

早産児の痛みのアセスメントのためのフェース・スケール (FSPAPI) の信頼性と妥当性の検証

阿部明子¹⁾, 横尾京子²⁾

キーワード (Key words) : 1. 早産児 2. 痛み 3. アセスメント・ツール
4. フェース・スケール 5. 信頼性と妥当性

本研究の目的は、上部顔面運動の三次元行動解析により、臨床で活用しやすい早産児の痛みのアセスメント・ツールとして開発したフェース・スケール (FSPAPI) の信頼性と妥当性を検証することである。

調査は、Ⅲ次レベルのNICUにおいて16名の早産児に実施された61の処置場面 (採血、皮下注射、栄養チューブ交換、電極除去) について、34名の看護師の観察によって行った。評定者間信頼性はベッドサイドでの看護師間の観察結果の一致度を、評定者内信頼性はベッドサイドの観察と同じ場面のビデオ録画との一致度を分析した。妥当性については、PIPPによる併存妥当性を分析した。

評定者間信頼性について、一致度は0.809 (信頼できる一致度)、 κ 係数は0.707 ($p < 0.001$) で「かなりの一致」を得ることができ、FSPAPIの同等性を確認できた。評定者内信頼性については、一致度は評定者間よりも低値で0.761と「満足できる一致度」ではなかったが、 κ 係数は0.626 ($p < 0.001$) であり、「かなりの一致」を得ることができた。

FSPAPIで重要となる皺形成の判断は、ベッドサイドで直接観察するよりもビデオ画像のほうが難しく、そのため評定者内一致度は低く出ることが予想されていたが、本研究ではかなりの一致を得ることができ、FSPAPIの安定性が得られたものと考えられる。

併存妥当性については、FSPAPIレベルとPIPPスコアの相関係数 (ρ) は0.766 ($p < 0.001$) であり、両者間に強い相関を確認することができた。

以上から、FSPAPIはPIPPよりも少ない項目、すなわち上部顔面表情のみを用いて、早産児に実施されるベッドサイド処置に伴う痛みのアセスメントに使用できると考えられる。

I. はじめに

早産児は、子宮外環境に適応するために、繰り返し痛みを伴う処置を受けなければならない。早産児の長期予後への影響を考えると、痛み体験を適切にアセスメントすることが重要である。そのためには、ベッドサイドで容易に使用でき、かつ集中治療中であっても制限を受けない指標を用いたアセスメント・ツールが必要である。

この要件に合致したツールとして、「早産児の痛みのアセスメントのためのフェース・スケール Faces Scale for Pain Assessment for Preterm Infants (FSPAPI)」を開発した¹⁾。

FSPAPIは、気管チューブ等の固定によって顔面表情筋の動きが制限されない上部顔面の表情に注目し、上部顔面表情運動の定量 (顔面積変化率) と定性 (皺形成) による3分類に、先行研究²⁾による2分類を加え、処置に対する反応を5つのレベルに分類したものである。

先行研究から加えたレベル0は「処置に対して皺形成

を認めない」、レベル4は「痛みの反応が消失し、顔面蒼白・全身弛緩の状態」である。レベル1の皺形成は「眉間」、レベル2「眉間・鼻根・下眼瞼下」、レベル3は「眉間・鼻根・額・下眼瞼下・上眼瞼」であった。

FSPAPIが臨床で用いられるには、皺形成を正しく確認しレベルの判断ができること、および痛みを伴う処置に対する反応を適切に捉えていることを確認する必要がある。そこで本研究では、FSPAPIの信頼性および妥当性を検証することとした。

II. 研究方法

1. 対象者

対象者は、総合周産期母子医療センターおよび小児専門病院のNICUに勤務する看護師と入院中の早産児とした。

早産児は、在胎32週未満 (相当体重) で出生し、先天異常や中枢神経系異常・鎮痛剤や鎮静剤の投与・手術

・ Verification of reliability and validity of a face scale for pain assessment of preterm infants (FSPAPI)

・ 所属 : 1) 広島大学病院 2) 広島大学大学院保健学研究科

・ 日本新生児看護学会誌 Vol.16, No.1 : 19 ~ 24, 2010

既往がなく、経過が安定し、両親より研究協力の承諾が得られた者とした。看護師は、研究協力への同意が得られ、FSPAPIの説明会に参加し、FSPAPIを完全に使用できるようになった者とした。説明会では、FSPAPIの使用法の説明と動画像を用いた観察の練習を行った。

2. データ収集法

データ収集は、3施設のNICUにおいて、平成16年8月4日～同年11月4日および平成17年3月8日～同年3月17日に行った。早産児の背景や経過は診療録から、看護師の背景は質問紙により収集した。心拍数と酸素飽和度はパルスオキシメータ（Oximax N595 NELLCOR）で測定し、数値を2秒毎にパソコンに取り込んだ。処置場面と顔面表情は、早産児の頭側左右に1台ずつビデオカメラ（SONY LCH-VX2000）を三脚で固定し、撮影した。処置開始までに、ビデオカメラ、パルスオキシメータ、パソコンの内蔵時計を合わせた。

処置は、採血（穿刺・搾り）、皮下注射（針刺し・注入・採みこみ）、栄養チューブ抜去、栄養チューブ挿入、電極除去とした。研究者は、痛み刺激が加わった時点と特定するためにライトを点灯させ、また顔面表情の観察時点を定めるために、処置中に顔面表情に変化を認めた時点を口頭で示し（合図）、ビデオカメラに録音した。処置実施者は、手順を口頭で示し、それをビデオカメラに録音した。

NICU入室時および撮影中の操作は、当該施設の感染防止マニュアルに従った。室内および保育器内の照度と音レベルは、撮影開始前に、照度計（HIOKI 3423ルクスハイテスタ）、騒音計（RION積分騒音計NL-06）を用いて測定した。

FSPAPIを用いた観察

看護師によるFSPAPIを用いた観察は、ベッドサイドとビデオ画面で行った。ベッドサイドでは、1～3名の看護師が処置場面に立ち会い、各々独自で、研究者が合図した時点の顔面表情を観察した。ビデオ画面（当該看護師が立ち会った処置の録画）での観察は、看護師の都合に合わせて、ベッドサイド観察から数日後に行った。看護師は、ビデオに録音された合図で、ベッドサイドと同様に顔面表情の観察を行った。ベッドサイドでの記憶が影響しないよう、当該看護師が観察していない録画も加え、観察した。

PIPPを用いた観察

PIPP（Premature Infant Pain Profile）の項目中、修正年齢・state・心拍数・酸素飽和度・顔面表情（3要素）の7項目は、診療録、ビデオ画面、パルスオキシメータによって収集した値を用いて採点した。なお、ビデオ画面を用いたstate・顔面表情の分析は、研究者と協力者計3名が一致するまで検討した。

3. 分析方法

各観察場面のFSPAPIレベルの決定は、研究者と協力者計3名が観察し、一致した結果を用いた。検定には、統計ソフトSPSS11.5Jを用い、5%水準をもって有意とした。

1) 信頼性の検証

評定者間信頼性の検証は、レベル1, 2, 3について、ベッドサイドで同じ処置場面に立ち会った看護師2名の観察結果から、一致度とkappa統計量を算出した。評定者内信頼性の検証は、レベル1, 2, 3について、各看護師のベッドサイドとビデオ画面での観察結果から、一致度とkappa統計量を算出した。一致度の算出には、単純な一致度＝一致した観察数／全観察数を用い、kappa統計量の算出には、 $\kappa = (\text{判断の一致率} - \text{偶然の一致率}) / (1 - \text{偶然の一致率})$ を用いた。不一致があった場合には、当該看護師に観察や判断状況の聞き取りを行った。

2) 併存妥当性の検証

併存妥当性の検証は、研究者がFSPAPIとPIPPを用いて観察した結果を用い、FSPAPIレベルとPIPPスコアにおけるSpearmanの順位相関係数を算出した。

FSPAPIでは、4通りのデータを用いた。レベル1～4については、「皺形成なし（0-1）」と「収縮性の変化以外の反応（0-2）」「皺形成なし（0-1）」と「収縮性の変化以外の反応（0-2）」レベル0の場合には、「処置前途変わらず皺形成なし（0-1）」と「収縮性の変化以外の反応（0-2）」に分類し、統合した場合と分割した場合の2通りのデータを用いた。

PIPPでは、総得点と顔面表情3要素の得点の2通りのデータを用いた。

4. 倫理的配慮

本研究は、広島大学大学院保健学研究科倫理審査委員会および協力施設の倫理委員会へ申請し、承認を得た。看護師長から看護師および早産児の紹介を受けた。早産児の父親または母親、看護師に、文書を用いて研究目的・方法・意義・守秘義務・研究協力への任意性および中断の自由・結果の公表について説明し、研究協力への承諾を文書で得た。研究協力を撤回する場合のために、同意取り消し書の使い方も説明した。

観察対象とする処置は予定されたものに限り、調査のために加えることはしなかった。また、調査が経過に影響すると判断された場合は、データ収集開始前やデータ収集途中で調査を中止した。

プライバシー保護のため、入手した情報は本研究目的に限り使用し、調査用紙やビデオテープは研究者が責任をもって鍵のかかる場所に保管した。データ保管や結果公表の際に対象者が特定できないよう、コード番号によ

る識別などの対処をした。

IV. 結果

1. 対象者の背景とデータ収集場面の概要

観察に参加した看護師は、3施設のNICU看護師35名であったが、その内1名はFSPAPIレベル0のみの観察であったため除外し、34名の観察結果を分析に用いた。看護師34名の年齢は 28.4 ± 5.6 歳（平均値 \pm 標準偏差）、臨床経験年数は 7.0 ± 6.0 年、NICU経験年数は 5.3 ± 5.1 年であり、3施設間で差は認められなかった。（表1）

早産児18名に72場面で処置が実施された。レベル1, 2, 3が出現したのは72場面中55場面（早産児16名）であった。この内、評定者間信頼性の検証には、2名以上の看護師が立ち会った29場面、評定者内信頼性の検証には、1名以上の看護師が観察に立ち会った39場面を用いた。妥当性の検証には、72場面中、生理学的データの収集が可能であった61場面を用いた。表2に、観察時の概要を示した。

ベッドサイドでは、看護師1人あたり 2.1 ± 1.5 場面、 3.4 ± 2.8 回の観察を行った。ビデオ画面での観察が行われたのは、ベッドサイド観察後 6.0 ± 3.4 日であった。

2. 信頼性および妥当性の分析結果

表3に、信頼性の分析結果を示した。評定者間の一致度は、レベル1は1.000、レベル2は0.750、レベル3は0.826、レベル1～3をまとめた一致度は0.809であった。 κ 係数は0.707 ($p < 0.001$)であった。観察結果の不一致は、レベル1と2（5場面）、レベル2と3（観察数7）で認められた。

評定者内の一致度は、レベル1は0.762、レベル2は0.640、レベル3は0.891、レベル1～3をまとめた一致度は0.761であった。 κ 係数は0.626 ($p < 0.001$)であった。観察結果の不一致は、レベル1と2（観察数12）、レベル2と3（観察数16）で認められた。不一致の理由として、判断基準となる皺形成を見逃していたこと、およびビデオ画面が暗く、眉弓のふくらみや皺形成がはっきりしなかった／強くしかめているように見えたなど、ビデオ画像を用いた方法上の理由が述べられた。

表4にFSPAPIのレベル別にみたPIPPの得点、表5にFSPAPIとPIPPの相関係数を示した。FSPAPIのレベルとPIPPの総得点の相関は、レベル1の0-1と0-2を統合し、「最も強く現れたレベル」でみると $\rho = 0.766$ ($p < 0.001$)、「最も長く現れたレベル」は $\rho = 0.725$ ($p < 0.001$)であった。レベル1の0-1と0-2を分割し、「最も強く現れたレベル」でみると $\rho = 0.762$ ($p < 0.001$)、「最も長く現れたレベル」は $\rho = 0.762$ ($p < 0.001$)、「最も長く現れたレベル」は $\rho = 0.762$ ($p < 0.001$)、

表1. 看護師の背景

	施設1	施設2	施設3	全体
看護師数	12名	12名	10名	34名
年齢	27.6 ± 4.9 歳	28.4 ± 7.4 歳	29.4 ± 4.2 歳	28.4 ± 5.6 歳
臨床経験年数	6.3 ± 5.2 年	7.0 ± 7.7 年	7.7 ± 5.1 年	7.0 ± 6.0 年
NICU経験年数	3.9 ± 2.9 年	6.1 ± 6.9 年	6.0 ± 5.0 年	5.3 ± 5.1 年

数値：平均値 \pm 標準偏差

表2. 観察時の概要

		評定者間信頼性	評定者内信頼性	併存妥当性	
早産児数		12名	16名	13名	
処置場面数		29場面	39場面	61場面	
早産児の背景	出生時	在胎週数	26.8 ± 2.4 週	27.5 ± 2.6 週	27.6 ± 2.1 週
		出生体重	906 ± 331 g	992 ± 341 g	990 ± 322 g
		Ap 1分値	4.8 ± 2.3 点	4.9 ± 2.3 点	4.9 ± 2.2 点
		Ap 5分値	6.7 ± 2.6 点	7.0 ± 2.4 点	7.2 ± 2.2 点
	観察時	修正齢	32.3 ± 3.1 週	32.5 ± 3.0 週	32.0 ± 2.8 週
		日齢	40.8 ± 23.5 H	38.8 ± 23.9 日	38.1 ± 22.3 日
		体重	$1,127 \pm 515$ g	$1,166 \pm 516$ g	$1,116 \pm 470$ g
	人工呼吸器装着	18場面	24場面	41場面	
環境条件	器内	温度	$32.3 \pm 1.5^\circ\text{C}$	$32.2 \pm 1.4^\circ\text{C}$	$32.2 \pm 1.4^\circ\text{C}$
		照度	127.8 ± 86.4 Lx	122.1 ± 82.2 Lx	110.6 ± 74.0 Lx
		音レベル	62.9 ± 9.0 dB	63.4 ± 8.9 dB	64.8 ± 8.7 dB
	室内	温度	$26 \sim 28^\circ\text{C}$	$26 \sim 28^\circ\text{C}$	$26 \sim 28^\circ\text{C}$
		照度	201.4 ± 113.1 Lx	188.8 ± 109.3 Lx	174.9 ± 107.6 Lx
		音レベル	65.6 ± 23.9 dB	64.1 ± 21.0 dB	63.7 ± 17.1 dB

数値：平均値 \pm 標準偏差 Ap：アプガースコア

表3. 評定者間および評定者内信頼性の分析結果

施設		レベル	評定者間信頼性				評定者内信頼性			
			レベル1	レベル2	レベル3	全体	レベル1	レベル2	レベル3	全体
施設1	観察数	3	7	7	17	9	13	12	34	
	一致数	3	5	5	13	7	11	10	28	
	一致度	1.000	0.714	0.714	0.762	0.778	0.846	0.833	0.824	
	κ 係数	-	-	-	0.642	-	-	-	0.734	
施設2	観察数	2	10	7	19	4	15	11	30	
	一致数	2	7	7	16	2	8	11	21	
	一致度	1.000	0.700	1.000	0.842	0.500	0.533	1.000	0.700	
	κ 係数	-	-	-	0.749	-	-	-	0.502	
施設3	観察数	3	15	9	27	8	22	23	53	
	一致数	3	12	7	22	7	13	20	40	
	一致度	1.000	0.800	0.778	0.815	0.875	0.591	0.870	0.755	
	κ 係数	-	-	-	0.721	-	-	-	0.612	
全体	観察数	8	32	23	63	21	50	46	117	
	一致数	8	24	19	51	16	32	41	89	
	一致度	1.000	0.750	0.826	0.809	0.762	0.640	0.891	0.761	
	κ 係数	-	-	-	0.707	-	-	-	0.626	

表4. FSPAPIのレベル別にみたPIPPの得点

FSPAPI レベル		PIPPの得点	総得点	顔面3要素	BB	ES	NLF
レベル0	15場面		6.6 ± 1.2	0.1 ± 0.3	0	0	0.1 ± 0.3
0-1	(5)		6.6 ± 0.5	0	0	0	0
0-2	(10)		6.6 ± 1.4	0.1 ± 0.3	0	0	0.1 ± 0.3
最強	レベル1	5場面	6.8 ± 2.2	1.0 ± 0.7	1.0 ± 0.7	0	0
	レベル2	18場面	9.3 ± 1.6	3.6 ± 1.2	1.6 ± 0.8	1.3 ± 0.7	0.4 ± 0.5
	レベル3	23場面	12.4 ± 2.8	5.5 ± 2.4	2.3 ± 0.9	1.6 ± 1.1	1.7 ± 1.1
最長	レベル1	13場面	8.8 ± 2.8	2.6 ± 1.8	1.7 ± 1.0	0.6 ± 0.8	0.3 ± 0.5
	レベル2	16場面	9.6 ± 2.0	3.3 ± 1.4	1.6 ± 0.7	1.3 ± 0.7	0.5 ± 0.6
	レベル3	17場面	12.9 ± 2.9	6.1 ± 2.2	2.3 ± 0.8	1.8 ± 0.8	2.0 ± 0.9

BB: Brow Bulge (眉弓のふくらみ) ES: Eye Squeeze (眼の圧縮) NLF: Naso-Labial Furrow (鼻唇溝)
 レベル0-1: 処置前と変わらず皺形成なし レベル0-2: 収縮性以外の動きを認める
 最強: 30秒間で最も強く現れたレベル 最長: 30秒間で最も長く現れたレベル 数値: 平均値 ± 標準偏差

表5. 併存妥当性の分析結果 (FSPAPIとPIPPの相関)

レベル0-1と0-2	分析視点	PIPP 総得点との相関係数	PIPP 顔面3要素の得点との 相関係数
統合した場合	最強	$\rho = 0.766^*$	$\rho = 0.840^*$
	最長	$\rho = 0.725^*$	$\rho = 0.827^*$
分割した場合	最強	$\rho = 0.762^*$	$\rho = 0.837^*$
	最長	$\rho = 0.722^*$	$\rho = 0.824^*$

レベル0-1: 処置前と変わらず皺形成なし レベル0-2: 収縮性以外の動きを認める
 最強: 30秒間で最も強く現れたレベル 最長: 30秒間で最も長く現れたレベル
 ρ : Spearmanの相関係数 * : $p < 0.001$

($p < 0.001$)であった。

また、顔面表情3要素の得点との相関は、レベル1の統合で「最も強く現れたレベル」では $\rho = 0.840$ ($p < 0.001$), 「最も長く現れたレベル」は $\rho = 0.827$ ($p < 0.001$)であった。レベル1の分割で「最も強く現れたレベル」では $\rho = 0.837$ ($p < 0.001$), 「最も長く現れたレベル」は $\rho = 0.824$ ($p < 0.001$)であった。

V. 考察

本研究の目的は、臨床で早産児の痛みのアセスメントが容易にできるよう開発されたFSPAPIの信頼性および妥当性を検証することであった。

評定者間信頼性について、一致度は0.809であり、「信頼性のある一致度」と言える³⁾。また、 κ 係数は0.707

($p < 0.001$)であり、「かなりの一致」を得ることができた⁴⁾。したがって、FSPAPIの同等性が確認できたと言える。一方、評定者内信頼性については、一致度は評定者間よりも低値で0.761であり、これは「満足できる一致度」ではなかった。しかし、 κ 係数は0.626 ($p < 0.001$)であり、「かなりの一致」を得ることができた。

評定者内の一致度が低値であった理由の一つに、ビデオ画像を用いた方法上の理由が述べられた。ビデオ画像による観察では、カメラの角度によって事象や状況について偏った見方がなされうるという欠点がある⁵⁾。特に、FSPAPIで重要となる皺形成の判断は、ベッドサイドで直接観察するよりもビデオ画像のほうが難しく、そのため評定者内一致度は低く出ることが予想されていた。しかし本研究では、かなりの一致を得ることができ、FSPAPIの安定性が得られたものと考ええる。

妥当性については、修正28～40週の早産児において信頼性および妥当性が検証されているPIPP⁶⁻⁷⁾を用いて併存性について検証した。PIPPは優れたツールであり、研究のためによく用いられている。しかし、顔面表情については、3つの要素(Brow Bulge・Eye Squeeze・Naso-Labial Furrow)各々の30秒間での出現割合を算出し得点化しなければならないため、臨床で処置の進行に合わせ前向きに痛みをアセスメントすることは難しく、また心拍数や酸素飽和度をモニタリングする必要がある。こうしたPIPPの臨床的限界があることからFSPAPIを開発し、PIPPとの併存妥当性の検証を計画した。

FSPAPIとPIPPの相関係数は、8通りの組み合わせすべてにおいて0.700以上であり、両者間に強い相関を確認することができた。したがって、FSPAPIはPIPPよりも少ない項目、すなわち上部顔面表情のみを用いて、早産児のベッドサイド処置に伴う痛みのアセスメントに使用できると考える。

今後の課題はFSPAPIの臨床への普及と実用化であり、それとともに、医療チーム全体で痛みのアセスメントやケアに取り組めるよう働きかけていく必要があると考える。

VI. 結 論

本研究において、FSPAPIの高い評定者間信頼性およ

び併存的妥当性が確認され、評定者内信頼性についても、ビデオテープによる場面の再現という方法上の限界はあったが、確認することができた。したがってFSPAPIは、早産児に実施されるベッドサイド処置に伴う痛みのアセスメントに使用できると考える。

謝 辞

稿を終えるにあたり、本研究に協力してくださいました赤ちゃんのご両親、看護師の皆様方、施設の先生方や看護師長様に深謝いたします。

本研究は、平成17年度科学研究費補助金(基盤研究(B)(2)) (課題番号15390672)を受けて実施した研究の一部である。

引用文献

- 1) 横尾京子, 阿部明子 (2010): 早産児の痛みのアセスメント・ツール (FSPAPI) の開発: 上部顔面表情運動の定量に基づいたフェース・スケール. 日本新生児看護学会誌, 16(1): 11-18.
- 2) 横尾京子 他 (2001): 気管内挿管中の早産児の痛みを伴う処置に対する顔表情: フェース・スケールの試案作成. 日本新生児看護学会誌, 8(1): 41-48.
- 3) 中澤 潤 (2001): 時間見本法の理論と技法. 中澤 潤ら編: 心理学マニュアル; 観察法, 第1章, 北大路書房, 京都, p21.
- 4) Fleiss JL, Cohen J & Everitt BS (1969): Large sample standard errors of kappa and weighted kappa. Psychological Bulletin, 72, 323-327.
- 5) Polit DF & Hungler BP, 近藤潤子監訳 (1999): 看護研究原理と方法. 医学書院, 東京, pp239-264.
- 6) Stevens B, Johnston C, Grunau R (1996): Premature infant pain profile: Development and initial validation. Clinical Journal of Pain 12:13-22.
- 7) Ballantyne M, et al. (1999): Validation of the premature infant pain profile in the clinical setting. Clinical Journal of Pain 15:297-303.

Verification of reliability and validity of a face scale for pain assessment of preterm infants (FSPAPI)

Akiko Abe¹⁾ Kyokoo Yokoo²⁾

1) Hiroshima University Hospital

2) Hiroshima University, Graduate School of Health Sciences

Key words : 1. preterm infant 2. pain 3. assessment tool
4. face scale 5. reliability and validity

The purpose of this study is to examine reliability and validity of the face scale for pain assessment of preterm infants developed by the applicant based on quantification of facial motions by 3-D behavioral analysis in order to facilitate pain assessment of preterm infants in clinical settings.

The investigation was conducted at three tertiary-level NICU facilities, with 61 recorded images of scenes from 16 preterm infants and observational data of 34 nurses. As for intra-evaluator reliability, agreement was 0.761, κ coefficient 0.626 (pretty much agree), whereas for inter-evaluator reliability, agreement was 0.809 (agreement with reliability) and κ coefficient 0.707 (pretty much agree). Dependence validity was compared with PIPP and was strongly correlated, with correlation coefficient (ρ) of 0.766 ($p < 0.001$).

The reason why intra-evaluator reliability was lower than inter-evaluator reliability, which was high, was considered to be because of the poor quality of recorded images which were used for reproducing the scenes, due to low intensity of illumination at NICU. However, because of the good inter-evaluator reliability and good correlation with PIPP used for research, the scale developed by the applicant was considered to be convenient in clinical settings.