ORIGINAL RESEARCH

A novel pain assessment tool incorporating automated facial analysis: interrater reliability in advanced dementia

Mustafa Atee[1]
Kreshnik Hoti[1,2]
Richard Parsons[1]
Jeffery D Hughes[1]

[1]School of Pharmacy and Biomedical Sciences, Faculty of Health Sciences, Curtin University, Bentley, WA, Australia; [2]Division of Pharmacy, Faculty of Medicine, University of Pristina, Prishtina, Kosovo

Objectives: Regardless of its severity, dementia does not negate the experience of pain. Rather, dementia hinders self-reporting mechanisms in affected individuals because they lose the ability to do so. The primary aim of this study was to examine the interrater reliability of the electronic Pain Assessment Tool (ePAT) among raters when assessing pain in residents with moderate-to-severe dementia. Secondly, it sought to examine the relationship between total instrument scores and facial scores, as determined by automated facial expression analysis.

Study design: A 2-week observational study.

Setting: An accredited, high-care, and dementia-specific residential aged care facility in Perth, Western Australia.

Participants: Subjects were 10 residents (age range: 63.1–84.4 years old) predominantly with severe dementia (Dementia Severity Rating Scale score: 46.3±8.4) rated for pain by 11 aged care staff. Raters (female: 82%; mean age: 44.1±12.6 years) consisted of one clinical nurse, four registered nurses, five enrolled nurses, and one care worker.

Measurements: ePAT measured pain using automated detection of facial action codes and recordings of pain behaviors.

Results: A total of 76 assessments (rest =38 [n=19 pairs], movement =38 [n=19 pairs]) were conducted. At rest, raters' agreement was excellent on overall total scores (coefficient of concordance =0.92 [95% CI: 0.85–0.96]) and broad category scores (κ=1.0). Agreement was moderate (κ=0.59) on categorical scores upon movement, while it was exact in 68.4% of the cases. Agreement in actual pain category scores gave κ_w=0.72 (95% CI: 0.58–0.86) at rest and κ_w=0.69 (95% CI: 0.50–0.87) with movement. All raters scored residents with higher total scores post-mobilization compared to rest. More facial action unit codes were also detected during pain (mean: 2.5 vs 1.9; $p<0.0012$) and following mobilization (mean: 2.5 vs 1.7; $p<0.0001$) compared to no pain and rest, respectively.

Conclusions: ePAT, which combines automated facial expression analysis and clinical behavioral indicators in a single observational pain assessment tool, demonstrates good reliability properties, which supports its appropriateness for use in residents with advanced dementia.

Keywords: interrater reliability, ePAT, pain, pain assessment, PainChek®, dementia, facial action units, automated facial expression analysis, total pain scores

Correspondence: Mustafa Atee
School of Pharmacy and Biomedical Sciences, Curtin University, PO Box U1987, Perth, WA 6845, Australia
Tel +61 8 9266 7369
Fax +61 8 9266 2769
Email mustafa.atee@curtin.edu.au

Introduction

In residential aged care facilities (RACFs), dementia is common in >50% of residents.[1–3] Up to 97% of individuals with advanced dementia exhibit behavioral (eg, aggression) and psychological (eg, anxiety) symptoms that lead to poor quality of life in this setting.[4] More than 90% of aged care staff had been exposed to physical

submit your manuscript | www.dovepress.com

Dovepress

[Facebook] [Twitter] [LinkedIn] [YouTube]

http://dx.doi.org/10.2147/CIA.S168024

Clinical Interventions in Aging 2018:13 1245–1258

1245

or emotional aggression from residents.[5] This problem is in part due to lack of self-report and inadequate identification of pain resulting in its subsequent poor management in this vulnerable population.[6] There is strong evidence to suggest that behavioral and psychological symptoms of dementia (BPSD) are often associated with uncontrolled underlying pain from clinical and observational studies.[7–9] Moreover, in a large Swedish cohort (n=120,067) study of older adults (≥75 years) with advanced dementia, 38.6% received at least one medication of questionable benefit including psychotropic drugs in their final year of life.[10] Delayed identification of pain may also influence drug-prescribing patterns. In a recent Italian study of nursing home patients, psychotropic drugs were among the top 10 most commonly prescribed drugs (quetiapine ranked fifth).[11] Pain has also been reported to be significantly associated with BPSD, higher number of antipsychotic prescriptions, reduced quality of life, and premature mortality.[12]

Regardless of its severity, dementia does not negate the experience of pain although there is evidence that pain processing may be altered.[13–15] It is inferred that pain experience might be increased in individuals with Alzheimer's dementia as evident in pain responses recorded from brain activity and facial expressions.[13,16–19] Of particular importance, patients with dementia are more facially expressive than healthy subjects when they display pain.[20] In the absence or lack of self-rating report, facial expressions become an essential component of communicating the existence of pain, particularly for those with dementia.[20] Facial expressions provide instant and brief signals to alert the onlooker. Facial descriptors are also valid indicators in observational pain scales for nonverbal patients with dementia.[21] However, reliability of observers in identifying these descriptors is often low because included items are generic, vague, and not able to be consistently recorded.[22] Further, these descriptors such as "grimacing" in the Pain Assessment Checklist for Seniors with Limited Ability to Communicate (PACSLAC) and Abbey Pain Scale (APS) are not specific to pain as they could overlap with other emotions such as sadness.[23–25] Thus, it has been suggested that objective and comprehensive criteria, such as the Facial Action Coding System (FACS), should be considered among these tools to improve their reliability.[22,26–28] FACS is an anatomical catalog of facial expressions that annotates each individual facial action unit (AU) with a unique numerical label and specific description.[29] Pain-related AUs include eyelid tightening (AU7) and lips parting (AU25). Proficiency in the manual decoding of these AUs requires at least 100 hours of training, while each minute of video requires generally 1 hour of expert's observations.[29] It is, hence, preferred to use automated facial decoding because it reduces the reliance on human rating, which may introduce subjectivity and is likely to be associated with judgment bias.

To address the suboptimal management of pain in people with dementia, novel means of detecting pain in clinical practice are urgently needed. This is because none of the currently available observational pain assessment tools used for people with dementia possess sufficient evidence of validation and reliability to be considered the gold standard.[30] Attempts to integrate computer vision (eg, artificial intelligence or AI) and facial recognition technologies into clinical tools have been made possible with the introduction of smart devices that provide agile platforms for software applications or apps. These intricacies have inspired us to develop the electronic Pain Assessment Tool (ePAT).[30,31,61] In this study, we aimed to examine interrater reliability of ePAT as a means of evaluating pain in aged care residents with moderate-to-severe dementia. Further, we examined the relationship between facial scores (which are determined using automated facial analysis) and total pain scores.

Materials and methods
Ethics
This study is part of a larger clinical trial (Australian New Zealand Clinical Trials Registry Number: ACTRN12616001003460), which was approved by the ethics review board of the participating aged care facility and the Human Research Ethics Committee (HR 10/2014) of Curtin University, Bentley, Western Australia. The study was conducted according to the Declaration of Helsinki, Alzheimer's Australia Guidelines and the Australian National Statement for Ethical Conduct in Human Research.

All participating staff provided written informed consent. For residents, the capacity to consent was determined by the level of cognitive impairment. All residents had moderate-to-severe dementia or cognitive impairment, which makes them incapable of providing consent. Therefore, proxy informed written consent was obtained for each participant (resident) through their authorized and legal representatives. Consent was also given to the publication of images displayed in this manuscript.

Pain measure
The ePAT
The ePAT was designed by Curtin University researchers after reviewing the literature of pain, dementia, geriatric care, and pain facial expressions.[30,31,61] ePAT is a smart device application (App) that uses a combination of a selected set of facial AU codes and common pain behaviors reported in the literature (eg, items included in the American Geriatric
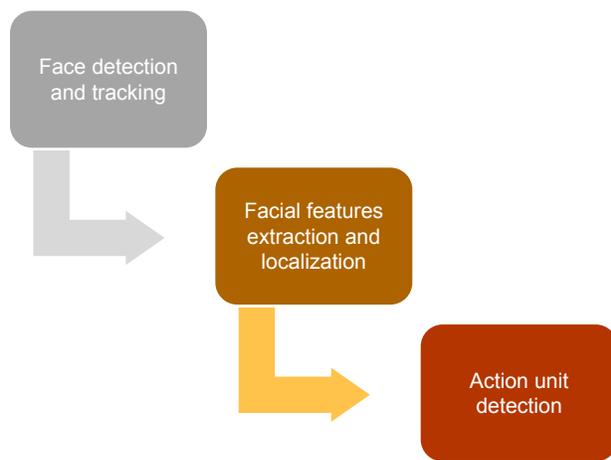
Figure 1 Steps of automated facial analysis in the Face domain of the ePAT to identify pain-related facial action units in patients with dementia.
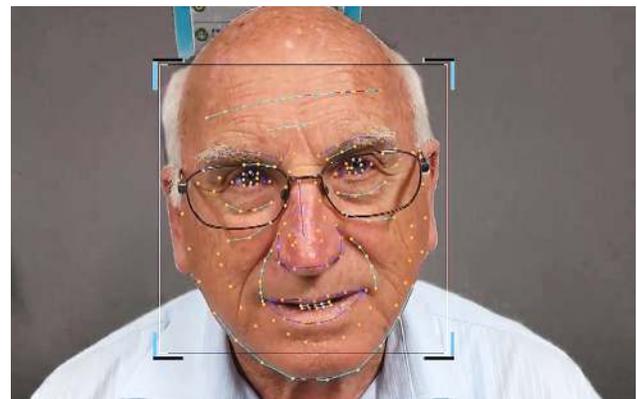**Abbreviation:** ePAT, electronic Pain Assessment Tool.

Figure 3 Automated facial recognition and extraction of facial action units (step 2) using active appearance model and facial landmarking.

Society [AGS] Indicators of Persistent Pain [2002]) to assess pain at the point of care.[29,32] A predefined set of facial AUs were included in the tool because they were associated with pain.[17,20,33–35]

The tool uses digitization, real-time automated facial recognition and decoding using a deep learning (AI) approach, as a means of identifying and evaluating pain.[61] Digitization and smart device technology serve as a platform to facilitate documentation, while automated FACS decoding is integrated in the tool with the view to improve objectivity through reducing human observation errors.[61] Automated facial analysis identifies subtle facial muscle movements called AUs, which represent the smallest building blocks responsible for exerting microexpressions, each of which lasts for 100–500 milliseconds.[36] The automated facial assessment consists of three steps (Figure 1):

1) Face detection and tracking (Figure 2)
2) Localization and extraction of facial features (Figure 3)
3) Detection of facial AUs (Figure 4).

We have tested the ePAT application on Samsung Note 3 (SM-N9005) operating on Android 4.4 KitKat using the lowest available frame per second mode (ie, 30 fps). However, a frame rate of ≤ 5 fps is adequate for the application to perform its facial analysis. The duration of automated facial analysis to process the detection of pain-related facial AUs is ~10 seconds.

The output of the processing is a list of numerical values that represent the confidence level for each AU that we detect. The application will then combine an "x" number of reports obtained for a processed grabbed image to create a consolidated report for the 10 second recordings.

Once detected, facial AUs related to pain are then used in conjunction with other observation-based clinical data (eg, vocalization parameters) recorded by the user to obtain a pain intensity score.

The ePAT is composed of six domains (Face, Voice, Movement, Behavior, Activity, and Body), which contains a total of 42 items.[31,61] Table 1 describes the ePAT domains and the corresponding items along with their operational definitions and primary conceptual basis. Each domain
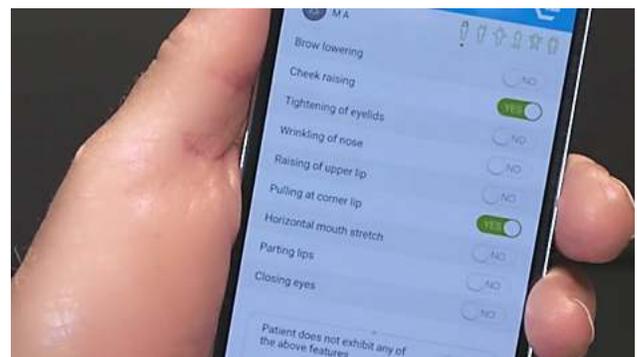

Figure 2 Face detection using the ePAT (step 1).
Abbreviation: ePAT, electronic Pain Assessment Tool.

Figure 4 Detection of facial actions using AU descriptors of FACS (step 3).
Abbreviations: AU, action unit; FACS, Facial Action Coding System.

Table 1 The six domains, corresponding items, and conceptual basis of the ePAT

Domain	Item number	Item description	Operational definition of item	Primary conceptual basis of selected items
Domain 1: The Face	1	Brow lowering	**Facial action unit and the corresponding descriptor** AU4: brow lowerer **Muscular basis** Depressor glabellae, depressor supercilii, and corrugator **Anatomical changes in upper Face (lower central forehead)** • Lowering down of both eyebrows • Movement of eyebrows toward each other • Appearance of vertical or oblique wrinkles between eyebrows in the lower central part of the forehead	FACS[29]
	2	Cheek raising	**Facial action unit and the corresponding descriptor** AU6: cheek raiser **Muscular basis** Outer portion of orbicularis oculi (pars orbitalis) **Anatomical changes in central face (infraorbital region)** • Pulling of skin toward the eye • Pulling the cheeks upward by lifting of the infraorbital triangle • Narrowing the eye aperture and wrinkling the skin below the eye • Appearance of Crow's feet lines or wrinkles • Deepening of the lower eyelid furrow	
	3	Tightening of eyelids	**Facial action unit and the corresponding descriptor** AU7: lid tightener **Muscular basis** Inner portion of orbicularis oculi (pars palpebralis) **Anatomical changes in upper face (orbital region)** • Tightening of the eyelids • Narrowing of the eye aperture • Raising of lower lid	
	4	Wrinkling of nose	**Facial action unit and the corresponding descriptor** AU9: nose wrinkler **Muscular basis** Levator labii superioris alaeque nasi **Anatomical changes in central face** • Pulling of skin upward along the side of the nose toward the root of the nose • Appearance of wrinkles along the side and root of nose • Wrinkling of infraorbital furrow • Lowering the medial portion of the eyebrows • Pulling the center of the upper lip upward	
	5	Raising of upper lip	**Facial action unit and the corresponding descriptor** AU10: upper lip raiser **Muscular basis** Levator labii superioris **Anatomical changes in central-lower face (infraorbital, nasolabial, infranasal and upper lip regions)** • Prominent deepening or wrinkling of infraorbital furrow • Deepening of nasolabial furrow • Pouching at upper lip and nasal passages • Widening and raising of the nostril wings	

(Continued)

Table 1 (*Continued*)

Domain	Item number	Item description	Operational definition of item	Primary conceptual basis of selected items
	6	Pulling at corner lip	**Facial action unit and the corresponding descriptor** AU12: lip corner puller **Muscular basis** Zygomatic major **Anatomical changes in lower face (mouth/lips region)** Drooping or oblique movement of lateral corners of the lips	
	7	Horizontal mouth stretch	**Facial action unit and the corresponding descriptor** AU20: lip stretcher **Muscular basis** Risorius **Anatomical changes in lower face (mouth/lips region)** Bilateral stretch of lips	
	8	Parting lips	**Facial action unit and the corresponding descriptor** AU25: lips part **Muscular basis** Depressor Labii, or relaxation of Mentalis or Orbicularis oris **Anatomical changes in lower face (mouth/lips region)** Relaxed opening of the mouth, that is jaw drop	
	9	Closing eyes	**Facial action unit and the corresponding descriptor** AU43: eye closure **Muscular basis** Relaxation of levator palpebrae superioris **Anatomical changes in upper face (orbital region)** Shutting both eyes for at least half a second	
Domain 2: The Voice	10	Noisy pain sounds, for example, ouch, ah, mm	Sounds or utters related to pain, for example, ouch, ah, mm	AGS (verbalizations and vocalizations)[32]
	11	Requesting help repeatedly	Include one or more of the following: • Expressing numerous verbal requests of help within short periods of time, for example, "help me, help me" • Constant talking • Repetitive use of words or phrases (eg, echophrasia) • Verbal nonsense • Vocalizations with/without discernible meaning Exclude verbal requests for ADL purposes	
	12	Groaning	Using a deep, creaking, or incoherent sound	
	13	Moaning	Producing a long, low-pitched, and inarticulate sound	
	14	Crying	Weeping, sobbing, or whimpering	
	15	Screaming	Using excessively loud voice when communicating (as in shouting or yelling)	
	16	Loud talk	Making louder than normal pitched voice	
	17	Howling	Producing a long wailing cry sound	
	18	Sighing	Breathing in followed by long audible sound upon breathing out	
Domain 3: The Movement	19	Altered or random leg/arm movement	Changed or random movement of any of the limbs	AGS (body movements)
	20	Restlessness	Unable to relax, that is fidgeting	
	21	Freezing	Sudden stiffening, avoiding movement, holding breath	
	22	Guarding/touching body parts	An abnormally stiff, rigid, or interrupted movement while changing position Protecting affected area by holding body part	

(*Continued*)

Table I (*Continued*)

Domain	Item number	Item description	Operational definition of item	Primary conceptual basis of selected items
	23	Moving away	Avoiding being touched or staying away from the interaction	
	24	Abnormal (altered) sitting/ standing/walking	Distorted, asymmetrical or changed sitting, standing (eg, posture), and/or imbalanced gait (eg, limping)	
	25	Pacing/wandering	Roaming restlessly and aimlessly back and forth	
Domain 4: The Behavior	26	Introvert (unsocial) or altered behavior	Being unsocial or socially isolated, that is reluctant to be involved in social activities	AGS (changes in interpersonal interactions, mental status changes)
	27	Verbally offensive	Verbally abusive or aggressive, cursing, swearing, or using foul/insulting language	
	28	Aggressive	Involved in combative or violent behavior	
	29	Fear or extreme dislike of touch, people	Phobias of being touched or interaction with people including family members, other residents, and/or aged care staff	
	30	Inappropriate behavior	Aberrant or socially unacceptable behavior, for example, fiddling	
	31	Confused	Unclear in thinking or understanding, for example, unable to follow instructions or repetitive questioning	
	32	Distressed	Anxious, worried, and agitated	
Domain 5: The Activity	33	Resisting care	Unable to cooperate or become compliant, or refuse to receive care, for example, food, medicine	AGS (changes in activity patterns or routine)
	34	Prolonged resting	Long resting periods without apparent reasons	
	35	Altered sleep cycle	Changed sleep–wake cycle, for example, long sleeps during the day	
	36	Altered routines	Changed the order or timing of activities from the norm	
Domain 6: The Body	37	Profuse sweating	Excessive sweating in various parts of the body excluding circumstances due to environmental factors such as no air conditioning or lack of proper ventilation	AGS
	38	Pale/flushed (red) face	Color faded or red-colored face	
	39	Feverish/cold	Changes in body temperature either too hot or too cold	
	40	Rapid breathing	Fast rate of breathing	
	41	Painful injuries	Injuries are known to induce pain, for example, falls, bed sores, active wounds	
	42	Painful medical conditions	Conditions known to cause pain including currently presented, for example, dental infections, urinary tract infections, or previously documented chronic conditions in medical history, for example, arthritis	

Abbreviations: ADL, activities of daily living; AGS, American Geriatric Society; AU, action unit; ePAT, electronic Pain Assessment Tool; FACS, Facial Action Coding System.

contains a number of items. For example, the Face domain (Domain 1) consists of nine descriptors, which correspond to AUs 4, 6, 7, 9, 10, 12, 20, 25, and 43, automatically recognized by the App. Other domains (Domains 2–6) are based around descriptors drawn from the literature (eg, the AGS Indicators, other observational pain scales, recommendations by Herr et al, and Pasero and McCaffrey's Hierarchy of Pain Assessment Techniques).[32,37,38] The latter is also supported by American Society for Pain Management Nursing recommendations about patients unable to self-report.[37,39]

Each domain provides a checklist of pain indicators from which the user makes binary selections (ie, present yes/no) for each indicator on the smart device touch screen based on clinical observations of the patient.[31,61] Domain scores are automatically calculated to provide a final pain score. Based on the published results of the validation study where ePAT scores were compared to APS cutoff scores (no pain: 0–2; mild pain: 3–7, moderate pain: 8–13, severe pain: 14 or more), the following categorical ratings have been derived: no pain: 0–6; mild pain: 7–11; moderate pain: 12–15; severe pain: 16 or more.[31] Each domain also has a blank field at the

bottom of the screen for the user to record any additional and/or relevant observation(s).[31,61]

The Face domain (AU score) of ePAT was blindly evaluated against self-reporting (gold standard) measures (visual analog scale [VAS], numerical rating scale [NRS], and verbal rating scale [VRS]) of cognitively intact people with chronic pain (n=43 [21 male, 22 female], mean age=54±14 years) in unpublished study.[40] When the AU score was classified into two groups (0–2 vs 3 or more), it was highly correlated with the gold standard measures of pain (t-tests or Wilcoxon: $p<0.0001$ for each measure). These measures were then classified into two groups (low or high pain) as follows: VAS: 0–50 vs 51–100; VRS: 0–3 vs 3.5–5; NRS: 0–4 vs 5–10. Cross tabulations of the categorized AU score against these binary variables showed that a high AU score had over 95% sensitivity to identify high pain scores and high specificities (69%, 90%, and 95% for each measure, respectively). Participants were classified into those recording high pain on any of the three validated measures vs low pain on all measures. The AU score was able to identify high pain with 95.7% sensitivity and 95% specificity.[40]

In a published study by Atee et al, the complete ePAT tool was tested in 40 residents (aged 60–98 years) with moderate-to-severe dementia (Psychogeriatric Assessment Scale–Cognitive Impairment Scale scores: 10–21) from three RACFs in Western Australia. Based on 353 paired pain assessments, the tool demonstrated excellent concurrent validity (r=0.882, 95% CI: 0.857–0.903), good discriminant validity (random regression model is not timing-dependent, p=0.795), good interrater reliability (weighted κ=0.74, 95% CI: 0.68–0.80), and excellent internal consistency (Cronbach's alpha (α)=0.925).[31] This observational study assessed the psychometric properties of ePAT compared to the APS, which is the widely used observational pain scale for people with dementia in Australia.[31] These findings were also confirmed in another cohort of people (n=34, 68.0–93.2 years old) with similar demographics.[41]

Setting

Single-site, accredited, high-care, and dementia-specific RACF. The facility has a capacity of 65 beds and is located in Perth, Western Australia.

Participants

Pain raters (users)

Raters were aged care staff working in the facility using the ePAT as an assessment scale of pain. Staff were recruited if they had been working for 3 months or more in the facility, were familiar with residents, able to converse in English, and keen to participate in the study. Staff were excluded if they had fears associated with using technologies or were likely to be absent for any period during the study.

Residents (subjects)

Residents were included in the study if they had moderate-to-severe dementia as indicated by Dementia Severity Rating Scale (DSRS) scores >18, and had documented behavioral problems and a history of painful conditions. Patients were excluded from the study if they were deemed medically unfit for participation.

Protocol

This substudy was a 2-week observational study, in which a convenience sampling technique was employed. Staff who consented to participate attended an education and training program prior to the study. The program involved a single session, which was conducted by the principal investigator over 4.5 hours at the study site. Attendees received education about pain, pain and dementia, pain assessment, and pain management. The contents of the program were developed after reviewing the International Association for the Study of Pain Curricula[42] and current literature with modifications made appropriate to the setting and demographics of raters. The session also included a demonstration of the ePAT and practical training on its use.

Staff rater data were collected using a 14-item questionnaire, which included a mix of open- and closed-ended questions. The questionnaire was piloted using five test subjects prior to administration to ensure readability and ease of completion.

Testing of the ePAT was undertaken indoors at the RACF in September 2016. Testing involved the use of the ePAT by pairs of independent staff raters who were blinded to each other's assessments, scores, and to the use of analgesics. Raters were instructed to conduct their assessments independently using own ePAT device without consulting or conversing with the other rater involved in the study. No discussions were made regarding each assessment, and scores obtained were not shared nor exchanged between paired raters. One of the study authors monitored the data collection process to ensure that this was being followed throughout the study.

Automated facial analyses were conducted consecutively to allow each rater access to a full frontal view of the resident and prevent any possible discrepancies (eg, physical hindrance) that might arise during the process.

Paired ratings were scheduled randomly to reduce learning bias and subsequent systematic error. Ratings also occurred within a time frame of 2–3 minutes to ensure that the results obtained were comparable. As far as possible, recording conditions of automated facial analyses (eg, lighting, distance from subject) were essentially the same for all cases.

Residents with dementia were assessed for pain during routine nursing activities or activities of daily living (ADL) that involved mobilization and during periods of rest. Over the study period (ie, 2 weeks), each resident was assessed by two different raters on four separate occasions: on each of the 2 days ~1 week apart, the assessors rated the resident's pain while at rest and shortly afterward while receiving care activities. Raters were instructed to observe the resident under the assessment for pain-induced behaviors for at least 5 minutes before commencing pain scoring on the ePAT.

Statistical analyses

Descriptive statistics (eg, mean, range, standard deviation) were used to summarize the profiles of the raters, residents, and pain scores including automated facial scores. Agreement on categorical pain data was evaluated using kappa statistics. The kappa coefficient measures interrater reliability or the agreement between two observers and takes into account the agreement expected by chance. It is, therefore, a more robust measure than percentage agreement.[43] A value of 0.6 or above indicates moderate agreement or good interrater reliability.[43] Cohen's kappa (κ) statistic was used to assess agreement between raters on the presence or absence of pain, whereas weighted kappa (κ_w) was employed to evaluate agreement when pain was divided into >2 categories. Agreement on continuous pain data (ie, total pain scores) was measured by Lin's concordance correlation coefficient (CCC).[44] Values of CCC range from 0 to ±1 where +1 is perfect concordance and −1 is perfect discordance. To assess the strength of agreement, we used Altman's criteria as a guide to interpret CCC values: <0.20="poor" and >0.80="excellent".[45] Using a published chart of the score range of the ePAT, total pain scores were allocated into broad pain categories: no pain (0–6), mild pain (7–11), moderate pain (12–15), and severe pain (16–42).[31] Further, a regression model was used to examine the relationship between automated facial scores and total instrument scores (pain vs no pain) of ePAT under various testing conditions.

Level of significance was expressed by 95% CI range or p-value <0.05. All data were analyzed using the Statistical Package for the Social Sciences (SPSS), Version 24 Software (SPSS Inc., Apache Software Foundation, Chicago, IL, USA).

Results

Demographic data

Demographics of resident sample

Ten residents with an age range of 63.1–84.4 years (mean: 74.4±5.9 years) were recruited into the study. The gender ratio of residents was 50:50 and the vast majority (90%) were Caucasians. Half of the residents had Alzheimer's dementia and 80% were classified as having severe dementia (mean DSRS score: 46.3±8.4). Table 2 provides demographic characteristics of resident sample.

Movement-based activities ranged from independent (eg, walking) to assisted (eg, transfer) events.

Demographics of rater cohort

A cohort of 11 staff with a mean age of 45.3±13.4 years were recruited into the study, two of whom were male. Working hours in the facility ranged from 20 to 38 hours per week with five staff employed as fulltime (ie, 38 hours/week). The average length of staff employment in the facility was 10.6±9.1 years. The cohort included 10 nurses of various hierarchical roles (one clinical nurse, four registered nurses, and five enrolled nurses), plus a trained carer. Range of nursing or caring experience among staff varied from 1 to 30 years, while aged care experience was 1–33 years. The mean years of experience in cognitive impairment or dementia care were 11.5±7.9 years.

Table 2 Demographic characteristics of resident sample at baseline (n=10)

Characteristics	Number (%)	Mean (SD)
Age (years) (range: 63.1–84.4)		
Female		78.2 (5.2)
Male		70.6 (5.0)
Overall		74.4 (5.9)
Gender		
Female	5 (50)	
Male	5 (50)	
Ethnicity		
Caucasian	9 (90)	
Other	1 (10)	
Primary language		
English	10 (100)	
Mobility		
Limited	4 (40)	
Immobile	4 (40)	
Bedridden	2 (20)	
Dementia severity (DSRS range: 30–54)		46.3 (8.4)
Diagnosis of dementia		
Alzheimer's disease	5 (50%)	
Alcoholic-related dementia	2 (20%)	
Frontotemporal dementia	1 (10%)	
Parkinson's dementia	1 (10%)	
Vascular dementia	1 (10%)	

Abbreviations: DSRS, Dementia Severity Rating Scale; SD, standard deviation.

Table 3 Demographic characteristics of rater cohort (n=11)

Characteristics	Number (%)	Mean (SD)
Age (years)		45.3 (13.4)
Gender (female)	9 (81.8)	
Ethnicity		
Caucasian	6 (54.5)	
Asian	3 (27.3)	
Other	2 (18.2)	
Primary language		
English	8 (72.7)	
Other	3 (27.3)	
Employment status (hours)		33.3 (7.2)
Part time (range: 20–26)	6 (54.5)	
Full time (38 hours)	5 (45.5)	
Years of experience		
Nursing/caregiving		15.5 (11.8)
Aged care		15.1 (11.1)
Cognitive impairments/dementia care		11.5 (7.9)
Employment in facility		10.6 (9.1)
Role in facility		
Enrolled nurse	5 (45.4)	
Registered nurse	4 (36.4)	
Clinical nurse	1 (9.1)	
Personal care worker	1 (9.1)	
Past pain education		
Yes	11 (100)	
No	0 (0)	
Last received pain education		
<3 months	1 (9.1)	
<12 months	1 (9.1)	
>12 months	1 (9.1)	
>3 years	1 (9.1)	
Not specified	7 (63.6)	

Abbrevaition: SD, standard deviation.

All staff reported receiving pain education in the past. Demographics of raters are shown in Table 3.

Pain data

All residents had four pairs of ePAT ratings over the 2-week study period except one resident who had only two pairs during the same period. This resulted in a total of 76 assessments for the sample. Of these, almost two-thirds (65.8%) were scored as "no pain" while less than a third (29%) scored "mild pain" as shown in Table 4. Pain-associated conditions documented for residents were diverse with 80% of the sample having two or more chronic painful conditions. Residents had a mean pain score of 5.6±3.5 (median=5) with a range of 1–18. Table 4 provides a description of pain-related data in residents who underwent pain assessment using the ePAT.

Interrater reliability data of the ePAT instrument

Kappa statistics

Rater agreement in broad categories of pain (no pain, mild, moderate, or severe pain) using kappa statistics was classified

Table 4 Pain-related data of residents (n=10) who underwent pain assessment using the ePAT

Variables	Number (%)	Mean (SD)
Pain assessments	76 (100)	7.6 (1.3)
Rest	38 (50)	3.8 (0.6)
Movement	38 (50)	3.8 (0.6)
Pain scores (median: 5, range: 1–18)		5.6 (3.5)
Pain categories		
No pain	50 (65.8)	
Mild pain	22 (29)	
Moderate pain	2 (2.6)	
Severe pain	2 (2.6)	
Number of documented chronic painful diagnoses (median: 3, range 1–5)		3.0 (1.6)
1	2 (20)	
2	4 (40)	
3	1 (10)	
4	1 (10)	
5	2 (20)	
Prescribed analgesia		
Regular		2.0 (1.5)
PRN		0.4 (0.8)
Non-opioid analgesics		
Regular		
Celecoxib capsules 100 mg	1 (10)	
Diclofenac gel 11.6 mg/g	1 (10)	
Paracetamol tablets 500 mg	5 (50)	
Paracetamol tablets SR 665 mg	1 (10)	
Paracetamol oral liquid 240 mg/5 mL	1 (10)	
PRN		
Paracetamol tablets 500 mg	2 (20)	
Opioid analgesics		
Regular		
Fentanyl patches 12 mcg/hour	3 (30)	
Oxycodone tablets 5 mg	2 (20)	
PRN		
Oxycodone tablets 5 mg	1 (10)	

Abbreviations: ePAT, electronic Pain Assessment Tool; PRN, pro re nata [when necessary]; SD, standard deviation.

as excellent (κ=1.0) at rest, where both raters agreed on the absence of pain on 17 occasions, and mild pain on two occasions (Table 5). With movement, agreement was moderate (κ=0.59), but assessments were in complete agreement for 13 (68.4%) out of the 19 paired assessments; the remaining six pairs differed only by one category.

Lin's concordance analysis

Lin's concordance correlation coefficient (CCC) was used to calculate agreement between total score values produced by the paired raters. The value of CCC was calculated to be 0.92 (95% CI: 0.85–0.96), which is classified as an excellent agreement.[46]

Means and standard deviations of total pain scores and facial scores of ePAT at various occasions

The difference between the pairs of measurements in producing raw total pain scores while performed at rest and

Table 5 Agreement between raters in their assessments of total pain: kappa statistics for total pain scores within pain categories, and using raw pain scores

Activity	Broad pain categories Cohen's kappa (κ)	95% CI	Raw total pain scores weighted kappa (κ$_w$)	95% CI
Rest	1.0	Not applicable	0.72	0.58–0.86
Movement	0.59	0.27–0.91	0.69	0.50–0.87

Note: The figures are based on the 19 pairs of measurements with movement and 19 pairs without (ie, at rest).

with movement appeared to be small, as suggested by the kappa statistics (Table 5). A linear model confirmed that the agreement was very good, and that the agreement did not depend on the conditions (rest vs movement; $p=0.91$). In this model, the resident identifier was named as a random effect, the dependent variable was the difference in measurements made by the two raters on each occasion, and the independent variable was the condition (rest/movement). Because the p-value associated with "condition" was not significant, this suggested that the agreement between raters was similar for both conditions. In addition, the intercept obtained from the model (overall mean) was close to zero (0.05; $p=0.87$), suggesting that there was no consistent bias between raters.

The mean of the pain assessments made on each occasion by the two raters was calculated (n=38 occasions), and these were entered into a random-effects model to compare the measurements made at rest with those taken with movement. The model showed that the mean scores seen with movement (7.3±3.7) were significantly higher than those observed at rest (4.0±2.2; $p<0.0001$; standard error [SE] estimated from the regression model: 0.81). Similarly, the scores on the Face domain were significantly different ($p<0.0001$; SE: 0.17) between those taken with movement (mean: 2.5±0.6) and those at rest (mean: 1.7±0.7). These data are presented in Table 6.

Table 6 Means and standard deviations of total pain scores and facial scores of ePAT at various occasions

	N (%)	Total ePAT scores (mean±SD)	Automated facial (AU) scores (mean±SD)
Episode			
No pain	50 (65.8)	3.6±1.7[a]	1.9±0.8[b]
Pain	26 (34.2)	9.3±2.9[a]	2.5±0.6[b]
Occasion			
Rest	38 (50%)	4.0±2.2[c]	1.7±0.7[c]
Movement	38 (50%)	7.3±3.7[c]	2.5±0.6[c]

Notes: [a]The two-tailed p-value is <0.0001. [b]The two-tailed p-value equals 0.0012. [c]The two-tailed p-value is <0.0001.
Abbreviations: AU, action unit; ePAT, electronic Pain Assessment Tool; SD, standard deviation.

Discussion

Our study aimed to investigate the reliability of a new tool, named the ePAT in individuals with moderate-to-severe dementia living in RACFs. The tool takes advantage of advanced computational capabilities together with the cameras in smart devices and automated facial recognition technology to identify the presence and severity of pain.[31,61]

Our findings suggested that agreement between raters was greater during rest (Table 5) because fewer behaviors were observed and hence recorded. In contrast, during movement, pain-induced behaviors are likely to increase because of the experienced nociceptive stimuli associated with movement-related activities (eg, turning).[47,48] Some pain behaviors incorporated in ePAT are subtle and difficult to identify by raters, which may contribute to some degree of interrater variability.[49] This variation in agreement between rest and movement is consistent with other observational pain assessment tools as indicated by these kappa value ranges: The Checklist of Nonverbal Pain Indicators (CNPI): 0.625–0.819, Mahoney Pain Scale (MPS): 0.55–0.77, Mobilization-Observation-Behavior-Intensity-Dementia-2 (MOBID-2): 0.44–0.90 for observed behaviors.[50]

With regard to automated facial scores produced by the ePAT, the mean values were significantly higher ($p<0.001$) for "pain" events compared to those recorded under "no pain" (Table 6). This indicates that the facial AUs are sensitive to aversive events that trigger painful stimuli. This was also supported by Lints-Martindale et al in a study investigating facial reactions to experimentally induced pain stimuli.[35] They found that noxious electrical stimuli produce much greater FACS activity with "pain" compared to "no pain" events.[35] Automated software scoring is also a reliable way of recognizing expressions in comparison with human observation. In a study by Peter Lewinski, automated facial coding software (eg, FaceReader) outperformed human observers in recognizing neutral faces by 31%.[51] Bartlett et al also found that automated decoding of facial expressions was far superior (85% accuracy) in identifying genuine from fake pain compared to untrained and trained observers (50% and 55%, respectively).[52]

A significantly higher mean facial score using automated facial analysis was also observed on movement compared to that during rest (Table 6). Research suggests that joint movements generate shear forces on the axolemma of the "free" nerve endings resulting in nociceptive signals as pain.[53] Our results are similar to those of Hadjistavropoulos et al who found that more facial activity was produced in movement activities.[17] They reported that the FACS score was significantly greater during walking compared to reclining or transferring.[17] In addition, the difference in the average number of AUs detected for residents after movement was significantly greater ($p<0.0001$) than for those at rest.

Current research suggests that integrating automated FACS descriptors with observational tools is psychometrically sound and clinically useful.[27,54] Beach et al also support this endeavor, reporting that pain-relevant FACS scores and modified PAINAD scale scores were highly correlated in older adults with Alzheimer's dementia.[54] Pain-related FACS was also found to be clinically relevant for inclusion in observational pain assessment scales designed for people with dementia.[54] Observational tools with pain-related AUs have also shown higher sensitivity and better psychometric properties than those that contain generic facial expression descriptors.[21]

Lin's CCC is a relative index of reliability where agreements on total pain numerical scores are compared. Our statistical analysis showed an excellent agreement (CCC=0.92). As far as we know, there are no CCC values of pain assessment scales in dementia reported in the literature. CCC values were previously reported for observational pain scales in other noncommunicative populations such as the Nonverbal Pain Assessment Tool (NPAT): 0.21–0.72 (95% CI).[55] Our results demonstrated higher values than NPAT.

Our study has tested the interrater reliability of a novel tool that integrates pain-relevant FACS items (ie, facial AU codes) with other communicative (eg, vocalization items), protective (eg, guarding), and subtle (eg, resistance to care) pain behaviors. This approach has emerging support in the literature.[27,54,56] The total number of pain behaviors is also significantly related to self-reported pain intensity in older adults.[57] It is essential to highlight that using an observational pain assessment tool improves detection of presence and severity of pain in people with cognitive impairment.[58] Further, ePAT uses automated facial recognition and analysis to detect pain-relevant AU codes.[31,61] Given that patients with dementia have an enhanced facial activity as illustrated in previous studies[16,20,35] and that observational tools improve pain recognition in this population,[58] we believe that ePAT can facilitate the process of pain detection in these patients.

Strengths and limitations

This study had several merits and limitations. Generalizations to other settings and populations are limited by the sampling method (ie, convenience, purposive sampling) and sample size. Therefore, the risk of committing Type II errors in this study remains a possibility. Despite the small sample size, an equal number of pain assessments were performed on most (ie, 9 out of 10) residents. The resident cohort was homogenous although it lacked ethnic diversity. Gender and cultural disparities were only evident in the rater group. This group had a diverse range of skills representative of the hierarchical workforce in the residential aged care setting. Learning effect associated with repeated use of the tool on the same subject is inevitable in agreement studies. The short time frame of the study may have influenced how raters remembered pain-related behaviors and how they may carry forward this information to the following week because of memory bias. Pain assessments were delivered during clinical rounds while residents were receiving their standard care, in order to minimize interference to workflow. As such, this perhaps contributed to variations in pain scores, which are associated with consecutive delivery of the assessments, individual observation skills of a rater to record nonfacial pain-related behaviors, and the general subjective nature of pain. In addition, assessments were delivered during ADL, such as walking, to provide a real-world context of actual use of the tool in clinical scenarios.

In the current study, we tested the interrater reliability by comparing two ePAT users. Head-to-head comparative studies of observational pain assessment scales can provide valuable data to guide the process of tool development and refinement.[56] Further, interrater reliability is one of the key psychometric properties of observational scales because arriving at similar pain scores by different clinicians provides confidence in the tested tool.[27] In our study, there was a small number (26 out of 76) of "pain" cases detected, perhaps due to the adequate pain management in the sample. Although we acknowledge this limitation, we believe the testing discussed here is sufficient to address the objectives of the study. In fact, identifying "pain" from "no pain" or neutral cases consistently is considered a useful criterion of reliability in judgment studies.[27] Our findings were based on clinical validation (ie, clinical pain from ADL) and, therefore, results obtained from experimental studies (pressure or temperature pain-induced modalities) may vary. Notwithstanding, there is some evidence that experimental pain response is different from clinical pain response and that the predictive value of experimental pain for clinically induced pain is weak and not

reliable.[59] Lichtner et al in their systematic meta-review and Closs et al in their meta-review recommended that validation work should be conducted in clinical settings, so that it informs the applicability of the tool and its potential value in everyday clinical practice.[50,60] This is because pain assessment tools that are experimentally tested in research do not necessarily transfer easily and effectively in clinical settings.[60]

In the study design, we allowed access of raters to all available information (except for analgesics) to minimize the chances of underestimating pain. An equal access to medical profiles by both raters means that raters were well informed about the patients' diagnoses of possible painful chronic conditions. This strategy may have strengthened raters' evaluation when conducting clinical pain assessments.

Another strength of the study is that various reliability measures were used including kappa, weighted kappa, and CCC. Reliability statistics that consider chance agreement between raters will account for the variation in frequency of AUs distribution. This is important because it will assist in extrapolating the findings into other populations. However, measurement errors are still possible because of confounding effects linked to uncontrolled conditions inside the aged care facility such as lighting, shadowing, and random movement that might have affected the performance of the tool.

Conclusion

Facial scores were significantly higher during "pain" compared to those scores clinically recorded as "no pain". Similarly, automated scoring of facial AUs was higher for residents with movement compared to rest. This indicates that the Face domain of the ePAT has a good sensitivity to the presence of pain. Combining automated facial expression analysis and clinical behavioral indicators in a single observational pain assessment scale affords ePAT good reliability properties. This supports its appropriateness for use in nonverbal residents with advanced dementia. Reliable clinical tools particularly for pain assessment are desired to improve therapeutic outcomes. It should be stressed, however, that currently there is no gold standard pain assessment tool available for noncommunicative people with dementia, and any attempt to work toward this goal must be encouraged. Innovative approaches of pain assessment such as those included in the ePAT can assist clinicians to more objectively assess pain in challenging populations, such as those with dementia.

Author contributions

MA, KH, and JDH conceived the idea, designed the study, and wrote the protocol. MA conducted the literature search, recruited the subjects and collected the data, and organized and written first draft of the manuscript. RP and MA wrote statistical methods. RP conducted the statistical analyses. All authors contributed toward data analysis, drafting and critically revising the paper, gave final approval of the version to be published, and agree to be accountable for all aspects of the work.

Disclosure

MA, KH, and JDH are shareholders in PainChek Ltd (previously known as EPAT Technologies Ltd), which is commercializing the ePAT instrument as PainChek®. They also have a patent application titled "A pain assessment method and system" (PCT/AU2015/000501), which is currently under national phase examination since February 2, 2017. MA is a research scientist for PainChek Ltd and is a research fellow and PhD candidate with the School of Pharmacy and Biomedical Sciences, Curtin University. KH is employed as a consultant by PainChek Ltd and is an assistant professor at the University of Pristina. JDH holds the position of chief scientific officer of PainChek Ltd and is a professor in the School of Pharmacy and Biomedical Sciences, Curtin University. RP has no competing or financial interest in PainChek Ltd. The authors report no other conflicts of interest in this work.

References

1. Australian Institute of Health and Welfare (AIHW). *Residential Aged Care in Australia 2010–11: A Statistical Overview. Cat. no. AGE 68.* Canberra: AIHW; 2012.

2. Harris-Kojetin L, Sengupta M, Park-Lee E, Valverde R. Long-term care services in the United States: 2013 overview. *National Health Care Statistics Reports*. Vol 1. Hyattsville: National Center for Health Statistics; 2013.

3. Hoffmann F, Kaduszkiewicz H, Glaeske G, van den Bussche H, Koller D. Prevalence of dementia in nursing home and community-dwelling older adults in Germany. *Aging Clin Exp Res*. 2014;26(5):555–559.

4. Steinberg M, Shao H, Zandi P, et al. Point and 5-year period prevalence of neuropsychiatric symptoms in dementia: the Cache County Study. *Int J Geriatr Psychiatry*. 2008;23(2):170–177.

5. Rodwell J, Demir D. Addressing workplace violence among nurses who care for the elderly. *J Nurs Admin*. 2014;44(3):152–157.

6. Achterberg WP, Pieper MJC, Dalen-Kok AH, et al. Pain management in patients with dementia. *Clin Interv Aging*. 2013;8:1471–1482.

7. Husebo BS, Ballard C, Sandvik R, Nilsen OB, Aarsland D. Efficacy of treating pain to reduce behavioural disturbances in residents of nursing homes with dementia: cluster randomised clinical trial. *BMJ*. 2011;343:d4065.

8. Sampson EL, White N, Lord K, et al. Pain, agitation, and behavioural problems in people with dementia admitted to general hospital wards: a longitudinal cohort study. *Pain*. 2015;156(4):675–683.

9. Tosato M, Lukas A, van der Roest HG, et al. Association of pain with behavioral and psychiatric symptoms among nursing home residents with cognitive impairment: results from the SHELTER study. *Pain*. 2012;153(2):305–310.

10. Morin L, Vetrano DL, Grande G, Fratiglioni L, Fastbom J, Johnell K. Use of medications of questionable benefit during the last year of life of older adults with dementia. *J Am Med Dir Assoc*. 2017;18(6): 551. e1–551.e7.

11. Allegri N, Rossi F, Del Signore F, et al. Drug prescription appropriateness in the elderly: an Italian study. *Clin Interv Aging*. 2017;12:325–333.

12. Rajkumar AP, Ballard C, Fossey J, et al. Epidemiology of pain in people with dementia living in care homes: longitudinal course, prevalence, and treatment implications. *J Am Med Dir Assoc*. 2017;18(5): 453.e1–453.e6.

13. Defrin R, Amanzio M, de Tommaso M, et al. Experimental pain processing in individuals with cognitive impairment: current state of the science. *Pain*. 2015;156(8):1396–1408.

14. Husebo BS, Strand LI, Moe-Nilssen R, Borgehusebo S, Aarsland D, Ljunggren AE. Who suffers most? Dementia and pain in nursing home patients: a cross-sectional study. *J Am Med Dir Assoc*. 2008; 9(6):427–433.

15. van Kooten J, Smalbrugge M, van der Wouden JC, Stek ML, Hertogh CMPM. Prevalence of pain in nursing home residents: the role of dementia stage and dementia subtypes. *J Am Med Dir Assoc*. 2017;18(6):522–527.

16. Kunz M, Mylius V, Scharmann S, Schepelman K, Lautenbacher S. Influence of dementia on multiple components of pain. *Eur J Pain*. 2009;13(3):317–325.

17. Hadjistavropoulos T, LaChapelle DL, MacLeod FK, Snider B, Craig KD. Measuring movement-exacerbated pain in cognitively impaired frail elders. *Clin J Pain*. 2000;16(1):54–63.

18. Cole LJ, Farrell MJ, Duff EP, Barber JB, Egan GF, Gibson SJ. Pain sensitivity and fMRI pain-related brain activity in Alzheimer's disease. *Brain*. 2006;129:2957–2965.

19. Cole LJ, Gavrilescu M, Johnston LA, Gibson SJ, Farrell MJ, Egan GF. The impact of Alzheimer's disease on the functional connectivity between brain regions underlying pain perception. *Eur J Pain*. 2011; 15(6):568.e1–11.

20. Kunz M, Scharmann S, Hemmeter U, Schepelman K, Lautenbacher S. The facial expression of pain in patients with dementia. *Pain*. 2007;133: 221–228.

21. Sheu E, Versloot J, Nader R, Kerr D, Craig KD. Pain in the elderly: validity of facial expression components of observational measures. *Clin J Pain*. 2011;27(7):593–601.

22. Craig KD, Prkachin KM, Grunau RVE. The facial expression of pain. In: Turk DC, Melzack R, editors. *Handbook of Pain Assessment*. 3rd ed. New York, NY, USA: The Guilford Press; 2011:117–133.

23. 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.

24. 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.

25. Prkachin KM, Solomon PE. The structure, reliability and validity of pain expression: evidence from patients with shoulder pain. *Pain*. 2008;139(2):267–274.

26. Prkachin KM, Berzins S, Mercer SR. Encoding and decoding of pain expressions: a judgement study. *Pain*. 1994;58(2):253–259.

27. Hadjistavropoulos T, Herr K, Prkachin KM, et al. Pain assessment in elderly adults with dementia. *Lancet Neurol*. 2014;13(12):1216–1227.

28. Prkachin KM. The consistency of facial expressions of pain: a comparison across modalities. *Pain*. 1992;51(3):297–306.

29. Ekman P, Friesen WV. *The Facial Action Coding System*. Palo Alto, CA: Consulting Psychologists Press; 1978.

30. Hughes J, Hoti K, Atee MAW, Inventors; Electronic Pain Assessment Technologies (ePAT) Pty Ltd, assignee. A pain assessment method and system. US patent 2017/156661 (A1), 2017.

31. Atee M, Hoti K, Parsons R, Hughes JD. Pain assessment in dementia: evaluation of a point-of-care technological solution. *J Alzheimers Dis*. 2017;60(1):137–150.

32. AGS Panel on Persistent Pain in Older Persons. The management of persistent pain in older persons. *J Am Geriatr Soc*. 2002;50:S205–S224.

33. Kunz M, Mylius V, Schepelmann K, Lautenbacher S. On the relationship between self-report and facial expression of pain. *J Pain*. 2004;5(7):368–376.

34. Porter FL, Malhotra KM, Wolf CM, Morris JC, Miller JP, Smith MC. Dementia and response to pain in the elderly. *Pain*. 1996;68(2–3):413–421.

35. Lints-Martindale AC, Hadjistavropoulos T, Barber B, Gibson SJ. A psychophysical investigation of the facial action coding system as an index of pain variability among older adults with and without Alzheimer's disease. *Pain Med*. 2007;8(8):678–689.

36. Yan W-J, Wu Q, Liang J, Chen Y-H, Fu X. How fast are the leaked facial expressions: the duration of micro-expressions. *J Nonverbal Behav*. 2013;37(4):217–230.

37. Herr K, Coyne PJ, McCaffery M, Manworren R, Merkel S. Pain assessment in the patient unable to self-report: position statement with clinical practice recommendations. *Pain Manag Nurs*. 2011;12(4):230–250.

38. Pasero C, McCaffery M. *Pain Assessment and Pharmacologic Management*. St. Louism MO: Mosby: Elsevier Health Sciences; 2011.

39. Pasero C, McCaffery M. Pain control–no self-report means no pain-intensity rating. *Am J Nurs*. 2005;105(10):50–53.

40. Atee M, Hoti K, Parsons R, Hughes J. Evaluation of an automated facial recognition software application for assessment of pain among chronic pain sufferers. Australian Pain Society 35th Annual Scientific Meeting 17th March 2015, 2015; Brisbane.

41. Atee M, Hoti K, Hughes JD. Psychometric evaluation of the electronic Pain Assessment Tool (ePAT): an innovative instrument for individuals with moderate to severe dementia. *Dement Geriatr Cogn Disord*. 2017;44(5–6):256–267.

42. International Association for the Study of Pain (IASP). *IASP Curricula*; 2016. Available from: http://www.iasp-pain.org/Education/CurriculaList.aspx?navItemNumber=647. Accessed June 2, 2016.

43. Viera AJ, Garrett JM. Understanding interobserver agreement: the kappa statistic. *Fam Medicine*. 2005;37(5):360–363.

44. Lin LI. A concordance correlation-coefficient to evaluate reproducibility. *Biometrics*. 1989;45(1):255–268.

45. Altman D. Inter-rater agreement. *Pract Stat Med Res*. 1991;5:403–409.

46. McBride G. A proposal for strength-of-agreement criteria for Lin's concordance correlation coefficient. *NIWA Client Report: HAM2005–062*. 2005. Available from: https://www.medcalc.org/download/pdf/McBride2005.pdf. Accessed January 10, 2018.

47. Schaible HG, Schmidt RF. Effects of an experimental arthritis on the sensory properties of fine articular afferent units. *J Neurophysiol*. 1985;54(5):1109–1122.

48. Herr K, Bursch H, Ersek M, Miller LL, Swafford K. Use of pain-behavioral assessment tools in the nursing home expert consensus recommendations for practice. *J Gerontol Nurs*. 2010;36(3):18–29.

49. Herr K. Pain assessment in cognitively impaired older adults. *Am J Nurs*. 2002;102(12):65–67.

50. Lichtner V, Dowding D, Esterhuizen P, et al. Pain assessment for people with dementia: a systematic review of systematic reviews of pain assessment tools. *BMC Geriatr*. 2014;14:138.

51. Lewinski P. Automated facial coding software outperforms people in recognizing neutral faces as neutral from standardized datasets. *Front Psychol*. 2015;6:1386.

52. Bartlett MS, Littlewort GC, Frank MG, Lee K. Automatic decoding of facial movements reveals deceptive pain expressions. *Curr Biol*. 2014;24:738–743.

53. McDougall JJ. Arthritis and pain: neurogenic origin of joint pain. *Arthritis Res Ther*. 2006;8(6):220.

54. Beach PA, Huck JT, Miranda MM, Foley KT, Bozoki AC. Effects of Alzheimer disease on the facial expression of pain. *Clin J Pain*. 2016;32(6):478–487.

55. Klein DG, Dumpe M, Katz E, Bena J. Pain assessment in the intensive care unit: development and psychometric testing of the nonverbal pain assessment tool. *Heart Lung*. 2010;39(6):521–528.

56. Herr K, Zwakhalen S, Swafford K. Observation of pain in dementia. *Curr Alzheimer Res*. 2017;14(5):486–500.

57. Horgas AL, Elliott AF, Marsiske M. Pain assessment in persons with dementia: relationship between self-report and behavioral observation. *J Am Geriatr Soc*. 2009;57(1):126–132.

58. Lukas A, Barber JB, Johnson P, Gibson SJ. 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.

59. Kim H, Neubert JK, Rowan JS, Brahim JS, Iadarola MJ, Dionne RA. Comparison of experimental and acute clinical pain responses in humans as pain phenotypes. *J Pain*. 2004;5(7):377–384.

60. Closs SJ, Dowding D, Allcock N, et al. Health Services and Delivery Research. *Towards Improved Decision Support in the Assessment and Management of Pain for People with Dementia in Hospital: A Systematic Meta-Review and Observational Study*. Southampton, UK: NIHR Journals Library; 2016.

61. Atee M, Hoti K, Hughes JD. A technical note on the PainChek™ system: a web portal and mobile medical device for assessing pain in people with dementia. *Front Aging Neurosci*. 2018;10:117.