

## **Paramedic adult pain assessment: Pilot Study**

### **Abstract**

#### **Objective**

Our study aimed to determine the inter-rater reliability of the Wong-Baker FACES® Pain Rating Scale in the prehospital setting in the State of Qatar with five adult standardised patients.

**Methods:** This prospective, quantitative pilot study gathered primary data using survey questionnaires. Five members of staff were prepared as standardised adult patients presenting with differing reference levels of pain. 35 consenting paramedics assessed and indicated the pain intensity score of the five standardised adult patients using the Wong-Baker FACES® Pain Rating Scale. Each participant was exposed to the same five standardised adult patients and the same range of facial expressions in a randomized order.

**Results:** The paramedics recorded the pain score of the 5 standardised patients based on their observations of the patient's facial expression. Overall, the Inter-Rater Reliability as determined through Fleiss Kappa indicates only a poor to slight agreement of the allocated pain scores, as described against the reference standards. There was a wide grouping of the pain score levels around the reference standard. Most of the allocations were 1 to 2 pain score levels away from the reference standard, although not in a normal distribution with some of the higher reference pain levels receiving lower scores, and vice versa. Overall the sensitivity was poor to very poor throughout.

**Conclusion:** This study has shown that the Inter-Rater Reliability of the participant sample when applying the Wong-Baker FACES® Pain Rating Scale to the study population in order to determine the pain levels of the five standardised adult patient cases was extremely poor because the tool was not appropriately utilized by the clinician. This could be attributed to various factors including the multi-national population, language barriers, lack of familiarization with the Wong-Baker FACES® Pain Rating Scale and other environmental factors.

### **Key point**

Paramedic assessment of patients' pain is essential to their appropriate management in the prehospital setting. Further, multicultural environments present communication challenges and create biases when it comes to pain assessment.

### **Reflective questions**

1. Is assessing a patient's pain intensity in the prehospital setting essential?
2. What are the most appropriate methods or tools utilized to assess the pain intensity of an adult patient in the prehospital setting?
3. Is the Wong-Baker FACES® Pain Rating Scale appropriate for the assessment of an adult patient's pain intensity?
4. Is it essential to manage a patient's pain in the prehospital setting?
5. Is the Wong-Baker FACES® Pain Rating Scale as intuitive to use as one might think it is?

## **Introduction**

Qatar has a population that exceeds 2.6 million residents (MDPS 2016). However, nearly 90% of the population comprises emigrant workers mainly from Asia and North Africa (Goodman 2015). Hamad Medical Corporation Ambulance Service (HMCAS) employed Critical Care Paramedics (CCPs) and Ambulance Paramedics (APs) (Wilson et al. 2017) receive diversified training in pain assessment in their home countries and are also trained in various pharmacological and non-pharmacological remedies for pain. CCPs are primarily recruited from western countries. In the prehospital environment in Qatar, CCPs work with APs from various linguistic and cultural backgrounds. These may include APs from Tunisia, India, Philippines, Jordan, Morocco, Egypt, and Britain. Qatar's multinational population adds to the diversity of emergency medical care practice (Gangaram et al. 2017).

Globally, on average 4 out of 5 (80%) of all patients seeking emergency medical service (EMS) help, experience pain (Iqbal et al. 2015). Recent studies show that pain is poorly assessed in the prehospital setting (Lynde and Zorab 2015). Delays in prehospital pain assessment and treatment are further prolonged in the Emergency Centre due to the initial triage processes (Hodkinson 2016).

Researchers conducted two significant ( $p < 0.001$ ) studies in California (USA) to determine the effects of an Educational Intervention (EI) on prehospital pain management (French et al. 2006, French et al. 2013). Both studies provided paramedics with a three-hour long EI, with surveys completed before the EI and one month after. In 2001 (French et al. 2006) the researchers reviewed 297 surveys and 439 EMS patient care reports (PCRs) with pain complaints. They found that following

the EI, paramedics' knowledge of basic pain management principles increased by 17.5% (57.3% to 74.9%). Their use of non-pharmacological pain therapies improved by 32.2%, documentation of pain severity by 51.0%, and pain characteristics by 24.0%. Overall the reassessment of pain following the EI improved by 13.0%. Even before the EI conducted in 2007 (French et al. 2013) the researchers found an improvement since 2001 in the basic knowledge of pain management by 18.2%, perceptions of pain by 9.2%, and management of pain by 13.8%. The researchers concluded that ongoing continuing education of pain management is key to improved effectiveness of prehospital pain management.

The appropriate assessment and treatment of pain, in the prehospital setting in Qatar, has been identified as a key performance indicator (KPI) by HMCAS. Currently, patients' pain is assessed by HMCAS paramedics using the Wong-Baker FACES® Pain Rating Scale (Figure 1). This pain rating scale translates facial pain expression into a numerical pain scale (NPS) rating which is then recorded on the electronic patient case report (ePCR) using a zero to ten numerical value. The tool was primarily designed for paediatric patients who are unable to verbalize their pain intensity score but can pictorially reflect their pain intensity. It is mandated by HMCAS policy that all patients presenting with a pain intensity score of more than 4/10 receive prescribed analgesia based on the clinical practice guidelines (CPG) and the paramedic's specific scope of practice. However, recent HMCAS findings indicated that the assessment of patients presenting with acute pain was sub-optimal.

Use of the Wong-Baker FACES® Pain Rating Scale at HMCAS has not previously been researched. The plethora of evidence suggests that once pain is assessed and

documented accurately, patients are more likely to receive appropriate analgesia. An Australian emergency centre study was conducted to assess pain score documentation and the treatment thereof (Furyk and Sumner 2008). The researchers conducted a retrospective evaluation of 145 charts from patients with confirmed appendicitis. Pain scores were documented for 13 children and 79 adults. Eleven children as compared to 79 adults received intravenous morphine. The study suggested that once pain is assessed and documented accurately, the likelihood of patients receiving analgesia is increased.

A further retrospective cross-sectional study was conducted on emergency medical service (EMS) PCR after the introduction of a prehospital pain assessment protocol (McLean et al. 2004). Data extracted included Verbal Rating Scale (VRS), NPS, and emergency call related information. In total 1,227 PCRs were studied. 907 (75%) were non-trauma EMS transports. Two percent (n=27) of the study population were unconscious. Pain was assessed using the EMS protocol in 1,002 of 1,200 (84%) patients. Of the 518 patients reporting pain, 104 (20%) completed a VRS but not an NPS. A total of 31% of patients reported moderate or severe pain. Prehospital pain assessment using a VRS and NPS was thus feasible.

In addition, studies show that ethnicity affects the appropriate assessment of pain (Todd et al. 2000, Tamayo-Sarver et al. 2003). Given the local context with cultural and language differences, varying expectations of Qatar's population regarding the treatment provided by EMS professionals and their scope of practice, and suspected differences in the assessment of pain by HMCAS paramedics, this pilot research study was deemed to be of interest. Further, no inter-rater reliability studies were found that

evaluated the use of the Wong-Baker FACES® Pain Rating Scale on adult patients. Searches of databases including Science Direct, Medline, EMBASE and CINAHL revealed that the use of the Wong-Baker FACES® Pain Rating Scale has not been assessed in the prehospital setting on adult patients. Although the Wong-Baker FACES® Pain Rating Scale has been designed for use by paediatric patients to self-report their pain intensity level, at HMC it has been adopted for use to include adult patients.

## **Methods**

### **Study design**

A prospective, quantitative pilot study was conducted. Primary data on the paramedics' assessment of pain was gathered utilizing survey questionnaires following simulation-based interactions with five standardised adult patients.

### **Study setting, population and sample**

This pilot study was conducted through HMCAS in Qatar. The study was approved by the organisations Medical Research Centre (16155/16). To direct and coordinate HMCAS emergency resources, the organisation utilizes the hub and spoke model. This model was designed to ensure that the public have rapid access to emergency care. The country has 6 hubs with 29 spokes (Wilson et al. 2017). A sample size of 3.0% (35/1159) of APs and CCPs was deemed the minimum appropriate requirement for this inter-rater reliability pilot study. Participant recruitment was randomized based on staff presence at the various locations during the study data collection period.

## **Study protocol**

Five members of staff from the HMCAS training department were prepared as standardised adult patients presenting with differing reference levels of pain. These standardised adult patients were transported to all HMCAS hubs and spokes over a period of two weeks. All paramedics present at these locations were invited to voluntarily participate in the study. No advance invitations were circulated to prospective participants to prevent participants refreshing on the use of the Wong-Baker FACES® Pain Rating Scale prior to data collection. On the day of data collection, information letters regarding the study were circulated to all prospective participants. Only consenting participants were recruited into the study. The data collection tool included demographic questions and a section in five parts regarding pain scoring for the different cases. The standardised adult patients' presentation sequence to participants was done using a randomization table. The paramedics were then required to assess the standardised adult patients' pain utilizing the Wong-Baker FACES® Pain Rating Scale and record the score on the data collection tool for each case. The participants were required to explain the procedure to the standardised adult patients, get their consent, explain the use of the Wong-Baker FACES® Pain Rating Scale, and get the patient to identify their pain intensity score. Anonymised completed questionnaires were placed in a sealed box. The data collection process did not impact on the paramedics' availability to respond to emergency calls.

The standardised adult patient scenarios were validated by a focus group comprising of instructors from the HMCAS training department, Consultant Paramedics from HMCAS, academics from the Durban University of Technology's Department of Emergency Medical Care and Rescue. The actors for the simulated adult patient

scenarios remained constant throughout the data collection process. The five cases included;

- Case A: 47-year-old, Indian national male patient presenting with acute coronary syndrome (ACS) and mimicking a facial expression of pain corresponding to an intensity score of 7/10.
- Case B: 25-year-old, Qatari national male patient presenting with a fractured left femur and a pain mimicking a facial expression of pain corresponding to an intensity score of 10/10.
- Case C: 65-year-old, Pakistani national female patient presenting with appendicitis and a pain mimicking a facial expression of pain corresponding to an intensity score of 4/10.
- Case D: 38-year-old, Egyptian national male patient presenting with frontal headaches and mimicking a facial expression of pain corresponding to an intensity score of 2/10.

Case E: 30-year-old, South African national male patient presenting with renal colic and mimicking a facial expression of pain corresponding to an intensity score of 6/10.

## **Data analysis**

Microsoft Excel (© Microsoft Office, Palo Alto, CA) was used as the primary analytical software. An add-in analysis tool from Real Statistics (© Charles Zaiontz, Trento) was used to supplement statistical computations. A 95% Confidence Interval and Statistical Significance of  $\alpha = 0.05$  where  $p < \alpha$  was chosen to reflect the statistical power of the study. Inter-Rater Reliability is the degree of agreement among raters, that is how closely the pain scores (ratings) given by each participant for each patient are to each



other. The data obtained was ranked and thus ordinal, requiring mostly non-parametric statistical measures. Furthermore, only a single rating was awarded by each rater on each patient at only one time. A confusion Matrix was applied to determine the sensitivity, specificity, over-rating, under-rating, and degree of variance as they were important measures to give direction and perspective to the Inter-Rater Reliability results. As there were 35 participants, each scoring 5 cases, it provided a total of 175 pain scores to be analysed and compared.

## **Results**

The researchers observed that the participants were scoring the standardised patients' pain intensity based on their facial expression of pain rather than the patients selecting the desired face on the Wong-Baker FACES® Pain Rating Scale. There were 30 (85.7%) male and 5 (14.3%) female paramedic participants. A total of 32 (91.4%) of the 35 participants were APs and 3 (8.6%) were CCPs. 13 (37.1%) of the participants received their initial basic degree medical training in Tunisia, 7 (20.0%) in Jordan, 5 (14.3%) in India, 4 (11.4%) in Philippines, 2 (5.7%) each in Morocco and South Africa and 1 (2.9%) each in the United States of America and Yemen. The mean years of practicing in the State of Qatar of the participants were 5 (range of 1-14). Overall Fleiss Kappa values indicate only a poor to slight agreement of the allocated pain scores among participants (raters) (Table 1). Not only are there poor agreement overall, the five patient cases individually showed equally poor agreement. Only the simulated case with pain score 10/10 received moderate/good agreement between participants (74.3%) (Table 2).

Each simulated case had a pre-determined reference pain score (i.e. rank), thus the correlation of ratings distributed among these cases provided some reference point. In all cases for both Spearman's and Kendall's correlation coefficients, values remained below 0.50, indicating a poor correlation of pain score distributions throughout all the cases (Table 3). The distribution of pain score allocations were equally varied throughout the participants' allocation, signifying that they were equally poor at agreeing or allocating the correct pain score throughout the group.

The null hypothesis for this study states that there is no significant difference between the raters, or between the cases. If  $p < \alpha$  ( $\alpha = 0.05$ ) and  $F\text{-distribution} > F\text{-critical}$  then the null hypothesis can be rejected, and vice versa. The ANOVA result indicates, raters ( $p = 3.381\text{E-}05$ )  $< (\alpha = 0.05)$  and  $F\text{-distribution} (2.667) > F\text{-critical} (1.516)$ , cases ( $p = 5.88\text{E-}39$ )  $< (\alpha = 0.05)$  and  $F\text{-distribution} (97.479) > F\text{-critical} (2.438)$ . Thus, the null hypothesis can be rejected and therefore there is a significant difference of pain score allocations between the raters and between the cases. The latter proves the hypothesis testing accurate as there were different pre-determined reference standards set for each case. There is no reliability between the participants (raters) when it comes to the allocation of pain scores based on the Wong-Baker FACES® Pain Rating Scale.

The confusion matrix indicates similar results to what was found through the Kappa and Correlation statistics (Table 4). It further describes the distribution of scores as seen in Table 2. Overall, sensitivity is poor to very poor throughout, except for case B (10/10 pain reference score) where 74.3% is regarded as good sensitivity (Fleiss and Cohen 1973, Landis and Koch 1977). Sensitivity as an exclusionary measure in this

instance shows how poorly the Wong-Baker FACES® Pain Rating Scale was applied by these participants as they could not accurately determine the correct pain score. Specificity on the other hand is very good throughout, as an inclusionary measure indicating poor delineation of pain through varied case presentations. The under- and over-score values coupled with the sensitivity and specificity values clearly indicate that the participants were not able to allocate the correct pain scores, and their scores were notably wide-spread throughout the Wong-Baker FACES® Pain Rating Scale (Table 2).

The degree off-set distributions were clustered around the reference standard (also evident in Table 2). Most of the allocations were 1 to 2 pain score levels away from the pre-determined reference standard, although not a normal distribution. However, some of the higher reference pain levels received lower scores, and vice versa (Table 5).

## **Discussion**

This pilot study aimed to determine the inter-rater reliability of the Wong-Baker FACES® Pain Rating Scale when applied on adult patients in the prehospital setting. Participants in this study scored the patients pain intensity using the tool based on the patient's facial expressions of pain rather than the patients themselves identifying their pain score based on the Wong-Baker FACES® Pain Rating Scale . Overall, the Inter-Rater Reliability as determined through Fleiss Kappa indicated only a poor to slight agreement of the allocated pain scores, as described against the reference standards. There was a wide grouping of the pain score levels around the reference standard. Most of the allocations were 1 to 2 pain score levels away from the reference standard,

although not in a normal distribution with some of the higher reference pain levels receiving lower scores, and vice versa. Ideally if the standardised patients scored their pain intensity, then their scores would have been the same throughout. Overall the sensitivity was poor to very poor throughout. Being able to appropriately assess the intensity of pain is essential to its effective management in the prehospital setting (Garra et al. 2010). Contextually, paramedics must overcome crucial barriers such as environmental factors, communication differences, cultural assumptions, and bias and ineffective use of the Wong-Baker FACES® Pain Rating Scale to successfully assess pain. The plethora of evidence demonstrates that paramedics can utilize various pain assessment tools/scales, and based on their respective scopes of practice; manage the patients' pain appropriately.

APs and CCPs in Qatar undergo extensive training on the use of the Wong-Baker FACES® Pain Rating Scale during their induction programme phase. These clinicians are then assessed theoretically and practically on the use of the Wong-Baker FACES® Pain Rating Scale prior to being certified to utilize the tool in operational duty. Further, memory aides in the form of laminated cards are distributed to all paramedics during their induction training to assist them with effective pain assessment. The Wong-Baker FACES® Pain Rating Scale is also contained in the HMCAS CPGs which is hosted on the ePCR for quick reference.

Although the tool was standard, the Inter-Rater Reliability results demonstrate that this sample had poor agreement when allocating pain scores using the Wong-Baker FACES® Pain Rating Scale, except for one standardised adult patient where the agreement was moderate to good. The Wong-Baker FACES® Pain Rating Scale, when

used with adults and not as intended, results in poor assessment of pain intensity of 5 simulated patients (French et al. 2013).

The pilot study findings could be attributable to the specific barriers within the prehospital setting in Qatar. Communicating with patients is challenging in certain instances. Although the HMCAS makes every effort to ensure that paramedic teams are multi-lingual, patients from certain linguistic groups are disadvantaged. The lack of identifying and understanding non-verbal cues may also have a negative effect on their ability to appropriately assess pain and use the tool as intended. There are also prevalent assumptions associated with certain ethnic groups and nationalities with regards to tolerance to pain, but nonetheless clinicians should irrespectively assess patients' pain according to their reported level of discomfort rather than based on subjective assumptions.

The paramedic's currency of knowledge on the appropriate use of the Wong-Baker FACES® Pain Rating Scale for assessing pain is also doubtful. Further EI on the correct use and interpretation of the Wong-Baker FACES® Pain Rating Scale is required at HMCAS to meet the International Patient Safety Goals as set out by the Joint Commission International. Essentially, achieving effective pain relief and patient comfort is critical to efficient emergency medical care. However, inaccurate pain scores and the inability to appropriately assess pain may translate into poor treatment, or incorrect treatment (Schyve 2007).

In addition, this study shows that the Wong-Baker FACES® Pain Rating Scale is possibly an inaccurate tool to determine the intensity of pain levels in adults. And if

utilized incorrectly, it will not detail an appropriate pain intensity score. More commonly utilized tools for assessing pain intensity in adults include the numerical rating scales and verbal rating scales (Hjermstad et al. 2011).

## **Limitations**

This was a pilot study, and thus only a limited number of clinicians and standardised patient cases were used. For statistical analysis, this provides problems when applying Fleiss Kappa and Correlation Coefficients as the picture can be skewed due to limited data and its variation. Six of the ten possibilities did not have a pre-determined reference simulation case; thus, the ranking order was inconsistent. The results only have bearing to this sample of cases at a single point in time. Further research to include a variety of reference cases inclusive of all the possible pain score allocations, and possibly multiples thereof would be recommended to get a clearer picture of the phenomenon. Test-retest can also be considered to factor in the effect of possible training on the use of the Wong-Baker FACES® Pain Rating Scale. Although the pain rating scale is specifically included in the HMCAS CPG, its instructions for use during patient care need to be carefully reviewed to ensure it explicitly specifies that it is the patient who is meant to show the facial expression corresponding to their level of pain.

## **Conclusion**

Based on the study population, the data has shown that the participants were not accurate in determining the correct pain score. Not only were they inaccurate, they were also unable to agree on the pain score to be allocated (regardless of the pre-determined reference standard) and thus were not precise. Pain score allocations were widely spread throughout the scale showing poor consistency and inter-rater

agreement on the score to be allocated. This could be attributed to the inaccurate use of the tool. It is this study finding that the participating paramedics were unreliable in the application of the Wong-Baker FACES® Pain Rating Scale to determine the pain levels of these 5 simulated cases. Since the pain score has direct impact on the treatment provided, it is notably concerning that incorrect pain score allocations can lead to inappropriate under- or over-management of pain and administration of analgesic therapy.

Further training for HMCAS staff on the Wong-Baker FACES® Pain Rating Scale is recommended. A further study should then be conducted with a larger sample size to determine their ability to accurately utilize the Wong-Baker FACES® Pain Rating Scale in determining the adult patient's pain intensity. HMCAS should also consider changing to an appropriate adult pain rating scale in the prehospital care setting with a culturally very diverse patient population.

## References

- Fleiss, J. L. & Cohen, J. 1973. The equivalence of weighted kappa and the intraclass correlation coefficient as measures of reliability. *Educational and Psychological Measurement*, 33, 613-619.
- French, S. C., Chan, S. B. & Ramaker, J. 2013. Education on prehospital pain management: a follow-up study. *Western Journal of Emergency Medicine*, 14, 96.
- French, S. C., Salama, N. P., Baqai, S., Raslavicus, S., Ramaker, J. & Chan, S. B. 2006. Effects of an educational intervention on prehospital pain management. *Prehospital Emergency Care*, 10, 71-76.

- Furyk, J. & Sumner, M. 2008. Pain score documentation and analgesia: a comparison of children and adults with appendicitis. *Emergency Medicine Australasia*, 20, 482-487.
- Gangaram, P., Alinier, G. & Menacho, A. 2017. Crisis resource management in relation to empowering people to speak up in emergency medical service clinical practice settings. *Journal of Paramedic Practice*, 9, 60-65.
- Garra, G., Singer, A. J., Taira, B. R., Chohan, J., Cardoz, H., Chisena, E. & Thode, H. C. 2010. Validation of the Wong-Baker FACES pain rating scale in pediatric emergency department patients. *Academic Emergency Medicine*, 17, 50-54.
- Goodman, A. 2015. The development of the Qatar healthcare system: a review of the literature. *International Journal of Clinical Medicine*, 6, 177.
- Hjermstad, M. J., Fayers, P. M., Haugen, D. F., Caraceni, A., Hanks, G. W., Loge, J. H., Fainsinger, R., Aass, N. & Kaasa, S. 2011. Studies comparing Numerical Rating Scales, Verbal Rating Scales, and Visual Analogue Scales for assessment of pain intensity in adults: a systematic literature review. *Journal of Pain and Symptom Management*, 41, 1073-1093.
- HMCAS 2017. Clinical Practice Guidelines. *In: Ambulance Service*. (ed.). Qatar: HMCAS.
- Hodkinson, M. 2016. Where is the paramedic profession going with pain management? *Journal of Paramedic Practice*, 8, 118-120.
- Iqbal, M., Spaight, P. A., Kane, R., Asghar, Z. & Siriwardena, A. N. 2015. Feasibility study of a novel pain assessment tool for improving prehospital pain management. *Emergency Medicine Journal*, 32.
- Landis, J. R. & Koch, G. G. 1977. The measurement of observer agreement for categorical data. *Biometrics*, 159-174.



- Lynde, J. & Zorab, O. 2015. Prehospital pain assessment and management: A quality improvement approach. *Emergency Medicine Journal*. 32.
- Mclean, S. A., Domeier, R. M., Devore, H. K., Hill, E. M., Maio & Frederiksen, S. M. 2004. The feasibility of pain assessment in the prehospital setting. *Prehospital Emergency Care*, 8, 155-161.
- MDPS 2016. Monthly figures on total population in Qatar. *In*: MDPS (ed.). Qatar.
- Schyve, P. M. 2007. Language differences as a barrier to quality and safety in health care: the Joint Commission perspective. *Journal of General Internal Medicine*, 22, 360-361.
- Tamayo-Sarver, J. H., Hinze, S. W., Cydulka, R. K. & Baker, D. W. 2003. Racial and ethnic disparities in emergency department analgesic prescription. *American Journal of Public Health*, 93, 2067-2073.
- Todd, K. H., Deaton, C., D'Adamo, A. P. & Goe, L. 2000. Ethnicity and analgesic practice. *Annals of Emergency Medicine*, 35, 11-16.
- Wilson, P., Alinier, G., Reimann, T. & Morris, B. 2017. Influential Factors on Urban and Rural Response Times for Emergency Ambulances in Qatar. *Mediterranean Journal of Emergency Medicine*, 26.

## Tables and Figures

Table 1: Fleiss Kappa per pain scores		
Pain Score	$\kappa$	CI (95%)
<b>Total</b>		(0.132 – 0.160)
<b>0</b>	No Value	
<b>1</b>	-0.005	(-0.041 – 0.030)
<b>2</b>	<b>0.019</b>	<b>(-0.016 – 0.055)</b>
<b>3</b>	0.202	(0.166 – 0.238)
<b>4</b>	<b>0.062</b>	<b>(0.026 – 0.098)</b>
<b>5</b>	0.027	(-0.008 – 0.063)
<b>6</b>	<b>0.017</b>	<b>(-0.018 – 0.053)</b>
<b>7</b>	<b>0.076</b>	<b>(0.040 – 0.112)</b>
<b>8</b>	-0.004	(-0.040 – 0.031)
<b>9</b>	0.037	(0.001 – 0.073)
<b>10</b>	<b>0.501</b>	<b>(0.465 – 0.537)</b>

Table 2: Characteristics of the ratings per simulated cases (n=35)												
Characteristic	All Cases (n=175 )		Case A (7/10)		Case B (10/10)		Case C (4/10)		Case D (2/10)		Case E (6/10)	
	N	%	N	%	N	%	N	%	N	%	N	%
Pain Score												
0		0.							2.			
1	1	6							1	9		
		4.					5.		<b>11</b>		2.	
2	7	0					2	7	<b>4</b>	<b>.4</b>	1	9
		16		5.			37		37			
3	28	.0	2	7			13	.1	13	.1		
		16		5.			<b>22</b>		25		25	
4	28	.0	2	7			<b>8</b>	<b>.9</b>	9	.7	9	.7
		14		11		2.	25		11		22	
5	26	.9	4	.4	1	9	9	.7	4	.4	8	.9
		8.		17		2.	5.		2.		<b>14</b>	
6	15	6	6	.1	1	9	2	7	1	9	<b>5</b>	<b>.3</b>
		10		<b>25</b>		2.	2.		2.		20	
7	19	.9	<b>9</b>	<b>.7</b>	1	9	1	9	1	9	7	.0
		6.		11		8.			5.		5.	
8	11	3	4	.4	3	6			2	7	2	7
		4.		11		8.						
9	7	0	4	.4	3	6						
		18		11		<b>74</b>					8.	
10	33	.9	4	.4	<b>26</b>	<b>.3</b>					3	6

Table 3: Correlation Coefficients per simulated cases					
Correlation	A	B	C	D	E
<b>Spearman's</b>					
A	-	0.05	0.29	0.43	0.47
B	0.05	-	0.02	-0.14	0.20
C	0.29	0.02	-	0.04	0.38
D	0.43	-0.14	0.04	-	0.29
E	0.47	0.20	0.38	0.29	-
<b>Kendall's</b>					
A	-	0.04	0.21	0.32	0.35
B	0.04	-	0.01	-0.12	0.17
C	0.21	0.01	-	0.04	0.31
D	0.32	-0.12	0.04	-	0.23
E	0.35	0.17	0.31	0.23	-
<b>Inter-Class Raters</b>	0.67 (0.41 – 0.95)				
<b>Inter-Class Cases</b>	0.08 (0.01 – 0.21)				

Table 4: Confusion Matrix per simulated case				
Case	Sensitivity, %	Specificity, %	Under-score, %	Over-score, %
<b>All</b>	29.7	92.9	32.6	37.7
<b>A</b>	25.7	92.9	40.0	34.3
<b>B</b>	74.3	95.0	25.7	
<b>C</b>	22.9	85.7	42.9	34.3
<b>D</b>	11.4	97.9	2.9	85.7
<b>E</b>	14.3	92.9	51.4	34.3

Table 5: Distribution pain score allocation off-set per simulated case														
Case	Under-score						Correct		Over-score					
	>-2		-2		-1		0		+1		+2		>+2	
	N	%	N	%	N	%	N	%	N	%	N	%	N	%
A	4	11.4	4	11.4	6	17.1	9	25.7	4	11.4	4	11.4	4	11.4
B	3	8.6	3	8.6	3	8.6	26	74.3						
C			2	5.7	13	37.1	8	22.9	9	25.7	2	5.7	1	2.9
D					1	2.9	4	11.4	13	37.1	9	25.7	8	22.9
E	1	2.9	9	25.7	8	22.9	5	14.3	7	20.0	2	5.7	3	8.6



**Figure 1. Wong-Baker FACES Pain Rating Scale(HMCAS 2017)**