient when using the instrument has been described by others[202]. It also makes the APS-SE less useful in environments such as the emergency department, unless the patient is accompanied by a HCP or a family member, as noted by other studies[215,216].

An unreported problem until Study I was the perceived difficulty of applying the items 'flushing' and 'pallor' in people with darker skin tones. While 'flushing' was sometimes accompanied by a higher skin temperature, which made it easier to detect, this was not the case for 'pallor'. This made the HCPs doubt their assessment. Although halogen or natural lightning is recommended to avoid a misleading bluish tint[217], when using the APS-SE there is a need for education on assessment of pallor and flushing in darker skin tones.

Methodological considerations

A strength of this thesis is that its conclusion, that the use of the APS-SE is inadequate in patients with advanced cancer, has been scientifically demonstrated with both qualitative and quantitative research methods.

Study I

When planning a translation and cultural adaptation of an instrument to another language and country, it is important to consider the setting and population in which the translation will be used[130]. Even though the overall aim of this project was to evaluate APS for people with cancer in a palliative care setting, the original APS was created for people with end-stage dementia in nursing homes. Thus, the decision was made to keep the original population, people with end-stage dementia, but to broaden the original setting by including hospitals and specialized home care units as well, to make the instrument more diversified.

Transferring the instrument to both a different language and another population simultaneously would have added an unnecessary aspect of uncertainty. Specifically, if the new version of the APS failed, it would be difficult to determine whether this was due to an inferior translation or to the shift to another population.

The translation and cultural adaptation of the APS was primarily based on two different guidelines. The first is a well-established general guideline for healthcare instrument translation by Beaton et al.[130], which according to the Web of Science has been cited almost 8000 times[218]. An established guideline from a palliative care context was chosen as a complement, a guideline from Antunes et al.[131].

As part of the translation and adaptation process, a first conceptual phase was used to assess whether each item in the APS had an equivalent item in the targeted cultural care context. The HCPs involved in the conceptual phase and during the rest of the translation phase did not work exclusively with people with dementia; instead, they worked in specialized palliative care settings where they often met people with cognitive impairment and/or delirium. These HCPs were selected since they also had extensive experience of another Swedish version of the APS. The outcome of the conceptual discussion and the translation might have been different if HCPs working only with people with dementia had been included, but this was considered unlikely.

A strong point of Study I was the repeated interviews to test the comprehension of the APS-SE, not only in a nursing home setting but also in a hospital and a specialized palliative home care setting.

When testing a translation along with the comprehension of a newly translated instrument, it is important to reflect on who the actual administrator or user of the instrument is[130]. Thus, the sampling was aimed at identifying and recruiting participants who would most likely provide the best information to answer the research question, as well as achieving variation in gender, age, linguistic and educational background, and clinical setting.

During the initial interviews it became evident that in some cases the comprehension of the APS-SE depended on educational background in health care and/or whether the participants had Swedish as a first language. Hence, a conscious decision was made to deliberately include more participants with diverse backgrounds in terms of native language and education in health care. Nevertheless, recruitment of additional nursing assistants would have better represented the majority of the administrators in nursing homes.

To simulate the reality of the context in which the APS-SE would actually be used, both experienced and inexperienced users of the APS were recruited. A high turnover of HCPs means that many users will be inexperienced, and without including inexperienced users, the voices of many HCPs would have been lost.

Each interview was transcribed verbatim and then read through, after which ambiguous items in the instrument were identified and revised. After the first interview it became obvious that a revision had to be made to explain the meaning of the word 'contractures'. During the following interviews, minor revisions were made to find more comprehensible synonyms for some of the items in order to enhance the comprehension for non-Swedish speakers. The word 'fidgeting' needed several revisions; this was as anticipated, since the difficulty of finding an equivalent Swedish synonym was evident even in the translation phase.

The revisions were made between the different interviews to allow prompt testing of whether the revision made the next version more understandable. Once discarded, old versions were not tried again. If all the interviews had been analysed simultaneously, the influence of the participants' native language and healthcare education would still have been noticeable.

However, the chance to actively recruit more diverse participants on a continuous basis would have been lost. As a result, a greater number of interviews would have been required to evaluate the impact of each new revision.

The inclusion in Study I was set to continue until no new essential data were exposed during three consecutive interviews. Eleven interviews were performed in total, giving a sample size in line with previous suggestions for the number of interviews needed[129]. Nevertheless, although this Swedish version of the APS was linguistically revised to enhance comprehension for people with native languages other than Swedish, there are no guarantees that the revisions will allow all users to understand every word in the instrument.

The interviews were performed to test the comprehension of the translation of the APS and to assess if the items were acceptable in a Swedish care context. However, during the readthrough of the interviews it became obvious that they contained material on the use of the APS in HCPs' everyday clinical practice. A thematic analysis was therefore also conducted to investigate factors important to the clinical implementation of the instrument.

The research question was derived from the coding process, and so an inductive analysis was used to create data-driven themes strongly connected to the data collected in the interviews. Since the research question was concerned with investigating factors for implementation, a semantic analysis seemed to be the most appropriate approach. Moreover, as the analysis was not originally intended when the interviews took place, the decision was made to perform a more elementary analysis. If a thematic analysis had been planned from the beginning, a more elaborate approach would probably have been adopted. The theme 'implementation' could also have been named differently, to further distinguish it from the research question.

Study II

When there is little or no information and knowledge available on a phenomenon, qualitative research can be a useful and appropriate way to approach the research question[219]. Study II used an established method, qualitative content analysis as described by Graneheim and Lundman(2004)[122]. I had the opportunity to take courses with Graneheim and Lindgren both in qualitative content analysis and in interviews and observations as qualitative data collection methods.

Both the semi-structured interview guide in Study II and the analysis of the first pilot interview in Study II were used during these courses, and received valuable feedback from Graneheim and Lindgren.

As previously discussed in the introduction to this thesis (see pages 20–22), trustworthiness is important when evaluating the results of a qualitative study. Trustworthiness can be approached in different ways: through credibility, authenticity, dependability, confirmability, and transferability[122-126].

Credibility concerns the data, and whether the interpretation of the data is representative of the participants' views. It can be enhanced in several ways[122,123,125,126]. As qualitative content analysis emphasizes diversity and variation within the data[122,123], a conscious decision was made to include both experienced and inexperienced HCPs in order to capture any diverging points of view between the different groups. If only experienced users had been invited, the voices of the many HCPs who have only just encountered the APS-SE would not have been heard.

Several other approaches have been suggested to enhance credibility in qualitative research in general. *Prolonged engagement* allows the researcher to establish trust and become familiar with the context and setting[125]. In this case, the researcher had an established relationship with the majority of the participants and was already familiar with the clinical context and setting. All the interpreted data were also originally collected by the same researcher(s), further ensuring familiarity with the data.

Another such approach is *triangulation*; for example, using different ways of gathering data and having several researchers involved in analysing the data. Different types of triangulation exist, including data triangulation, investigator triangulation, and method triangulation[125]. In Study II, data triangulation was implemented by collecting data both in different clinical settings (a standalone specialized palliative inpatient care unit, hospital wards, and a specialized home care unit) and from different professions (nurses and physicians). To broaden the investigation, interviews could also have been performed with nursing assistants. Although the views of patients and family members are always of interest, they were not part of the aim of this study. Investigator triangulation was also applied, as the coding, analysis, and interpretation of the data took place in collaboration between several researchers.

Data were collected through individual interviews. Since one important goal of the research was to understand the experience of using each of the separate item in the APS-SE, a semi-structured interview guide was used instead of a narrative approach. The structured aspect of the guide ensured that all participants addressed their experience of each separate item in the APS-SE, but the guide also allowed more in-depth follow-up questions. An alternative to the individual interviews would have been focus groups, allowing more participants to be interviewed.

An additional way to achieve credibility in qualitative research in general is by using a *member check*; that is, getting feedback from the participants themselves to ensure that the results reflect their views[124,125]. No formal member check took place during the analysis in this study, but an informal check was conducted when the final results were reported back to the participating clinical settings. Multiple participants confirmed that they recognized the results and conclusion according to their own work experience. When the findings were presented at a national conference in palliative medicine, the results were once more verbally confirmed by several audience members from different parts of Sweden. Hence, it is reasonable to conclude that a member check during the analysis would not have led to any major differences in the final results.

To enhance both credibility and authenticity in qualitative content analysis, it is important to provide examples of the interpretation and abstraction process, along with representative quotations from different participants when presenting the results[123]. Both of these suggested measures were used when reporting the results in Study II.

When evaluating dependability in qualitative content analysis, i.e. the consistency of the research over time, there is a risk of inconsistency if the data are gathered over longer periods of time[122]. The interviews in Study II took place during November and December 2019, and so inconsistency due to extended data collection was not considered a problem. As explained in a methodological overview of trustworthiness in qualitative research in general, both confirmability, i.e. whether the results can be confirmed by another researcher, and dependability can be assessed with the help of an audit[125]. However, no audit was performed in Study II.

The transferability of the results, either to another specific setting or to a different population, is to be determined by the reader. Both in qualitative content analysis and in qualitative research in general, this calls for a solid description of the setting and situation in which the research was performed[118,122-125]. However, the need and desire to protect the

confidentiality of the participants can sometimes be in conflict with the need to give a more explicit description to help the reader to evaluate the transferability[118].

A strength in Study II was the recruitment of both experienced and inexperienced participants, as well as participants from several different settings. By mimicking a 'real' clinical context, which will likely contain many inexperienced users of the APS-SE, and by using different settings, the possibility of being able to transfer the result to the desired clinical setting is increased[123,125]. Transferability could have been further enhanced by giving a more thorough description of the settings and the population; for example, the different types of cancer that are usually cared for in the standalone specialized palliative inpatient care unit, or the patients' average time until death.

During the interviews, the participants were asked to visualize a patient with advanced cancer whom they had previously cared for, but they were given no further instructions on which patient to choose. This procedure was chosen in order to ensure that the HCPs could reflect on how and in which patients they used the APS-SE without any pre-set boundaries. The results indicated that the APS-SE was used in several different scenarios that had not previously been anticipated. With more strict pre-set boundaries, this information may have been missed.

The aim of this thesis is of course no coincidence. When managing pain in patients with advanced cancer in everyday practice, I had a suspicion that the APS was not a good instrument for assessing pain in this population. This preconceived assumption was the reason for conducting this project - to either confirm or reject the suspicion. When conducting a study, the researcher's pre-understanding influences everything from how questions are asked to how the participants' answers are understood and interpreted[123]. Being aware of this, I tried my best to highlight the participants' experiences rather than my own. The use of an interview guide and analysing the results together with others will hopefully have helped with this.

It is possible that the participants who worked together in the same workplace might have discussed their interviews with each other, and thus influenced each other's answers. To reduce this problem, the study was conducted in three different settings. To further reduce the problem, participants could have been recruited from other parts of the country.

The inclusion in Study II was set to continue until no new essential data were exposed during three consecutive interviews. Twelve interviews were conducted in total, giving a sample size in line with previous suggestions for the number of interviews needed[129].

Study III

After a literature review focusing on how to psychometrically test the instrument, the statistical methods were chosen in collaboration with an experienced statistician. However, due to the unpredictability of what range of APS-SE scores could be expected within a cancer population, the sample size could not be determined. Instead, the literature was once more consulted to find the sample sizes used in earlier studies. The original APS development study included 113 patients, and other studies validating the APS in patients with dementia used 50–171 patients[77,103-106]. In similar studies validating two common assessment instruments in palliative care, the IPOS and the ESAS, the sample sizes varied from 23 to 376 patients[69,153-179].

The sample size for Study III was set to 50–100 patients. An initial analysis was performed after the inclusion of 50 patients, since studies imply that 30–50 is a reasonable number of patients for calculating reliability[121,220]. This analysis showed that the inter-rater reliability had an ICC with a confidence interval ranging from poor to good. Considering the wide confidence interval, an additional 22 patients were included. The next analysis, at 72 patients, still revealed a confidence interval ranging from poor to good inter-rater reliability, and thus further inclusion was ended. However, if further patients had been included without introducing additional raters, the confidence interval for the inter-rater reliability would likely have become narrower.

One of the inclusion criteria, delirium, was assessed with the SQiD, which has been previously tested in patients with cancer[148]. In the original setting, the question ‘Do you feel that [patient’s name] has been more confused lately?’ is answered by the patient’s family member or a friend[147]. In Study III, the question was instead assessed by the HCPs in charge. In all likelihood, family members and friends have a deeper understanding of the patient and can more precisely identify any delirium, so this alteration could have led to an over- or underestimation of delirium among the patients.

Furthermore, delirium per se fluctuates over time[143]. Thus, some of the patients could have been assessed as cognitively intact at the time of inclusion but suffered from delirium at the time of assessment. As an attempt at minimizing this problem, the assessments were made approximately one hour after the inclusion. However, the situation could also have been the other way around, with patients being confused at inclusion but not at assessment. As in clinical reality, there was no certainty that the patients were cognitively intact when self-reporting pain.

Patients with delirium may struggle to accurately self-report their pain when using the NRS, potentially leading to NRS ratings that might not properly reflect their actual level of pain. To test whether the inclusion of delirious patients could explain the results, a subgroup analysis was performed including only those patients without delirium. This analysis revealed a similar outcome, with a slight correlation between self-reported pain and observed pain, indicating that this was not a major problem when analysing the criterion validity.

Pain may also fluctuate over time. The pain assessments were performed at different times throughout the day, but for practical reasons none of them took place between 5 pm and 8.30 am. Any recurring pain during evenings and nights would subsequently have been lost.

Fluctuation of pain over time might also have affected the test-retest analysis. In general, test-retest measurements are performed to analyse whether an instrument is stable over time, and are therefore often repeated weeks or months apart[114]. However, in Study III the second assessment took place approximately an hour after the first assessment. Since the targeted population had a median time of 11 days until death, it was not practical to repeat the test weeks or months later. In an end-of-life clinical situation, the APS-SE is believed to be stable over at a least a short, limited time span.

For a proper evaluation of the instrument's responsiveness, patients had to have at least a four-hour washout time after being given any as-needed opioid. In a few cases (n=5) this criterion meant that patients could not be included, since they never had four hours or more between opioids during the period from screening until death. One can only speculate that the APS-SE and NRS scores might have been different in these patients compared to the patients who were included. In hindsight, one way to avoid this problem would have been to include these patients, but only make a first assessment, this actively sacrificing the opportunity to measure responsiveness in order to ensure the inclusion of these specific patients.

As in Studies I and II, the intent when planning this study was to mimic the reality of an ordinary clinical setting, and hence to involve HCPs working their ordinary shift as raters instead of using two previously selected raters as in other APS validation studies[77,103,104]. Regardless of how well they were acquainted with the APS-SE, all HCPs received a short review of the protocol and the APS-SE prior to the first assessment. During this time, between February 2022 and March 2023, there was a high turnover of HCPs. This led to 40 different raters assessing 72 patients, which is a truthful reflection of the reality of the clinical settings.

One overall methodological problem was that the raters were aware of whether or not the patient received opioids between the first and the second assessment. In the majority of the assessments, the nurse in charge of the patient was also one of the raters, and it was this nurse who made the clinical decision of whether to administer opioids based on the first APS-SE assessment. In a busy clinical environment, it was not achievable to have a second nurse make this decision and administer the opioids within a reasonable amount of time.

The two raters at the second assessment could have been unconsciously influenced by knowing whether or not the patient had received opioids, and this might have affected both the responsiveness and the test-retest results. However, since both raters had the same information about opioids, there was unlikely to have been any effect on the inter-rater reliability.

Conclusions

The Abbey Pain Scale (APS), an observational pain assessment instrument initially developed for patients with dementia, has been used in various Swedish versions to evaluate pain in patients with advanced cancer who are unable to vocalize their pain. However, the APS has not been evaluated in this specific population.

In this thesis, the APS was scientifically translated and culturally adapted to the Swedish dementia care context, resulting in the Swedish version APS-SE. The APS-SE underwent modifications to enhance its comprehensibility across diverse educational and linguistic backgrounds before being evaluated in patients with advanced cancer.

Both qualitative and quantitative methods revealed that the APS-SE is inadequate for assessing pain in patients with advanced cancer. It fails to detect moderate or severe cancer pain. This limitation appears inherent to the instrument itself, regardless of translation quality or the specific Swedish version used.

The thesis underscores the need for a specialized observational pain assessment instrument explicitly tailored for patients with advanced cancer. No recognized alternative currently exists, emphasizing the importance of developing such an instrument to address the critical gap in observational pain assessment in the palliative oncology setting.

Future research

The APS-SE is now scientifically translated and culturally adapted to a Swedish dementia care context. International studies show reasonable validity and reliability for patients with dementia, but the instrument has not been fully evaluated in a Swedish dementia context. A further step would be to examine the validity, reliability, and responsiveness of the APS-SE when used in this population.

Cancer is one of the most common causes of death, with millions of people worldwide suffering from pain at the end of life. The need for an observational pain assessment instrument for these affected patients is evident. The next step would be to create such an instrument.

Based on the low validity and reliability for the APS-SE in patients with advanced cancer, a mere revision of the APS-SE would not be sufficient to make the instrument more suitable for this population. The development of

a new instrument would have to start from the beginning, building a foundation through an intensive literature search and in collaboration with experienced HCPs working in anaesthesiology, oncology, and palliative medicine. When reviewing which items should be included, it is important that each item has been properly evaluated for patients with cancer pain (as, for example, vocalization has not; see below). Since cancer pain is a global problem, this could preferably be executed as an international multi-centre study.

The use of vocalization as an indicator of pain has not been investigated in patients with cancer. As vocalization is likely to be suggested as a part of a future observational pain assessment instrument, it would be beneficial to explore the association between vocalization and chronic cancer pain. Such an exploration might produce functional examples of vocalization to be used in the new instrument, such as 'sighing' or 'groaning'. One possible approach might be to use recording devices to gather large amounts of material for analysis.

Likewise, there is a need to investigate the association between facial expression and pain in patients with cancer. As AI has been used to create 'big data' research to establish which types of facial pain expressions are more common among patients with dementia, similar studies might be performed in patients with advanced cancer. Without such research, it would be difficult to create a robust observational pain assessment instrument for patients with cancer.

In all likelihood, a future observational pain assessment would include AI, in order to reduce differences between raters and to improve the accuracy of reassessment of pain after any intervention. Devices to assess facial expressions already exist, and perhaps AI could also be used to analyse vocalization. A commercially available device is already available for assessing pain in patients with dementia; this uses a combination of AI assessment and observational assessment from HCPs, and the same would probably be needed for a similar device for use in patients with cancer.

Regardless, it is important for options other than AI-driven devices to exist for assessing pain. An alternative instrument must also be developed, with written descriptions of what to look for, such as 'closing eyes' or 'sighing'. This is the only way to make the instrument available for *everyone*. Many of the patients suffering from cancer pain live in low- and middle-income countries without the opportunity of using an expensive device. As the WHO concludes, it is a human right to receive palliative care regardless of who you are and where you live.

Acknowledgements

From the bottom of my heart, I want to thank the following:

Mats Sjöström, my devoted husband and confidant. Thank you for your valuable responses after every readthrough during this entire journey, and for all the cups of tea that you brought me during my writing.

Linus, Emil, and Elin Tegenborg Sjöström, my lovely and loving children. You are my biggest achievements in life. I hope I can make you proud of me, the way I am always proud of you.

Anna Näslund and Anna Rosén, my best friends, who always generously listen and give encouraging advice.

My dear friends and colleagues at Palliative Medicine in Umeå – thank you for making this possible. A special thanks to Annika Kramer, Susanne Arvidsson, Karin Palmberger, Henrik Ångström, and Karin Lindahl; without your unwavering support, this thesis would not exist.

Mats Kärnestad and Margareta Fahlström and all the other HCPs in Axlagården Hospice, for your support of research in general, and me in particular.

All the HCPs in the Department of Oncology, the Department of Geriatrics, the Emergency Department, and the nursing homes, who participated and helped in every way during all the studies. And all the patients, although no longer with us, who so altruistically contributed to research in a palliative care setting.

Last, but not least, my supervisors Per Fransson and Camilla Thellenberg Karlsson, but especially my main supervisor Lisa Martinsson. Thank you, Lisa, for all your hard work in carrying me through this adventure to the end. I have cherished our lunch meetings with equal amounts of work and socializing, and I hope our journey continues together. There is still so much to do!

References

[1] Raja SN, Carr DB, Cohen M, et al. The revised International Association for the Study of Pain definition of pain: concepts, challenges, and compromises. Pain. 2020;161(9):1976-1982.

[2] Clark D. 'Total pain', disciplinary power and the body in the work of Cicely Saunders, 1958-1967. Soc Sci Med. 1999;49(6):727-736.

[3] Stewart J. The challenges of cancer pain assessment. Ulster Med J. 2014;83(1):44-46.

[4] Yang Y, Zhao X, Cui M, et al. Dimensions of spiritual well-being in relation to physical and psychological symptoms: a cross-sectional study of advanced cancer patients admitted to a palliative care unit. BMC palliat care. 2023;22(1):137.

[5] Cancer.Net. Coping with cancer. 2024. [Accessed 2024 28 February, at: https://www.cancer.net/coping-with-cancer/physical-emotional-and-social-effects-cancer/spiritual-support-when-you-have-cancer].

[6] Aktaş A, Uğur Ö. The effect of physical and psychological symptoms on spiritual well-being and emotional distress in inpatient cancer patients. Support care cancer. 2023;31(8):473-483.

[7] Evenepoel M, Haenen V, De Baerdemaecker T, et al. Pain Prevalence During Cancer Treatment: A Systematic Review and Meta-Analysis. J Pain Symptom Manage. 2022;63(3):317-335.

[8] Snijders RAH, Brom L, Theunissen M, et al. Update on Prevalence of Pain in Patients with Cancer 2022: A Systematic Literature Review and Meta-Analysis. Cancers. 2023;15(3):591.

[9] Copenhaver DJ, Huang M, Singh J, et al. History and Epidemiology of Cancer Pain. Cancer Treat Res. 2021;182:3-15.

[10] van den Beuken-van Everdingen MHJ, 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.

[11] 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.

[12] Roberto A, Greco MT, Uggeri S, et al. Living systematic review to assess the analgesic undertreatment in cancer patients. Pain Pract. 2022;22(4):487-496.

[13] Greco MT, Roberto A, Corli O, et al. Quality of cancer pain management: an update of a systematic review of undertreatment of patients with cancer. J Clin Oncol. 2014;32(36):4149-4154.

[14] Vuong S, Pulenzas N, DeAngelis C, et al. Inadequate pain management in cancer patients attending an outpatient palliative radiotherapy clinic. Support Care Cancer. 2016;24(2):887-892.

[15] Fink RM, Gallagher E. Cancer Pain Assessment and Measurement. Semin Oncol Nurs. 2019;35(3):229-234.

[16] Hachem GE, Rocha FO, Pepersack T, et al. Advances in pain management for older patients with cancer. ecancer. 2019;13:980.
[17] 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.
[18] Saifan AR, Bashayreh IH, Al-Ghabeesh SH, et al. Exploring factors among healthcare professionals that inhibit effective pain management in cancer patients. Cent Eur J Nurs Midw. 2019;10(1):967-976.
[19] 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.
[20] Bhadelia A, De Lima L, Arreola-Ornelas H, et al. Solving the Global Crisis in Access to Pain Relief: Lessons From Country Actions. Am J Public Health. 2019;109(1):58-60.
[21] Karcioglu 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.
[22] 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.
[23] Booker SQ, Haedtke C. Assessing pain in nonverbal older adults. Nursing. 2016;46(5):66-69.
[24] Lukas A, Barber JB, Johnson P, et al. Observer-rated pain assessment instruments improve both the detection of pain and the evaluation of pain intensity in people with dementia. Eur J Pain. 2013;17(10):1558-1568.
[25] 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.
[26] Ryan S, Wong J, Chow R, et al. Evolving Definitions of Palliative Care: Upstream Migration or Confusion? Curr Treat Options Oncol. 2020;21(3):20.
[27] Van Mechelen W, Aertgeerts B, De Ceulaer K, et al. Defining the palliative care patient: a systematic review. Palliat Med. 2013;27(3):197-208.
[28] World Health Organization. National Cancer Control Programmes. Policies and managerial guidelines. 2nd Edition. 2002. [Accessed 2023 27 October, at: https://iris.who.int/bitstream/handle/10665/42494/9241545577.pdf?sequence=1].
[29] World Health Organization. Palliative care. 2020. [Accessed 2023 25 October, at: https://www.who.int/news-room/fact-sheets/detail/palliative-care].

[30] International Association for Hospice and Palliative Care. Palliative Care Definition. 2019. [Accessed 2023 25 October, at: https://hospicecare.com/what-we-do/projects/consensus-based-definition-of-palliative-care/definition/].
[31] Radbruch L, De Lima L, Knaul F, et al. Redefining Palliative Care - A New Consensus-Based Definition. J Pain Symptom Manage. 2020;60(4):754-764.
[32] Knaul FM, Farmer PE, Krakauer EL, et al. Alleviating the access abyss in palliative care and pain relief-an imperative of universal health coverage: the Lancet Commission report. Lancet. 2018;391(10128):1391-1454.
[33] Payne S, Harding A, Williams T, et al. Revised recommendations on standards and norms for palliative care in Europe from the European Association for Palliative Care (EAPC): A Delphi study. Palliat Med. 2022;36(4):680-697.
[34] National Board of Health and Welfare [Swe: Socialstyrelsen]. Palliativ vård – förtydligande och konkretisering av begrepp [Palliative care – clarification and definitions of terms]. 2018. [Accessed 2023 25 October, at: https://www.socialstyrelsen.se/globalassets/sharepoint-dokument/artikelkatalog/ovrigt/2018-8-6.pdf].
[35] National Board of Health and Welfare [Swe: Socialstyrelsen]. Socialstyrelsens termbank [The term bank of the National Board of Health and Welfare]. 2021. [Accessed 2021 7 June, at: https://termbank.socialstyrelsen.se/?TermId=462&SrcLang=sv].
[36] World Health Organization. Assessing national capacity for the prevention and control of noncommunicable diseases. 2020. [Accessed 2024 22 March, at: https://iris.who.int/bitstream/handle/10665/331452/9789240002319-eng.pdf?sequence=1].
[37] Westley-Wise V, Moules S, Masso M, et al. Who needs, receives and misses out on palliative and end-of-life care? A population-based study to identify needs and gaps in a regional health service. Aust Health Rev. 2022;46(1):91-99.
[38] Morin L, Aubry R, Frova L, et al. Estimating the need for palliative care at the population level: A cross-national study in 12 countries. Palliat Med. 2017;31(6):526-536.
[39] Murtagh FE, Bausewein C, Verne J, et al. How many people need palliative care? A study developing and comparing methods for population-based estimates. Palliat Med. 2014;28(1):49-58.
[40] Etkind SN, Bone AE, Gomes B, et al. How many people will need palliative care in 2040? Past trends, future projections and implications for services. BMC Med. 2017;15(1):102.
[41] Knaul FM. Integrating palliative care into health systems is essential to achieve Universal Health Coverage. Lancet Oncol. 2018;19(11):566-567.

[42] Center to Advance Palliative Care. Cancer the Most Common Diagnosis in Palliative Care Patients. 2019. [Accessed 2023 25 October, at: https://www.capc.org/blog/news-bites-cancer-common-diagnosis-palliative-care-patients/?clickthrough_doc_id=blog.blogpostpage.930&clickthrough_req_id=&clickthrough_query=].
[43] Schoenherr LA, Bischoff KE, Marks AK, et al. Trends in Hospital-Based Specialty Palliative Care in the United States From 2013 to 2017. JAMA Netw Open. 2019;2(12):e1917043.
[44] World Health Organization. Cancer. 2022. [Accessed 2022 April 7, at: https://www.who.int/news-room/fact-sheets/detail/cancer#:~:text=Cancer%20is%20a%20leading%20cause%20of%20death%20worldwide%2C,breast%2C%20lung%2C%20colon%20and%20rectum%20and%20prostate%20cancers].
[45] Axelsson B. Palliative medicin i praktiken [Palliative medicine in practise]. Lund: Studentlitteratur; 2023. Chapter 1.
[46] Swedish Government Official Report (SOU) [Swe: Statens Offentliga Utredningar (SOU)]. Vårdens svåra val. Slutbetänkande av Prioriteringsutredningen (SOU 1995:5) [The difficult decision for health care. Final report of the Priority investigation (SOU 1995:5)]. 1995. [Accessed 2023 9 November, at:http://www.regeringen.se/rattsdokument/statens-offentliga-utredningar/1995/03/sou-19955/].
[47] National system for knowledge-based management within healthcare [Swe: Nationellt system för kunskapsstyrning hälso- och sjukvård]. Personcentrerat och sammanhållet vårdförlopp Palliativ vård [Careprocess for Palliative care]. 2023. [Accessed 2023 30 November, at: https://kunskapsstyrningvard.se/kunskapsstyrningvard/kunskapsstod/publiceradekunskapsstod/aldreshalsaochpalliativvard/vardforlopppalliativvard.64368.html].
[48] Swedish Register of Palliative Care [Swe: Svenska Palliativregistret]. Årsrapport för Svenska palliativregistret 2022 [The annual report from the Swedish Register of Palliative Care 2022]. 2023. [Accessed 2023 15 November, at: https://palliativregistret.se/media/pykbuogo/%C3%A5rsrapport-2022.pdf].
[49] Swedish Council for Palliative Care [Swe: Nationella Rådet för Palliativ Vård]. Palliativ guiden [The guide to Palliative care]. 2024. [Accessed 2024 27 February, at: https://www.nrpv.se/].
[50] National Board of Health and Welfare [Swe: Socialstyrelsen]. Statistik om dödsorsaker [Statistics on cause of death]. 2023. [Accessed 2024 3 March, at: https://www.socialstyrelsen.se/statistik-och-data/statistik/alla-statistikamnen/dodsorsaker/].
[51] National Board of Health and Welfare [Swe: Socialstyrelsen]. Dataanalyser av cancer 1975–2019 [Data analysis of cancer 1975-2019]. 2021. [Accessed 2024 3 March, at: https://www.socialstyrelsen.se/globalassets/sharepoint-dokument/artikelkatalog/ovrigt/2021-3-7285.pdf].

[52] Olsson Möller U, Beck I, Fürst CJ, et al. A Reduced Deck of Conversation Cards of Wishes and Priorities of Patients in Palliative Care. J Hosp Palliat Nurs. 2022;24(3):175-180.
[53] University at Buffalo. Instrument Tool. 2023. [Accessed 2023 18 December, at: https://publichealth.buffalo.edu/cat/kt4tt/best-practices/need-to-knowledge-ntk-model/ntk-instrument-tool.html].
[54] Teachers College Columbia University. Research Instrument Examples. 2023. [Accessed 2023 18 December, at: https://www.tc.columbia.edu/media/administration/institutional-review-board-/guide-amp-resources---documents/Published_Study-Material-Examples.pdf].
[55] Cicely Saunders Institute. The Palliative care Outcome Scale outline for new & current users. 2023. [Accessed 2023 24 November, at: https://pos-pal.org/maix/pos-and-ipos-summary.php].
[56] Watanabe S, Nekolaichuk C, Beaumont C, et al. The Edmonton symptom assessment system-what do patients think? Support Care Cancer. 2009;17(6):675-683.
[57] University of Texas MD Anderson Cancer Center. The MD Anderson Symptom Inventory. 2023. [Accessed 2023 27 November, at: https://www.mdanderson.org/research/departments-labs-institutes/departments-divisions/symptom-research/symptom-assessment-tools/md-anderson-symptom-inventory.html].
[58] MAPI Research Trust. McGill Pain Questionnaire (MPQ). 2023. [Accessed 2023 29 November, at: https://eprovide.mapi-trust.org/instruments/mcgill-pain-questionnaire#need_this_questionnaire].
[59] Kerns RD, Turk DC, Rudy TE. The West Haven-Yale Multi-dimensional Pain Inventory (WHYMPI). Pain. 1985;23(4):345-356.
[60] Fainsinger RL, Nekolaichuk C, Lawlor P, et al. An international multicentre validation study of a pain classification system for cancer patients. Eur J Cancer. 2010;46(16):2896-2904.
[61] Cleeland CS, Ryan KM. Pain assessment: global use of the Brief Pain Inventory. Ann Acad Med Singap. 1994;23(2):129-138.
[62] University of Texas MD Anderson Cancer Center. The Brief Pain Inventory. 2023. [Accessed 2023 27 November, at: https://www.mdanderson.org/research/departments-labs-institutes/departments-divisions/symptom-research/symptom-assessment-tools/brief-pain-inventory.html].
[63] 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.
[64] Euasobhon P, Atisook R, Bumrungchatudom K, et al. Reliability and responsivity of pain intensity scales in individuals with chronic pain. Pain. 2022;163(12):1184-1191.

[65] Jensen MP, Castarlenas E, Roy R, et al. The Utility and Construct Validity of Four Measures of Pain Intensity: Results from a University-Based Study in Spain. Pain Med. 2019;20(12):2411-2420.
[66] 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):217.
[67] Firdous S, Berger A, Jehangir W, et al. How Should We Assess Pain: Do Patients Prefer a Quantitative or Qualitative Scale? A Study of Patient Preferences. Am J Hosp Palliat Care. 2021;38(4):383-390.
[68] Beck I, Olsson Möller U, Malmström M, et al. Translation and cultural adaptation of the Integrated Palliative care Outcome Scale including cognitive interviewing with patients and staff. BMC palliat care. 2017;16(1):49.
[69] Bruera E, Kuehn N, Miller MJ, et al. The Edmonton Symptom Assessment System (ESAS): a simple method for the assessment of palliative care patients. J Palliat Care. 1991;7(2):6-9.
[70] Lundh Hagelin C, Klarare A, Fürst CJ. The applicability of the translated Edmonton Symptom Assessment System: revised [ESAS-r] in Swedish palliative care. Acta Oncol. 2018;57(4):560-562.
[71] Burckhardt CS, Bjelle A. A Swedish version of the short-form McGill Pain Questionnaire. Scand J Rheumatol. 1994;23(2):77-81.
[72] Bergström G, Jensen IB, Bodin L, et al. Reliability and factor structure of the Multidimensional Pain Inventory-Swedish Language Version (MPI-S). Pain. 1998;75(1):101-110.
[73] Gallagher E, Rogers BB, Brant JM. Cancer-Related Pain Assessment: Monitoring the Effectiveness of Interventions. Clin J Oncol Nurs. 2017;21(3):8-12.
[74] 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.
[75] Herr K, Bjoro K, Decker S. Tools for assessment of pain in nonverbal older adults with dementia: a state-of-the-science review. J Pain Symptom Manage. 2006;31(2):170-192.
[76] Prinsen CAC, Mokkink LB, Bouter LM, et al. COSMIN guideline for systematic reviews of patient-reported outcome measures. Qual Life Res. 2018;27(5):1147-1157.
[77] 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.
[78] Warden V, Hurley AC, Volicer L. Development and psychometric evaluation of the Pain Assessment in Advanced Dementia (PAINAD) scale. J Am Med Dir Assoc. 2003;4(1):9-15.
[79] Rat P, Jouve E, Pickering G, et al. Validation of an acute pain-behavior scale for older persons with inability to communicate verbally: Algoplus. Eur J Pain. 2011;15(2):198.

[80] Feldt KS. The checklist of nonverbal pain indicators (CNPI). Pain Manag Nurs. 2000;1(1):13-21.
[81] Lefebvre-Chapiro S. The Doloplus-2 scale: evaluating pain in the elderly. Eur J Palliat Care. 2001;8:191-194.
[82] Akbarzadeh M, Jakobsson U. Smärtbedömning hos äldre personer med kommunikationssvårigheter - en utvärdering av den svenska versionen av Doloplus-2 [Pain assessment in elderly people with communcation difficulties - an evaluation of the Swedish version of Doloplus-2]. Vård i Norden. 2007;27(4):26-31.
[83] Ekman P, Friesen WV. Facial action coding system. Environmental Psychology & Nonverbal Behavior. 1978.
[84] Husebo BS, Strand LI, Moe-Nilssen R, et al. Mobilization-Observation-Behavior-Intensity-Dementia Pain Scale (MOBID): development and validation of a nurse-administered pain assessment tool for use in dementia. J Pain Symptom Manage. 2007;34(1):67-80.
[85] Husebo BS, Strand LI, Moe-Nilssen R, et al. Pain in older persons with severe dementia. Psychometric properties of the Mobilization-Observation-Behaviour-Intensity-Dementia (MOBID-2) Pain Scale in a clinical setting. Scand J Caring Sci. 2010;24(2):380-391.
[86] Fuchs-Lacelle S, Hadjistavropoulos T. Development and preliminary validation of the pain assessment checklist for seniors with limited ability to communicate (PACSLAC). Pain Manag Nurs. 2004;5(1):37-49.
[87] Malmfors P. Utveckling av en svensk version av smärtbedömningsinstrumentet PACSLAC i en population av personer med demenssjukdom [Development of the Swedish version of the pain assessment tool PACSLAC in a population of people with dementia]. 2014. [Accessed 2023 28 November, at: https://www.demenscentrum.se/sites/default/files/globalassets/fou_pdf/14_fou-_narsjukvarden-i-ostergotland.pdf].
[88] Chan S, Hadjistavropoulos T, Williams J, et al. Evidence-based development and initial validation of the pain assessment checklist for seniors with limited ability to communicate-II (PACSLAC-II). Clin J Pain. 2014;30(9):816-824.
[89] National Board of Health and Welfare [Swe: Socialstyrelsen]. Nationella riktlinjer för vård och omsorg vid demenssjukdom [National guidelines for health and care in dementia]. 2017. [Accessed 2023 8 December, at: https://www.socialstyrelsen.se/kunskapsstod-och-regler/regler-och-riktlinjer/nationella-riktlinjer/riktlinjer-och-utvarderingar/demens/].
[90] Regional Cancer Centre [Swe: Regionala cancercentrum i samverkan]. Nationellt vårdprogram palliativ vård [National Clinical Practice Guideline for Palliative Care]. 2023. [Accessed 2023 9 November, at: https://kunskapsbanken.cancercentrum.se/diagnoser/palliativ-vard/].

[91] Edberg U-K, Söderqvist A. Validering av beteendeskalan: Ett instrument för smärtskattning av äldre individer med nedsatt kognitiv funktion inom akutsjukvård [Validating the Behavioural scale: A pain behaviour scale for pain assessment of older individuals with impaired cognitive function in emergency care]. Vård i Norden. 2012;32(4):14-17.
[92] World Health Organization. WHO Guidelines for the Pharmacological and Radiotherapeutic Management of Cancer Pain in Adults and Adolescents. 2019. [Accessed 2023 24 October, at: https://www.who.int/publications/i/item/9789241550390].
[93] Paice JA, Bohlke K, Barton D, et al. Use of Opioids for Adults With Pain From Cancer or Cancer Treatment: ASCO Guideline. J Clin Oncol. 2023;41(4):914-930.
[94] Fallon M, Giusti R, Aielli F, et al. Management of cancer pain in adult patients: ESMO Clinical Practice Guidelines. Ann Oncol. 2018;29(4):166-191.
[95] Regional Cancer Centre [Swe: Regionala cancercentrum i samverkan]. Nationellt vårdprogram för palliativ vård i livets slutskede [National Clinical Practice Guideline for Palliative Care in End-of-life]. 2016. [Accessed 2019 May 8. *No longer available from*: https://www.cancercentrum.se/globalassets/vara-uppdrag/rehabilitering-palliativ-vard/vardprogram/natvp_palliativvard_vers.2.1_dec2016.pdf].
[96] Hurley AC, Volicer BJ, Hanrahan PA, et al. Assessment of discomfort in advanced Alzheimer patients. Res Nurs Health. 1992;15(5):369-377.
[97] Simons W, Malabar R. Assessing pain in elderly patients who cannot respond verbally. J Adv Nurs. 1995;22(4):663-669.
[98] Taylor E. We Agree, Don't We? The Delphi Method for Health Environments Research. Herd. 2020;13(1):11-23.
[99] Nasa P, Jain R, Juneja D. Delphi methodology in healthcare research: How to decide its appropriateness. World J Methodol. 2021;11(4):116-129.
[100] Diamond IR, Grant RC, Feldman BM, et al. Defining consensus: a systematic review recommends methodologic criteria for reporting of Delphi studies. J Clin Epidemiol. 2014;67(4):401-409.
[101] Jünger S, Payne SA, Brine J, et al. Guidance on Conducting and REporting DElphi Studies (CREDES) in palliative care: Recommendations based on a methodological systematic review. Palliat Med. 2017;31(8):684-706.
[102] Abbey J. Putting pain scales to the test. Aust Nurs J. 2007;14(7):43.
[103] 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.
[104] 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.

[105] 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 pain detection in non-communcative patients]. Rev Soc Esp Dolor. 2013;20(1):3-7.
[106] Storti M. La validazione di una scala di accertamento del dolore in pazienti con decadimento cognitivo: la scala Abbey nella versione italiana [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.
[107] Neville C, Ostini R. A psychometric evaluation of three pain rating scales for people with moderate to severe dementia. Pain Manag Nurs. 2014;15(4):798-806.
[108] van Iersel T, Timmerman D, Mullie A. Introduction of a pain scale for palliative care patients with cognitive impairment. Int J Palliat Nurs. 2006;12(2):54-59.
[109] Liu JY, Briggs M, Closs SJ. The psychometric qualities of four observational pain tools (OPTs) for the assessment of pain in elderly people with osteoarthritic pain. J Pain Symptom Manage. 2010;40(4):582-598.
[110] Swedish Register of Palliative Care [Swe: Svenska Palliativregistret]. Abbey Pain Scale – för bedömning av smärta hos personer med demenssjukdom som har svårt att beskriva sin smärta i tal [Abbey Pain Scale - when assessing pain in people with dementia who cannot verbalise their pain]. 2011. [Accessed 2022 17 Jan, at: http://media.palliativ.se/2015/08/AbbeyPalliativ2011Sve1.pdf].
[111] 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.
[112] Karlsson C. Detection and assessment of pain in dementia care practice: Registered nurses' and certified nursing assistants' experiences. [PhD dissertation]. Jönköping: Jönköping University, School of Health and Welfare; 2015.
[113] De Vet HCW, Terwee CB, Mokkink LB, Knol DL. Measurement in medicine: a practical guide. Cambridge: Cambridge University Press; 2011. Chapter 5-7.
[114] DeVon HA, Block ME, Moyle-Wright P, et al. A psychometric toolbox for testing validity and reliability. J Nurs Scholarsh. 2007;39(2):155-64.
[115] StatsDirect Ltd. Validity. 2023. [Accessed 2023 11 December, at: https://www.statsdirect.com/help/basics/validity.htm].
[116] Hassan M. Internal Validity - Threats, Examples and Guide.2024. [Accessed 2024 2 April, at: https://researchmethod.net/internal-validity].
[117] Hassan M. External Validity - Threats, Examples and Types. 2024. [Accessed 2024 2 April, at: https://researchmethod.net/external-validity/].

[118] Polit DF, Beck CT. Generalization in quantitative and qualitative research: myths and strategies. Int J Nurs Stud. 2010;47(11):1451-1458.
[119] Gerber DE, Pruitt SL, Halm EA. Should criteria for inclusion in cancer clinical trials be expanded? J Comp Eff Res. 2015;4(4):289-291.
[120] American Society of Hematology. Who's In and Who's Out? 2017. [Accessed 2024 6 March, at: https://ashpublications.org/ashclinicalnews/news/2774/Who-s-In-and-Who-s-Out?searchresult=1].
[121] Koo TK, Li MY. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. J Chiropr Med. 2016;15(2):155-163.
[122] Graneheim UH, Lundman B. Qualitative content analysis in nursing research: concepts, procedures and measures to achieve trustworthiness. Nurse Educ Today. 2004;24(2):105-112.
[123] Graneheim UH, Lindgren BM, Lundman B. Methodological challenges in qualitative content analysis: A discussion paper. Nurse Educ Today. 2017;56:29-34.
[124] Elo S, Kääriäinen M, Kanste O, et al. Qualitative Content Analysis: A Focus on Trustworthiness. SAGE Open. 2014;4(1):1-10.
[125] Korstjens I, Moser A. Series: Practical guidance to qualitative research. Part 4: Trustworthiness and publishing. Eur J Gen Pract. 2018;24(1):120-124.
[126] Lincoln YS, Guba, EG. Naturalistic Inquiry. Thousands Oaks, CA: Sage; 1985. Chapter 11.
[127] Statisticssolutions.com. What is transferability in qualitative research and how do we establish it? 2024. [Accessed 2024 24 March, at: https://www.statisticssolutions.com/what-is-transferability-in-qualitative-research-and-how-do-we-establish-it/].
[128] Greenberg RS. Medical epidemiology: population health and effective health care (5th ed.). New York: McGraw-Hill Education LLC; 2015. Chapter 7.
[129] Hennink M, Kaiser BN. Sample sizes for saturation in qualitative research: A systematic review of empirical tests. Soc Sci Med. 2022;292:114523.
[130] 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.
[131] Antunes B, Daveson B, Ramsenthaler C, et al. The Palliative care Outcome Scale (POS) Manual for cross-cultural adaptation and psychometric validation. 2012. [Accessed 2018 3 September, at: https://pos-pal.org/doct/Manual_for_crosscultural_adaptation_and_psychometric_validation_of_the_POS.pdf].

[132] Simkus J. LA. Snowball Sampling Method: Techniques & Examples. 2023. [Accessed 2024 24 March, at: https://www.simplypsychology.org/snowball-sampling.html].
[133] Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol. 2006;3(2):77-101.
[134] Moser A, Korstjens I. Series: Practical guidance to qualitative research. Part 3: Sampling, data collection and analysis. Eur J Gen Pract. 2018;24(1):9-18.
[135] Lindgren BM, Lundman B, Graneheim UH. Abstraction and interpretation during the qualitative content analysis process. Int J Nurs Stud. 2020;108:103632.
[136] 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.
[137] Henson LA, Maddocks M, Evans C, et al. Palliative Care and the Management of Common Distressing Symptoms in Advanced Cancer: Pain, Breathlessness, Nausea and Vomiting, and Fatigue. J Clin Oncol. 2020;38(9):905-914.
[138] 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.
[139] MedlinePlus. Fatigue. 2023. [Accessed 2023 16 November, at: https://medlineplus.gov/ency/article/003088.htm].
[140] Radbruch L, Strasser F, Elsner F, et al. Fatigue in palliative care patients - an EAPC approach. Palliat Med. 2008;22(1):13-32.
[141] Al Maqbali M, Al Sinani M, Al Naamani Z, et al. Prevalence of Fatigue in Patients With Cancer: A Systematic Review and Meta-Analysis. J Pain Symptom Manage. 2021;61(1):167-189.
[142] International Association for Hospice and Palliative Care. ECOG score. 2022. [Accessed 2024 3 March, at: https://pallipedia.org/ecog-score/].
[143] American Physciatric Association. DSM-5-TR® Classification. Washington DC: American Psychiatric Association Publishing; 2022.
[144] Guo Y, Mu Y, Wu T, et al. Risk factors for delirium in advanced cancer patients: A systematic review and meta-analysis. Eur J Oncol Nurs. 2023;62:102267.
[145] Chishi KV, Patel BC, Umrania RA, et al. Prevalence of Delirium in Advance Cancer Patients Admitted in Hospice Centre and Outcome after Palliative Intervention. Indian J Palliat Care. 2023;29(1):82-88.

[146] Mercadante S, Masedu F, Balzani I, et al. Prevalence of delirium in advanced cancer patients in home care and hospice and outcomes after 1 week of palliative care. Support Care Cancer. 2018;26(3):913-919.
[147] Sands MB, Dantoc BP, Hartshorn A, et al. Single Question in Delirium (SQiD): testing its efficacy against psychiatrist interview, the Confusion Assessment Method and the Memorial Delirium Assessment Scale. Palliat Med. 2010;24(6):561-565.
[148] Sands MB, Sharma S, Carpenter L, et al. SQiD, the Single Question in Delirium; can a single question help clinicians to detect delirium in hospitalised cancer patients? BMC Cancer. 2021;21(1):75.
[149] Starmark JE, Stålhammar D, Holmgren E. The Reaction Level Scale (RLS85). Manual and guidelines. Acta Neurochir (Wien). 1988;91(1-2):12-20.
[150] 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. 2023;10(3):1367-1374.
[151] 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.
[152] Gliem JA, Gliem RR. Calculating, Interpreting, and Reporting Cronbach's Alpha Reliability Coefficient for Likert-Type Scales 2003. Midwest research-to-practise conference in adult, continuing, and community education. 2003 Vol. 1.
[153] Křemenová Z, Vlčková K. Translation, cultural adaptation, and validation of the Integrated Palliative Outcome Scale-renal (IPOS-r) to Czech. BMC palliat care. 2022;21(1):152.
[154] Costantini M, Rabitti E, Beccaro M, et al. Validity, reliability and responsiveness to change of the Italian palliative care outcome scale: a multicenter study of advanced cancer patients. BMC palliat care. 2016;15:23.
[155] Sakurai H, Miyashita M, Imai K, et al. Validation of the Integrated Palliative care Outcome Scale (IPOS) - Japanese Version. Jpn J Clin Oncol. 2019;49(3):257-262.
[156] Vlckova K, Hoschlova E, Chroustova E, et al. Psychometric properties of the Czech Integrated Palliative Outcome Scale: reliability and content validity analysis. BMC palliat care. 2020;19(1):39.
[157] Long VJE, Cheung YB, Qu D, et al. Validity and reliability of the English and translated Chinese versions of the Integrated Palliative care Outcome Scale (IPOS) in Singapore. BMC palliat care. 2021;20(1):40.

[158] Szeliga M, Kotlińska-Lemieszek A, Jagielski P, et al. Psychometric validation and cross-cultural adaptation of the Integrated Palliative care Outcome Scale in Polish (IPOS-Pol). Palliat Support Care. 2022;20(5):687-693.
[159] Neo SH, Tan JY, Sim DK, et al. Validity and Reliability of the Integrated Palliative Care Outcome Scale in Asian Heart Failure Patients. Palliat Med Rep. 2022;3(1):287-295.
[160] Sirati Nir M, Rassouli M, Ebadi A, et al. Psychometric Properties of the Persian Version of Palliative Care Outcome Scale (POS) in Adult Patients With Cancer. Front Psychol. 2022;13:858684.
[161] Ishii Y, Ito N, Matsumura Y, et al. Validity and reliability of the Integrated Palliative Care Outcome Scale for non-cancer patients. Geriatr Gerontol Int. 2023;23(7):517-523.
[162] Murtagh FE, Ramsenthaler C, Firth A, et al. A brief, patient- and proxy-reported outcome measure in advanced illness: Validity, reliability and responsiveness of the Integrated Palliative care Outcome Scale (IPOS). Palliat Med. 2019;33(8):1045-1057.
[163] Claessens P, Menten J, Schotsmans P, et al. Development and validation of a modified version of the Edmonton Symptom Assessment Scale in a Flemish palliative care population. Am J Hosp Palliat Care. 2011;28(7):475-82.
[164] Chinda M, Jaturapatporn D, Kirshen AJ. Reliability and validity of a Thai version of the Edmonton Symptom Assessment Scale (ESAS-Thai). J Pain Symptom Manage. 2011;42(6):954-60.
[165] Pautex S, Berger A, Chatelain C, et al. Symptom assessment in elderly cancer patients receiving palliative care. Crit Rev Oncol Hematol. 2003;47(3):281-286.
[166] Naik B, Nagaraju SP, Guddattu V, et al. Kannada Translation and Validation of the ESAS-r Renal for Symptom Burden Survey in Patients with End-Stage Kidney Disease. Indian J Palliat Care. 2023;29(2):195-199.
[167] Tricou C, Ruer M, Nekolaichuk C, et al. Transcultural Validation of the French Version of the Modified Edmonton Symptom Assessment Scale: The ESAS 12-F. J Palliat Med. 2019;22(12):1553-1560.
[168] Philip J, Smith WB, Craft P, et al. Concurrent validity of the modified Edmonton Symptom Assessment System with the Rotterdam Symptom Checklist and the Brief Pain Inventory. Support Care Cancer. 1998;6(6):539-541.
[169] Afsar N, Bhuiyan A, Alam A, et al. Validation of Bengali version of Edmonton Symptom Assessment Scale-Revised (ESAS-r Bengali): A multidimensional symptom assessment tool for patients with advanced incurable diseases receiving palliative care. SAGE Open Med. 2023;11:1-6.

[170] Gretarsdottir H, Fridriksdottir N, Gunnarsdottir S. Psychometric Properties of the Icelandic Version of the Revised Edmonton Symptom Assessment Scale. J Pain Symptom Manage. 2016;51(1):133-137.
[171] Kwon JH, Nam SH, Koh S, et al. Validation of the Edmonton Symptom Assessment System in Korean patients with cancer. J Pain Symptom Manage. 2013;46(6):947-956.
[172] Carvajal A, Centeno C, Watson R, et al. A comprehensive study of psychometric properties of the Edmonton Symptom Assessment System (ESAS) in Spanish advanced cancer patients. Eur J Cancer. 2011;47(12):1863-1872.
[173] Dong Y, Chen H, Zheng Y, et al. Psychometric Validation of the Edmonton Symptom Assessment System in Chinese Patients. J Pain Symptom Manage. 2015;50(5):712-717.
[174] Moro C, Brunelli C, Miccinesi G, et al. Edmonton symptom assessment scale: Italian validation in two palliative care settings. Support Care Cancer. 2006;14(1):30-37.
[175] Ripamonti C, Leporati R, De Feo G, et al. Italian version of the Edmonton Symptom Assessment System (ESAS)-Total Care (TC): development and psychometric validation in patients undergoing cancer treatment or follow-up. Support Care Cancer. 2022;30(3):1923-1933.
[176] Shamieh O, Alarjeh G, Qadire MA, et al. Validation of the Arabic Version of the Edmonton Symptom Assessment System. Int J Environ Res Public Health. 2023;20(3):2571.
[177] Paiva CE, Manfredini LL, Paiva BS, et al. The Brazilian Version of the Edmonton Symptom Assessment System (ESAS) is a Feasible, Valid and Reliable Instrument for the Measurement of Symptoms in Advanced Cancer Patients. PLoS One. 2015;10(7):e0132073.
[178] Dano S, Pokarowski M, Liao B, et al. Evaluating symptom burden in kidney transplant recipients: validation of the revised Edmonton Symptom Assessment System for kidney transplant recipients - a single-center, cross-sectional study. Transpl Int. 2020;33(4):423-436.
[179] Yokomichi N, Morita T, Nitto A, et al. Validation of the Japanese Version of the Edmonton Symptom Assessment System-Revised. J Pain Symptom Manage. 2015;50(5):718-723.
[180] McDermid F, Peters K, Jackson D, Daly J. Conducting qualitative research in the context of pre-existing peer and collegial relationships. Nurse Res. 2014;21(5):28-33.
[181] World Medical Association. WMA Declaration of Helsinki – Ethical principles for medical research involving human subjects 2013. [Accessed 2023 August 28, at: https://www.wma.net/policies-post/wma-declaration-of-helsinki-ethical-principles-for-medical-research-involving-human-subjects/].

[182] Bloomer MJ, Hutchinson AM, Brooks L, et al. Dying persons' perspectives on, or experiences of, participating in research: An integrative review. Palliat Med. 2018;32(4):851-860.
[183] White C, Hardy J. What do palliative care patients and their relatives think about research in palliative care?-a systematic review. Support Care Cancer. 2010;18(8):905-911.
[184] World Health Organization. ICD-11 International Classification of Diseases 11th Revision. The global standard for diagnostic health information. 2021. [Accessed 2023 23 October, at: https://icd.who.int/en].
[185] Källman U, Kimberly LB, Bååth C. Swedish translation and validation of the international skin tear advisory panel skin tear classification system. Int Wound J. 2019;16(1):13-18.
[186] Kassambara A. Kappa Coefficient Interpretation. 2019. [Accessed 2023 14 February, at: https://www.datanovia.com/en/blog/kappa-coefficient-interpretation/].
[187] International Association of the Study of Pain. IASP Announces Revised Definition of Pain. 2020. [Accessed 2024 27 January, at: https://www.iasp-pain.org/publications/iasp-news/iasp-announces-revised-definition-of-pain/].
[188] Cantón-Habas V, Rich-Ruiz M, Moreno-Casbas MT, et al. Correlation between Biomarkers of Pain in Saliva and PAINAD Scale in Elderly People with Cognitive Impairment and Inability to Communicate. J Clin Med. 2021;10(7)1424.
[189] Knoerl R, Giobbie-Hurder A, Berfield J, et al. Exploring Daily Salivary Cortisol Patterns as Biomarkers of Chronic Chemotherapy-Induced Peripheral Neuropathy Pain. Oncol Nurs Forum. 2022;49(3):207-211.
[190] Harding R, Simon ST, Benalia H, et al. The PRISMA Symposium 1: outcome tool use. Disharmony in European outcomes research for palliative and advanced disease care: too many tools in practice. J Pain Symptom Manage. 2011;42(4):493-500.
[191] 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.
[192] 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.
[193] Farrar JT, Portenoy RK, Berlin JA, et al. Defining the clinically important difference in pain outcome measures. Pain. 2000;88(3):287-294.
[194] 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.

[195] Osoba D, Rodrigues G, Myles J, et al. Interpreting the Significance of Changes in Health-Related Quality-of-Life Scores. J Clin Oncol. 2023;41(35):5345-5350.
[196] Thielking PD. Cancer pain and anxiety. Curr Pain Headache Rep. 2003;7(4):249-261.
[197] Antunes B, Harding R, Higginson IJ. Implementing patient-reported outcome measures in palliative care clinical practice: a systematic review of facilitators and barriers. Palliat Med. 2014;28(2):158-175.
[198] Collingridge Moore D, Payne S, Van den Block L, et al. Strategies for the implementation of palliative care education and organizational interventions in long-term care facilities: A scoping review. Palliat Med. 2020;34(5):558-570.
[199] The Swedish Goverment [Swe: Regeringen]. Språkkrav för personal i äldreomsorgen [Language requirements for staff in elderly care]. 2023. [Accessed 2024 17 January, at: https://www.regeringen.se/contentassets/99f2fd65ffe74be78b6d5ed493b5cb08/sprakkrav-for-personal-i-aldreomsorgen-dir--2023--44.pdf].
[200] National Board of Health and Welfare [Swe: Socialstyrelsen]. Enhetsundersökningen om äldreomsorg och kommunal hälso- och sjukvård. [The unit survey on elderly care and municipal healthcare]. 2023. [Accessed 2024 1 March, at: https://www.socialstyrelsen.se/globalassets/sharepoint-dokument/artikelkatalog/oppna-jamforelser/2023-10-8780.pdf].
[201] Brazil K, Brink P, Kaasalainen S, et al. Knowledge and perceived competence among nurses caring for the dying in long-term care homes. Int J Palliat Nurs. 2012;18(2):77-83.
[202] Ludvigsson C, Isaksson U, Hajdarevic S. Experiencing improved assessment and control of pain in end-of-life care when using the Abbey Pain Scale systematically. Nurs Open. 2020;7(6): 1798-1806.
[203] van Herk R, van Dijk M, Baar FP, et al. Observation scales for pain assessment in older adults with cognitive impairments or communication difficulties. Nurs Res. 2007;56(1):34-43.
[204] International Association of the Study of Pain. Pain Assessment in Dementia. 2019. [Accessed 2024 27 January, at: https://www.iasp-pain.org/resources/fact-sheets/pain-assessment-in-dementia/].
[205] Recchia A, Rizzi B, Favero A, et al. Prevalence of Delirium in End-of-Life Palliative Care Patients: An Observational Study. Med Princ Pract. 2022;31(2):118-124.
[206] Dunne RF, Roussel B, Culakova E, et al. Characterizing cancer cachexia in the geriatric oncology population. J Geriatr Oncol. 2019;10(3):415-419.
[207] Hadjistavropoulos T, Herr K, Prkachin KM, et al. Pain assessment in elderly adults with dementia. Lancet Neurol. 2014;13(12):1216-27.
[208] Helmer LML, Weijenberg RAF, de Vries R, et al. Crying out in pain-A systematic review into the validity of vocalization as an indicator for pain. Eur J Pain. 2020;24(9):1703-1715.

[209] Lautenbacher S, Kunz M. Facial Pain Expression in Dementia: A Review of the Experimental and Clinical Evidence. Curr Alzheimer Res. 2017;14(5):501-505.
[210] Lautenbacher S, Walz AL, Kunz M. Using observational facial descriptors to infer pain in persons with and without dementia. BMC geriatrics. 2018;18(1):88.
[211] Kunz M, Meixner D, Lautenbacher S. Facial muscle movements encoding pain-a systematic review. Pain. 2019;160(3):535-549.
[212] Atee M, Hoti K, Chivers P, et al. Faces of Pain in Dementia: Learnings From a Real-World Study Using a Technology-Enabled Pain Assessment Tool. Front Pain Res. 2022;3:827551.
[213] Lautenbacher S, Hassan T, Seuss D, et al. Automatic Coding of Facial Expressions of Pain: Are We There Yet? Pain Res Manag. 2022;6635496.
[214] Babicova I, Cross A, Forman D, et al. Evaluation of the Psychometric Properties of PainChek® in UK Aged Care Residents with advanced dementia. BMC geriatrics. 2021;21(1):337.
[215] Fry M, Arendts G, Chenoweth L. Emergency nurses' evaluation of observational pain assessment tools for older people with cognitive impairment. J Clin Nurs. 2017;26(9):1281-1290.
[216] Fry M, Chenoweth L, MacGregor C, et al. Emergency nurses perceptions of the role of family/carers in caring for cognitively impaired older persons in pain: a descriptive qualitative study. Int J Nurs Stud. 2015;52(8):1323-1331.
[217] Sommers MS. Color awareness: A must for patient assessment. 2011. [Accessed 2023 24 November, at: https://www.americannursetoday.com/color-awareness-a-must-for-patient-assessment/].
[218] Web of Science. Citations of Guidelines for the process of cross-cultural adaptation of self-report measures. 2024. [Accessed 2024 7 February, at: https://www.webofscience.com/wos/woscc/full-record/WOS:000166243700013].
[219] Hsieh HF, Shannon SE. Three approaches to qualitative content analysis. Qual Health Res. 2005;15(9):1277-1288.
[220] Hopkins WG. Measures of reliability in sports medicine and science. Sports Med. 2000;30(1):1-15.