Clinimetric properties of the electronic Pain Assessment Tool (ePAT) for aged-care residents with moderate to severe dementia

Kreshnik Hoti 1,2
Mustafa Atee 1
Jeffery D Hughes 1

1School of Pharmacy, Curtin University, Perth, Australia; 2Division of Pharmacy, Faculty of Medicine, University of Prishtina, Pristina, Kosovo

Purpose: Accurate pain assessment is critical to detect pain and facilitate effective pain management in dementia patients. The electronic Pain Assessment Tool (ePAT) is a point-of-care solution that uses automated facial analysis in conjunction with other clinical indicators to evaluate the presence and intensity of pain in patients with dementia. This study aimed to examine clinimetric properties (clinical utility and predictive validity) of the ePAT in this population group.

Methods: Data were extracted from a prospective validation (observational) study of the ePAT in dementia patients who were ≥65 years of age, living in a facility for ≥3 months, and had Psychogeriatric Assessment Scales – cognitive scores ≥10. The study was conducted in two residential aged-care facilities in Perth, Western Australia, where residents were sampled using purposive convenience strategy. Predictive validity was measured using accuracy statistics (sensitivity, specificity, positive predictive value, and negative predictive value). Positive and negative clinical utility index (CUI) scores were calculated using Mitchell’s formula. Calculations were based on comparison with the Abbey Pain Scale, which was used as a criterion reference.

Results: A total of 400 paired pain assessments for 34 residents (mean age 85.5±6.3 years, range 68.0–93.2 years) with moderate–severe dementia (Psychogeriatric Assessment Scales – cognitive score 11–21) were included in the analysis. Of those, 303 episodes were classified as pain by the ePAT based on a cutoff score of 7. Unadjusted prevalence findings were sensitivity 96.1% (95% CI 93.9%–98.3%), specificity 91.4% (95% CI 85.7%–97.1%), accuracy 95.0% (95% CI 92.9%–97.1%), positive predictive value 97.4% (95% CI 95.6%–99.2%), negative predictive value 87.6% (95% CI 81.1%–94.2%), CUI+ 0.936 (95% CI 0.911–0.960), CUI– 0.801 (95% CI 0.748–0.854).

Conclusion: The clinimetric properties demonstrated were excellent, thus supporting the clinical usefulness of the ePAT when identifying pain in patients with moderate–severe dementia.

Keywords: ePAT, PainChek™, pain assessment, dementia, predictive validity, clinical utility, automated facial analysis

Introduction

Accurate and objective pain assessment is a foundation stone of optimal pain management. This is particularly relevant for patients with impaired cognition, such as those with dementia, whose ability to report or verbalize pain is compromised.1,2 In this population group, there is currently a large unmet need in relation to pain assessment, resulting in undetected and untreated pain, which then contributes to worsening cognition (eg, delirium) and behavior (eg, agitation and aggression), inappropriate prescribing (eg, antipsychotics), and adverse drug effects (eg, death).1,2 Furthermore, suboptimal pain assessment and management compromise the quality of life of patients.
with dementia and constitute a significant burden for their families and carers. In order to improve the current situation, the electronic Pain Assessment Tool (eP AT; also known as PainChek™) was developed, which is a point-of-care smart device-enabled application that uses automated facial (video) analysis in conjunction with clinical indicators to evaluate the presence and intensity of pain in those with communication difficulties, including dementia. This application was designed to improve the objectivity and accuracy of assessing pain in patients with dementia, ultimately leading to effective pain management for this disadvantaged group.

In July 2017, the eP AT application was approved as a Class I medical device by the Therapeutic Goods Administration in Australia and received the CE mark in Europe “...to assess and monitor pain in people who cannot verbalise such as people with dementia or communication difficulties.”

To date, the application has been validated in a total of 74 residents aged 60–98 years across five residential aged-care facilities in Western Australia. In this setting, it has shown strong psychometric properties, including concurrent validity, discriminant validity, and interrater reliability.

Clinimetric properties, such as predictive validity and clinical utility (CU), are also important criteria in evaluating new pain-assessment tools, because they determine the usefulness and applicability of the tool in the clinical setting. Literature data indicate that a large number of psychometric evaluation studies of pain-assessment tools have neglected this aspect. Zwakhalen et al point out that further research should test validity, reliability, and CU of existing pain-assessment tools with the view of improving them. Furthermore, a more recent systematic review by Lichtner et al concluded that no single pain-assessment tool can be currently recommended in patients with dementia on the basis of lack of evidence concerning validity and CU to assess pain in patients with cognitive impairment.

The CU of a pain-assessment tool is an essential clinimetric parameter that goes beyond the analytical, technical, or even diagnostic accuracy performance of the tool. Testing CU in fact provides more insight into potential health benefits and outcomes, especially in comparison to existing options, in this case the Abbey Pain Scale (APS), which is currently a widely used (ie, silver) standard for assessing pain in patients with dementia in Australia and other countries. Furthermore, CU provides important information in relation to how useful the tool is in assisting the decision-making process about the patients, eg, administering an analgesic drug when a patient is in pain.

In addition to our aim of examining CU, the focus of this paper is also to assess the predictive validity of the eP AT. This is relevant, given the significance of data in evidencing the potential ability of the eP AT in predicting criterion-reference measures of the APS, which are a key reason for assessing this important clinimetric property. Specifically, this study aimed to analyze the predictive validity of the eP AT with a view toward discriminating and comparing between pain on movement and at rest with reference to the APS.

Methods

Ethics

This research was approved by the Human Research Ethics Committee of Curtin University, Western Australia (HREC: HR10/2014) and Mercy Health, Victoria (R15/50AC) in Australia. The study was also registered with the Therapeutic Goods Administration under the Clinical Trial Notification scheme (CT-2016-CTN-04886-1 v1). Proxy (third party) written informed consent was provided by legal representative(s) of participants involved, due to the latter’s impaired cognitive capacity to provide their own consent.

No additional data were collected for the current study, as all data were sourced from a previously published study. Therefore, there was no burden on staff or residents. Data were deidentified to ensure confidentiality.

Study design, setting, and inclusion criteria

Data were extracted from a previous validation (observational) study of the eP AT, which was conducted in two residential aged-care facilities in Perth, Western Australia over 10 weeks between January and April 2017. Pain assessments were done twice (at rest and postmovement) in pairs by eP AT rater and APS raters (nursing staff of the residential aged-care facilities). Raters performed assessments independently, and were blinded in terms of one another’s assessment and pain therapies received by the residents. All pain assessments were performed inside the facilities between 1-4pm while the residents were receiving their standard clinical care. This was either done during sitting, or recumbent positions to reflect rest (non-nociceptive) conditions, or during walking or transfer as examples of movement (nociceptive) conditions. Residents’ cognitive level was assessed prior to the study using Psychogeriatric Assessment Scales – cognitive (PAS-Cog), which has been validated in people with dementia. PAS-Cog scores categorize the degree of cognitive impairment according to increments of 0–3 = minimal, 4–9 = mild, 10–15 = moderate,
and 16–21 = severe. Cognitive data were extracted from the electronic medical records of the residents.

Residents were selected if they met the inclusion criteria: ≥65 years of age, living in facility for ≥3 months, PAS-Cog score ≥10, had a diagnosis of dementia by a geriatrician, and a medical history or presenting complaint(s) linked to painful condition(s), such as arthritis. Residents were excluded if they were medically unwell or unable to display intact facial expressions. Further details of the methodology employed in this study, including inclusion and exclusion criteria of residents and raters, have also been discussed in a previous publication.

Statistical methods
Descriptive statistics (eg, mean, median) were reported for pain scores, while inferential statistics (eg, CI) were used to determine the level of significance. The receiver-operating characteristic (ROC) curve was used to determine cutoff scores for the presence of pain, which were classified in a binary format (yes/no). The ROC curve is a graphical representation of the sensitivity versus 1 – specificity for a dichotomous variable. The curve is a method of estimating the best balance between sensitivity and specificity using multiple measurements, and enables the selection of an optimal threshold value (cutoff point). Besides cutoff points, ROC curve analysis can be used to convert continuous or ordinal variables into dichotomous outcomes. To derive the maximum value for these cutoff points, Youden’s index (sensitivity + specificity – 1) was also used. In our study, eP AT was the tool under investigation against the standard instrument, ie, the APS. Cross-tab calculations were made to enable univariate percentage agreement between eP AT and APS of the presence or absence of pain.

Predictive validity was measured using accuracy statistics (sensitivity, specificity, positive predictive value, negative predictive value, and positive and negative likelihood ratios, respectively). These values were calculated twice: before and after prevalence adjustment. Positive (+) and negative (-) CU index (CUI) values were calculated using Mitchell’s formula. In this study, we used the STARD (Standards for Reporting of Diagnostic Accuracy Studies) guidelines to report the clinimetric findings. Calculations were based on comparison with the APS, which was used as a criterion reference in our study. Presence of pain was defined as per published scores of the APS (≥3) or eP AT (≥7), with absence of pain defined as APS score of 0–2 or eP AT score of 0–6.

All analyses were done using Microsoft Excel 2013 for Windows 7 Enterprise (Microsoft Corporation, Redmond, WA, USA), Clinical Utility Index Calculator, and MedCalc statistical software (version 17.4 for Windows; MedCalc Software, Ostend, Belgium).

Results
Participants
A total of 34 residents aged 68.0–93.2 years were included in our sample, with a mean of 85.5±6.3 years. The sample comprised 20 females (58.8%) and 14 males (41.2%), of whom 35.3% had Alzheimer’s dementia and 44.1% an unspecified type of dementia. With the exception of one resident, all (n=33, 97.1%) were Caucasian. A total of 27 residents (79.4%) had a diagnosis of severe dementia based on PAS-Cog scores (mean 19.7±2.5, range 11–21). The vast majority of the sample were either nonambulant (n=14, 41.2%) or ambulant with assistance (n=19, 55.9%). A full description of the sample data was published in a previous study.

Pain data
A total of 400 paired pain assessments for 34 residents were included in the analysis. Of those, 303 episodes were classified as pain by the ePAT based on a cutoff score of 7 (Table 1). Values ≤6 were defined as no pain. On the other hand, 307 were deemed pain by APS based on a cutoff score of 3, whereby scores ≤2 were considered no pain (Table 2). Mean ePAT pain scores were significantly greater with movement (11.44±3.54, P<0.0001) than rest (8.33±3.34). During assessments, types of physical activities undertaken by residents were varied, and ranged from sitting to walking.

Clinimetric data
Predictive validity
Predictive validity was calculated using sensitivity, specificity, and accuracy before and after prevalence adjustment. In the context of this validation study (n=400 paired
assessments), sensitivity represents the true positive rate of detecting pain by the ePAT when pain actually exists, as indicated by the APS. Specificity denotes the true-negative rate of detecting pain (ie, absence of pain) by the ePAT when no pain is identified according to the APS. For the ePAT, the presence of pain was defined as a score of ≥7 on the final scale. According to published scores of the APS, a total score ≥3 refers to presence of pain, while a score ≤2 refers to the absence of pain.11

Accuracy statistics before prevalence adjustment

Diagnostic sensitivity and specificity were both calculated based on the results shown in Tables 1 and 2. Findings from accuracy statistics are presented in Tables 3 and 4. The ePAT has a pain-negative result if pain is absent, ie, in the case of no pain. Compared to the APS results in Table 2, 85 assessments were identified as true negative (TN) and 12 as false negative (FN). The ePAT has a positive pain result if pain exists, ie, the presence of pain. When compared to the APS, ePAT assessments were found to be 295 true positives (TPs) and eight false positives (FPs).

Determining accuracy, sensitivity, and specificity of the ePAT

Two steps were used to calculate these parameters. First, a 2×2 contingency table was set (Table 3). Next, sensitivity, specificity, accuracy, likelihood ratios, and predictive values were calculated (Table 4):

### Table 3 Contingency (2×2) table for accuracy statistics

<table>
<thead>
<tr>
<th>Pain score results</th>
<th>APS score</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Present (APS ≥3)</td>
<td>Absent (APS ≤2)</td>
</tr>
<tr>
<td>Pain</td>
<td>307</td>
<td>93</td>
</tr>
<tr>
<td>ePAT score</td>
<td>295 (TP)</td>
<td>8 (FP)</td>
</tr>
<tr>
<td>Positive (present, ePAT ≥7)</td>
<td>295 (TP)</td>
<td>8 (FP)</td>
</tr>
<tr>
<td>Negative (absent, ePAT 0-6)</td>
<td>12 (FN)</td>
<td>85 (TN)</td>
</tr>
</tbody>
</table>

Abbreviations: ePAT, electronic Pain Assessment Tool; APS, Abbey Pain Scale; TP, true positive; FP, false positive; FN, false negative; TN, true negative.

### Table 4 Calculations of sensitivity, specificity, accuracy, likelihood ratios, and predictive values before prevalence adjustment

<table>
<thead>
<tr>
<th>Clinimetric parameter</th>
<th>Formula</th>
<th>Value</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sensitivity</td>
<td>(\text{Sensitivity} = \frac{(TP + FN)}{(TP + FN)}) × 100</td>
<td>96.1%</td>
<td>93.9%–98.3%</td>
</tr>
<tr>
<td>Specificity</td>
<td>(\text{Specificity} = \frac{TN}{(TN + FP)}) × 100</td>
<td>91.4%</td>
<td>85.7%–97.1%</td>
</tr>
<tr>
<td>Positive likelihood ratio</td>
<td>(\text{Positive likelihood ratio} = \text{Sensitivity} / 100 - \text{Specificity})</td>
<td>11.2</td>
<td>5.8–21.7</td>
</tr>
<tr>
<td>Negative likelihood ratio</td>
<td>(\text{Negative likelihood ratio} = 100 - \text{Specificity/sensitivity})</td>
<td>0.04</td>
<td>0.02–0.07</td>
</tr>
<tr>
<td>Positive predictive value</td>
<td>(\text{Positive predictive value} = \frac{(TP + FP)}{(TP + FN + FP)}) × 100</td>
<td>97.4%</td>
<td>95.6%–99.2%</td>
</tr>
<tr>
<td>Negative predictive value</td>
<td>(\text{Negative predictive value} = \frac{TN}{(TN + FP)}) × 100</td>
<td>87.6%</td>
<td>81.1%–94.2%</td>
</tr>
<tr>
<td>Pain prevalence</td>
<td>(\text{Pain prevalence} = \frac{(TP + FN)}{(TP + TN + FP + FN)}) × 100</td>
<td>76.8%</td>
<td>72.3%–80.8%</td>
</tr>
<tr>
<td>Accuracy</td>
<td>(\text{Accuracy} = \frac{(TP + TN)}{(TP + TN + FP + FN)}) × 100</td>
<td>95.0%</td>
<td>92.9%–97.1%</td>
</tr>
</tbody>
</table>

Note: All values approximated to closest decimal point.

Abbreviations: TP, true positive; FP, false positive; FN, false negative; TN, true negative.
slightly varied. A choice of 50% prevalence was made to reflect the current data available in the literature.\(^\text{22}\) The optimal cutoff score where there was a balance of sensitivity and specificity was 7. This was confirmed by the area-under-the-curve value of 0.98 and \(P<0.0001\). A Youden’s index value of close to 1 (ie, 0.88) pointed to minimal FPs and FNs in the data set. The closeness of the curve to the left corner of the graph illustrates ePAT’s high sensitivity and specificity in distinguishing between pain/no pain cutoffs. Results from the ROC curve used to analyze cutoff scores are presented in more detail in Table 5 and Figure 1.

### Discussion

This study evaluated the clinimetric properties (predictive validity and CU) of the ePAT in residents with moderate–severe dementia. For the ePAT application, this provides more evidence to how clinically meaningful the test results are in relation to pain detection in the target-population group. Other psychometric properties, such as concurrent validity, discriminant validity, and interrater reliability, have been reported elsewhere.\(^\text{4,6}\)

| Table 5 ROC curve characteristics, optimal criterion, and calculated sensitivity and specificity after prevalence adjustment |
|-----------------|-----------------|
| Variable under investigation | ePAT |
| Classification variable (pain) | APS |
| Number of pain assessments (paired) | 400 |
| Positive group | 307 (76.8%) |
| Negative group | 93 (23.2%) |
| Pain prevalence (%) | 50 |
| Area under the ROC curve (AUC) | 0.98 |
| Standard Error | 0.00572 |
| 95% Confidence interval | 0.96–0.99 |
| z statistic | 83.9 |
| Significance level \(P\) (Area > 0.5) | <0.0001 |
| Youden index | 0.88 |
| 95% Confidence interval\(^\text{a}\) | 0.82–0.92 |
| Associated criterion | >7 |
| 95% Confidence interval\(^\text{b}\) | >6 to >8 |
| Sensitivity | 91.2 |
| Specificity | 96.8 |

**Optimal criterion**

\(\text{Youden index}\) = 0.88

### Notes:
- Taking into account disease prevalence (50%) and estimated costs: cost False Positive: 1; cost False Negative: 1 cost True Positive: 0; cost True Negative: 0. Positive group: pain present; Negative group: pain absent; BC, bootstrap confidence interval (1000 iterations; random number seed: 978).
- Abbreviations: ePAT, electronic Pain Assessment Tool; ROC, receiver-operating characteristic; BC, bias corrected and accelerated.

As reported by Herr et al\(^\text{9}\) and later by Lichtner et al,\(^\text{2}\) currently the evidence for existing tools in regard to their validity and CU is limited and unclear. Due to these psychometric and clinimetric limitations, clinician and carer guidance on effective pain assessment is currently at best compromised, leading to gaps in relation to informed treatment options and care plans.\(^\text{2}\) Some studies that have previously reported predictive validity of pain-assessment tools have been marred by failure actually to report the data, whereas others had significant scoring differences (pre- and postintervention).\(^\text{2}\) The situation is similar in regard to the CU of pain-assessment tools. In addition to clinical data being completely or substantially absent for some tools, conflicting data have been reported for other tools.\(^\text{2}\) Furthermore, CU dimensions, such as the cutoff scores (eg, pain vs no-pain cutoff described in the present study) needed for clinical decision making, have not been reported by many studies evaluating the currently available pain-assessment tools.\(^\text{2}\) For example, the total score of the Pain Assessment in Advanced Dementia Scale can be 0–10 points. However, there is no evidence of pain-intensity scoring.\(^\text{21}\) The original Pain Assessment Checklist for Seniors with Limited Ability to Communicate (PACSLAC) study did not report a specific cutoff score to determine the presence of pain. Rather, an increased PACSLAC score suggests that an increase in pain is likely, while a lower score suggests that pain has reduced.\(^\text{24}\) Later, it was found that PACSLAC scores greater than 12 (of 60) are indicative of high pain intensity, whereas scores of 0–5 represent usual pain.\(^\text{24}\) In the current study, the fact that the area under the curve was 0.98 (\(P<0.0001\)) indicated the ePAT’s strong ability to distinguish between cutoff scores for pain. Our current analysis confirmed a cutoff score of 7 for the presence of pain, also

![Figure 1 Electronic Pain Assessment Tool (ePAT) receiver-operating characteristic curve. ](image-url)
previously reported by Atee et al. This lends further support to our previous findings.

This is the first study to demonstrate that the ePAT is a useful and valid instrument to assess pain in patients with moderate–severe dementia. This is an important step in further evidencing this instrument in the context of better assessment and management of pain in patients with dementia, given their compromised ability to communicate pain, which results in underrecognition and undertreatment of their pain.

High sensitivity and specificity values for the ePAT in identifying pain yield strong support to its predictive validity and responsiveness to change. Even after disease-adjusted prevalence analysis had been performed, sensitivity and specificity values changed only slightly. This validates our approach to the conceptual foundation of the tool, previously reported by Atee et al. Pain assessments were performed during rest and movement, and the resultant scores reflected the change in timing, regardless of the order of testing. The overall CUI value of the ePAT reported in this study was excellent (0.95) as indicated by Mitchell’s analysis. This suggests that the ePAT is a useful tool to assist in the clinical decision-making processes related to pain management, including informing clinicians on specific actions, such as analgesic administration, emphasized as a necessary property by van Herk et al. when observing pain-assessment tools in patients with cognitive impairment. Currently, there are no published studies reporting Mitchell’s CUI of pain-assessment tools in dementia. Other clinical tools where Mitchell’s CUI has been used include the Patient Health Questionnaire (PHQ9 and PHQ2) for depression in primary care, and the Cornell Scale for Depression in Dementia. Our findings presented much better accuracy data than those tools.

Strengths and limitations
The present study has several points of strength. Statistical analyses were methodical and covered a number of variables. For example, Mitchell’s index was used for the first time (as far as we know) to investigate the CUI of a pain-assessment tool for people with dementia. The data were gleaned from participants with varying types of dementia and pain conditions, and this also addresses recommendations by systematic reviews. Full medical histories of residents were accessible to raters, which assisted in informing the clinical picture when doing pain assessments.

Among the limitations of our study (which was noted previously), it is important to highlight that the comparison was made with reference to the APS. Despite its frequent use in Australia, the APS is not a global gold standard. The original sample was relatively small, and contained disproportionate sex representation. Further, data obtained were limited to residential aged-care settings. Thus, the findings need to be interpreted with caution before applying them to other settings, eg, hospital. Also, there was no physical examination conducted for the sake of the study to identify the potential source of pain. In fact, pain data were based on medical history, where some aspects might not be a true representation of the current clinical status of residents. However, in the case of the ePAT, medical history was weighed as only 1 of the 42 points of the total scale. Therefore, it had a negligible effect on our data.

CU encompasses a wide range of effects that tests or tools can have on the patient. In this regard, it is worth emphasizing that a recent study conducted by van Kooten et al suggested that improving communication of results to physicians can lead to improvement in pain management in long-term care patients with dementia. Although the present study did not evaluate this aspect, it is worth emphasizing that the ePAT tool is the only pain-assessment tool available that uses automation and is able to document and manage clinical information electronically. This is thus expected to have positive effects on pain detection and management and communication to physicians. This aspect needs to be further researched in future as an additional potential indicator of the ePAT’s clinical usefulness. CU also has a significant value for implementing tools in clinical practice. More emphasis and research on the CU of other pain-assessment tools are needed, because these parameters provide data on practicalities, which may improve the uptake and utilization of tools in clinical settings. Use of an electronic pain-assessment tool, which eases the process of documentation and pain tracking over time, has the potential to positively impact the tool’s CU through facilitating information on the temporality (ie, trends and patterns) of pain. In this regard, it is worth highlighting that the temporality of pain was regarded as “most useful to the assessment of pain” by Lichtner et al. This potential strength of the ePAT application needs to be researched further in future.

Conclusion
The findings reveal strong clinimetric properties of the ePAT. We have demonstrated that both its predictive validity and CU are excellent, giving further evidence to the quality of pain assessment attained by the ePAT.

Acknowledgments
The authors express their gratitude to everyone involved in the study, including aged-care staff, residents, and their
families. The authors would like to acknowledge the contribution of an Australian Government Research Training Program Scholarship in supporting this research. The original research that led to the development of the PainChek tool is part of a PhD project that was also supported by the Dementia Australia Research Foundation (DARF) through grant funding and a stipend scholarship. The content of the article is solely the responsibility of the authors, and does not necessarily represent the official views of DARF. The project has been commercialized initially through a startup company (ePAT Pty Ltd), and since October 2016 the Australian Share Securities (ASX)-listed EPAT Technologies Ltd (now known as PainChek Ltd). This research was also sponsored by EPAT Technologies Ltd.

Author contributions
KH, MA, and JDH conceived the idea and designed the study. KH and MA performed the literature review and drafted the manuscript. KH, MA, and JDH revised the manuscript. MA conducted and interpreted the statistical analyses. All authors contributed to and approved the final version of the manuscript.

Disclosure
KH, MA, and JDH are shareholders in PainChek Ltd (formerly known as EPAT Technologies), which is commercializing the ePAT instrument titled PainChek™. They also have a patent application titled “A pain assessment method and system; PCT/AU2015/000501”, which has been under the national phase examination since February 2, 2017. KH is employed as a consultant by PainChek Ltd while serving as an assistant professor at the University of Prishtina and an adjunct senior lecturer at the School of Pharmacy, Curtin University. MA is a research scientist for PainChek Ltd while serving as a research fellow and PhD candidate with the School of Pharmacy, Curtin University. JDH is employed as chief scientific officer of PainChek Ltd while serving as a professor in the School of Pharmacy, Curtin University.

References

