Pain Management Nursing xxx (xxxx) xxx

Contents lists available at ScienceDirect

Pain Management Nursing

journal homepage: www.painmanagementnursing.org

Original Article

Content Validity of the Spanish Adaptation of the Premature Infant Pain Profile Revised

Irene Núñez-López, M.Sc., P.N.P., R.N.[*,†], María-Gema Cid-Expósito, Ph.D., M.Sc., P.N.P., R.N.[‡,§,##], Raquel Abalo, Ph.D.[‖,¶,#,**,††], Ana Serrano-Gutiérrez, P.N.P., R.N.[*], Lucía Jiménez-Fernández, M.Sc., NIDCAP & N.N.P., R.N.[*], Laura Collados-Gómez, Ph.D., M.Sc., P.N.P., R.N.[*,‡‡,§§]

[*] *Neonatal Intensive Care Department, Doce de Octubre University Hospital, Madrid, Spain*
[†] *Doctor of Health Sciences Programme, International Doctoral School, Rey Juan Carlos University, Madrid, Spain*
[‡] *Department of Nursing and Stomatology, Faculty of Health Sciences, Rey Juan Carlos University, Madrid, Spain*
[§] *Nursing Research Group, Gregorio Marañón Health Research Institute (IiSGM), Madrid, Spain*
[‖] *Department of Basic Health Sciences, Faculty of Health Sciences, Rey Juan Carlos University (URJC), Madrid, Spain*
[¶] *High Performance Research Group on Pathophysiology and Pharmacology of the Digestive Tract (NeuGut-URJC), Rey Juan Carlos University (URJC), Madrid, Spain*
[#] *R&D&I Unit associated with the Institute of Medicinal Chemistry (IQM), Spanish National Research Council (CSIC), Madrid, Spain*
[**] *Spanish Pain Society Working Group on Basic Sciences in Pain and Analgesia, Madrid, Spain*
[††] *Spanish Pain Society Working Group on Cannabinoids, Madrid, Spain*
[‡‡] *Department of Nutrition, Faculty of Biomedicine and Nursing, Universidad Europea de Madrid, Madrid, Spain*
[§§] *Invecuid Care Research Group, Hospital 12 de Octubre Health Research Institute (imas12), Madrid, Spain*

ARTICLE INFO

Article history:
Received 25 February 2023
Received in revised form 15 June 2023
Accepted 30 June 2023
Available online xxx

ABSTRACT

Background: To the best of our knowledge, there are no validated neonatal pain assessment scales in Spanish. Given the need for such a scale, a study was undertaken to adapt and validate the Premature Infant Pain Profile-Revised (PIPP-R) scale. After translation and back-translation, content validity was addressed, a crucial phase in validation studies, in which researchers examine whether the items that make up the scale represent the content that the scale is intended to assess.

Aims: The aim was to provide evidence for the content validity of the Spanish adaptation of the PIPP-R scale.

Method: The study used the Delphi technique with 10 experts. Data collection was anonymous and was conducted through an online platform. It was an ad hoc survey consisting of four questions, with a five-point Likert scale for each item on the scale and for the instruction table. An item-content validity index (I-CVI) and a scale-content validity index (S-CVI) were calculated for the analysis.

Results: After two rounds of the survey, all items exceeded an I-CVI of 0.9. The S-CVI value was 0.98 (±0.03) for the scale, and 1 for its instruction table. The kappa index yielded values indicating an excellent degree of agreement.

Conclusions: The Spanish version of the PIPP-R obtained a high degree of content validity according to the expert group and the Delphi technique.

Pain in the neonatal stage, from birth to 28 days of life, is a complex phenomenon from every point of view: biologic, psychological, and social. Acute episodic pain has recently been described by a group of experts (Ilhan et al., 2022) as an unpleasant sensory and emotional response associated with a procedure or event, eliciting physiologic characteristics in the neonate linked to the production of tissue damage, which may be actual or potential. The procedure is described as being associated with tissue damage, but this is not always necessarily the case. This group of experts also

[##] Address correspondence to María-Gema Cid-Expósito, Ph.D., M.Sc., P.N.P., R.N., Department of Nursing and Stomatology, Faculty of Health Sciences, Rey Juan Carlos University. Avenida de Atenas s/n, 28922, Madrid, Spain.
E-mail address: gema.cid@urjc.es (M.-G. Cid-Expósito).

https://doi.org/10.1016/j.pmn.2023.06.012

believe that the painful response could arise from a medical condition. Valid and reliable assessment tools are needed to assess pain among newborns given the multidimensionality of pain, the individual subjective nature of pain, and the inability of neonates to verbally express their pain level.

To the best of our knowledge, no validated scales are available for this purpose in the Spanish health care context. A study on the clinical assessment of pain in Spanish neonatal intensive care units (Avila-Alvarez et al., 2016) concluded that pain assessment scale use was low and that there was great variability in the scales used among the study units. Only 16.7% of the 468 neonates in the study sample were assessed for pain at least once. This trend is confirmed by a later study (Castillo Barrio et al., 2020) which concluded that clinical pain assessment scales in Spanish neonatal units are underused.

In view of the demand for validated scales for neonatal pain assessment, a multi-center study was undertaken (Núñez-López et al., 2022) to carry out the translation and cross-cultural validation of the Premature Infant Pain Profile-Revised (PIPP-R) neonatal pain assessment scale, which had been validated in the original language, English (Gibbins et al., 2014). Once the author, Stevens, had given her approval and the first stages of translation and back-translation of the scale had been completed, content validity was assessed.

The PIPP-R scale was selected for this study for the following reasons: it is suited to the characteristics of hospitalized neonatal patients in Spain; one of its scoring items takes the gestational age of the patient into consideration; its validity and reliability for use in English had been demonstrated; it is the most widely used scale (Giordano et al., 2019) both for assessing neonatal pain in hospitalized patients and conducting research on the topic.

The PIPP-R scale for the assessment of neonatal acute and procedural pain (Gibbins et al., 2014; Stevens et al., 1996) consists of 7 multidimensional items and a table detailing instructions for use. Three of the scale items are behavioral (facial gestures), two are physiological (heart rate and oxygen saturation), and two are contextual (corrected gestational age and baseline behavior). Behavior and physiologic items are scored numerically on a four-point scale (0, 1, 2, and 3) to reflect changes in each variable from reference or baseline values. Contextual items are also scored on a four-point scale (0, 1, 2, and 3) before the painful procedure and/or manipulation but will only be taken into consideration if the sum of the five aforementioned items is greater than 0 (supplementary material 1). This was one of the modifications present in the PIPP-R (Gibbins et al., 2014) with respect to the PIPP (Stevens et al., 1996), as the previous version tended to produce false positives due to the influence of the gestational age item. This was because the baseline behavior of an extremely preterm newborn is to sleep peacefully. However, while this category was given a score of 3, the gestational age item added another 3, resulting in a total score of 6 points, irrespective of the other indicators. On the PIPP-R, a score between 0 and 6 will be considered as no pain or mild pain; a score between 7 and 12 will be considered as moderate pain; and a score between 13 and 21 will be considered as severe pain.

Establishing content validity is a crucial step in the development of a new scale or the cross-cultural adaptation of an existing scale (Polit & Beck, 2006). It is an important issue for clinicians and researchers who require high-quality measurements. Content validity provides insight into whether the items that make up a scale represent the content that the scale is intended to assess. In our case, this is whether the items described in the Spanish version of the PIPP-R and its instruction table adequately represent the pain assessment in neonatal patients (Hyrkäs et al., 2003; Mokkink et al., 2010; Waltz et al., 2010).

Table 1
Adapted Translation by the Authors Based on the Literature on the Topic (Quatrini Carvalho Passos Guimarães et al., 2016).

• Criterion	Score
Doctorate	4 points
Doctoral thesis on neonatal pain	1 point
Clinical experience with neonates	1 point per year
Research projects in neonatal pain	1 point
Publications in neonatal pain	1 point
Specific training in neonatal pain	2 points
Participation in a working group on neonatal pain	1 point
Delivery of specific training in neonatal pain	2 points

There are various methods for assessing the content validity of a scale: multidimensional scaling techniques (Deville & Prometric, 1996), factor analysis (Dorans & Lawrence, 1987), and panels of experts (Lynn, 1986). The latter method, suggested by Lynn (1986), involves several measurements by a number of experts who confirm and approve the content validity of both the items of the scale and the instrument as a whole. This approach has been accepted and validated in the literature (Polit et al., 2007; Waltz et al., 2010).

The Delphi technique is a prospective, originally qualitative method that seeks to reach consensus among a group of experts on the subject matter under discussion through analysis and reflection. It was first defined in 1975 (Linstone & Turrof, 1975) as "a method for structuring a group communication process so that the process is effective in allowing a group of individuals, as a whole, to deal with a complex problem".

Methods

Objective

The aim of the study was to establish the content validity of the cross-cultural Spanish adaptation of the PIPP-R scale using an expert panel.

Study Design

A descriptive study was conducted from November 2020 to March 2021 to establish the content validity of the cross-cultural adaptation of the PIPP-R scale. The validation method adopted was expert group validation using the Delphi technique.

Expert Sample

The existing literature on the topic was consulted to determine the number of experts required to complete the sample. Lynn (1986) recommends that the proportion of experts agreeing on the content validity of an item and the entire instrument should be equal to the number of items in the scale, plus the standard error of that proportion. Given that the PIPP-R scale consists of seven items, a sample of 10 experts was required, as recommended by Lynn.

A total of 15 experts in pain and/or neonatal pain were identified who met one or more of the following criteria: being a member of working groups on neonatal pain (in collaboration with the Spanish Society of Neonatal Nursing); having published scientific papers and/or having conducted research projects on pain and/or neonatal pain; having extensive clinical experience in neonatology and/or experience validating pain scales. They were individually invited by e-mail to participate in the working group and to provide information about their expertise criteria (Quatrini Carvalho Passos Guimarães et al., 2016), as shown in Table 1. In the end, 10 experts confirmed their participation in the study.

Please cite this article as: I. Núñez-López, M.-G. Cid-Expósito, R. Abalo et al., Content Validity of the Spanish Adaptation of the Premature Infant Pain Profile Revised, Pain Management Nursing, https://doi.org/10.1016/j.pmn.2023.06.012

Table 2
Criteria for Evaluating Scale Items.

Validation criterion	Relevance	Clarity	Simplicity	Ambiguity
Changes in heart rate (beats per minute)	1 = not relevant 2 = item needs revision 3 = relevant but needs revision 4 = very relevant	1 = unclear 2 = item needs revision 3 = clear but needs revision 4 = very clear	1 = not simple 2 = item needs revision 3 = simple but needs revision 4 = very simple	1= ambiguous 2 = item needs revision 3 = unambiguous but needs revision 4 = unambiguous
Decrease in oxygen saturation (percentage %)	1 = not relevant 2 = item needs revision 3 = relevant but needs revision 4 = very relevant	1 = unclear 2 = item needs revision 3 = clear but needs revision 4 = very clear	1 = not simple 2 = item needs revision 3 = simple but needs revision 4 = very simple	1 = ambiguous 2 = item needs revision 3 = unambiguous but needs revision 4 = unambiguous
intensity of brow bulge (seconds)	1 = not relevant 2 = item needs revision 3 = relevant but needs revision 4 = very relevant	1 = unclear 2 = item needs revision 3 = clear but needs revision 4 = very clear	1 = not simple 2 = item needs revision 3 = simple but needs revision 4 = very simple	1 = ambiguous 2 = item needs revision 3 = unambiguous but needs revision 4 = unambiguous
Intensity of eye squeeze (seconds)	1 = not relevant 2 = item needs revision 3 = relevant but needs revision 4 = very relevant	1 = unclear 2 = item needs revision 3 = clear but needs revision 4 = very clear	1 = not simple 2 = item needs revision 3 = simple but needs revision 4 = very simple	1 = ambiguous 2 = item needs revision 3 = unambiguous but needs revision 4 = unambiguous
Intensity of naso-labial furrow (seconds)	1 = not relevant 2 = item needs revision 3 = relevant but needs revision 4 = very relevant	1 = unclear 2 = item needs revision 3 = clear but needs revision 4 = very clear	1 = not simple 2 = item needs revision 3 = simple but needs revision 4 = very simple	1 = ambiguous 2 = item needs revision 3 = unambiguous but needs revision 4 = unambiguous
corrected gestational age (weeks + days)	1 = not relevant 2 = item needs revision 3 = relevant but needs revision 4 = very relevant	1 = unclear 2 = item needs revision 3 = clear but needs revision 4 = very clear	1 = not simple 2 = item needs revision 3 = simple but needs revision 4 = very simple	1 = ambiguous 2 = item needs revision 3 = unambiguous but needs revision 4 = unambiguous
Baseline behavior	1 = not relevant 2 = item needs revision 3 = relevant but needs revision 4 = very relevant	1 = unclear 2 = item needs revision 3 = clear but needs revision 4 = very clear	1 = not simple 2 = item needs revision 3 = simple but needs revision 4 = very simple	1 = ambiguous 2 = item needs revision 3 = unambiguous but needs revision 4 = unambiguous

Table 3
Criteria for Evaluating Items in the Instructions for use Table.

Validation criterion	Relevance	Clarity	Simplicity	Ambiguity
Step 1	1 = not relevant 2 = item needs revision 3 = relevant but needs revision 4 = very relevant	1 = not relevant 2 = item needs revision 3 = relevant but needs revision 4 = very relevant	1 = not relevant 2 = item needs revision 3 = relevant but needs revision 4 = very relevant	1 = ambiguous 2 = item needs revision 3 = unambiguous but needs revision 4 = unambiguous
Step 2	1 = not relevant 2 = item needs revision 3 = relevant but needs revision 4 = very relevant	1 = not relevant 2 = item needs revision 3 = relevant but needs revision 4 = very relevant	1 = not relevant 2 = item needs revision 3 = relevant but needs revision 4 = very relevant	1 = ambiguous 2 = item needs revision 3 = unambiguous but needs revision 4 = unambiguous
Step 3	1 = not relevant 2 = item needs revision 3 = relevant but needs revision 4 = very relevant	1 = not relevant 2 = item needs revision 3 = relevant but needs revision 4 = very relevant	1 = not relevant 2 = item needs revision 3 = relevant but needs revision 4 = very relevant	1 = ambiguous 2 = item needs revision 3 = unambiguous but needs revision 4 = unambiguous
Step 4	1 = not relevant 2 = item needs revision 3 = relevant but needs revision 4 = very relevant	1 = not relevant 2 = item needs revision 3 = relevant but needs revision 4 = very relevant	1 = not relevant 2 = item needs revision 3 = relevant but needs revision 4 = very relevant	1 = ambiguous 2 = item needs revision 3 = unambiguous but needs revision 4 = unambiguous

Instrument

A questionnaire was constructed based on Lynn's recommendations (Lynn, 1986) consisting of four criteria: ambiguity, simplicity, clarity, and relevance. Each criterion was rated using the options provided in Table 2, for both each item and the instruction table detailed in Table 3. In addition, the experts were asked whether they thought any modification to any items or instructions was necessary. Data collection was carried out using an anonymous survey designed ad hoc by the study researchers on the LimeSurvey® platform (Carsten Schmitz, Jason Cleeland). The survey, consisting of four questions for each item of the scale with four possible answers for each, was sent for completion via e-mail to the group of pain and neonatal care experts.

The 10 experts were sent an e-mail inviting them to participate in the study and describing the data collection procedure, as well as explaining how the LimeSurvey® platform worked (Fig. 1).

Please cite this article as: I. Núñez-López, M.-G. Cid-Expósito, R. Abalo et al., Content Validity of the Spanish Adaptation of the Premature Infant Pain Profile Revised, Pain Management Nursing, https://doi.org/10.1016/j.pmn.2023.06.012

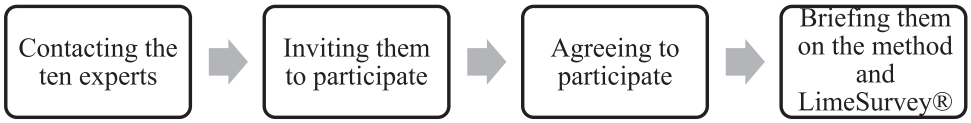

Figure 1. Flow chart of the instrument.

Ethical Considerations

This study is part of the doctoral thesis 'Adaptación cultural y validación de la escala de medición del dolor en neonatos Premature Infant Pain Profile-Revised (PIPP-R)' [Cultural adaptation and validation of the Premature Infant Pain Profile-Revised (PIPP-R) pain measurement scale in neonates'], which was approved by the Research Ethics Committee for the 12 de Octubre University Hospital (19/271) and by the Research Ethics Committee for the Rey Juan Carlos University in Madrid (2406201911219). The research protocol, which includes the overall analysis design for other psychometric properties (reliability, construct validity, inter-observer validity, intra-observer validity, and feasibility), was published in 2022 (Núñez-López et al., 2022). The data collection phase of the project is scheduled to be completed by December 2023.

Procedure

The experts were briefed and asked for their consent at the beginning of the first round of the Delphi questionnaire. They were provided with the PIPP-R scale and the instructions for use that resulted from the translation and back-translation phase of the study, which they were asked to evaluate.

The questionnaire contained a free-text response field so that experts could include any comments they deemed appropriate regarding modification of the item in question or relating to that item's responses. These comments were analyzed by the research team using content analysis and their interrelationships with the content validity index of the item in question. All of these comments were addressed and fed back to the experts in a post-round report. Some of the comments discussed confusion when measuring certain items—with suggestions being that the table of instructions be placed before the scale—while other comments mentioned data collection difficulties. These comments were taken into consideration for the planned training in subsequent phases of the study, where clinical validation of the questionnaire will be carried out.

The experts were given a period of one month to complete each round. An initial analysis of their responses was performed by the research team using a SPSS (version 25) database (IBM, Chicago, IL, USA). Modifications were made as appropriate.

Subsequently, the modified questionnaire was sent out once again alongside the results of the first evaluation round and the conclusions of all experts. Once the responses from this second round had been received, a second analysis was performed using the SPSS database and a final report was prepared and sent to all participating experts.

Assessment rounds were carried out until a consensus was reached by the experts. For consensus, an item-content validity index (I-CVI) greater than 0.8 for all items on the scale and a scale-content validity index (S-CVI) greater than 0.8 were required, as reported in the literature on the topic (Lynn, 1986; Polit & Beck, 2006; Yaghmaie, 2003).

The I-CVI is the number of experts who voted 'Relevant but needs revision' and 'Very relevant' divided by the total number of experts (n = 10). The S-CVI is the result of the sum of all I-CVIs divided by the number of items: n = 28 in the case of the scale and n = 16 in the case of the table of instructions for use.

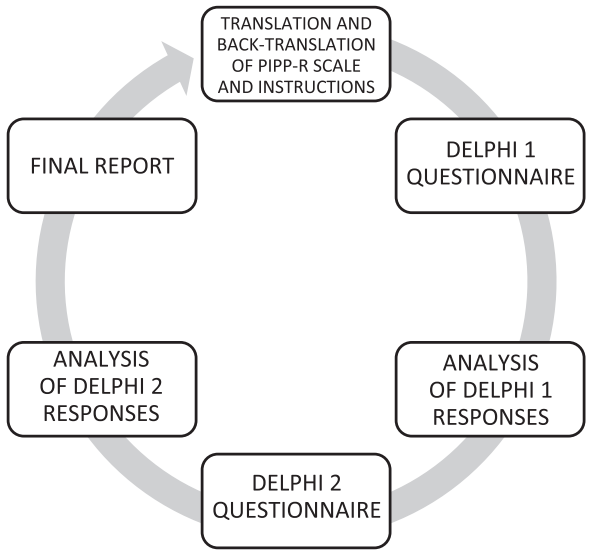

Figure 2. Flow chart of the procedure.

As proposed by Polit (2007), the kappa index was calculated to assess the proportion of agreements that were obtained beyond random chance, relative to the maximum proportion of agreements beyond the expected random chance: $k^* = (\text{I-CVI} - p_c) / (1 - p_c)$. The formula $p_c = 0.5^N$ was used to calculate the probability of chance universal agreement. k^* = modified kappa index according to Polit (2007). p_c = probability of chance universal agreement.

N = total number of experts participating in the consensus group.

The k^* values described by Polit (Cicchetti & Sparrow, 1981; Fleis, 1981; Polit & Beck, 2006) indicate a degree of agreement beyond random chance, which is excellent when k^* scores are above 0.74, good between 0.6 and 0.74, and acceptable between 0.40 and 0.59. Figure 2 illustrates the procedure described:

Results

The 10 experts selected included five nurses, three physicians, one nursing assistant, and one psychologist. The scores obtained for the number of experts who met each criterion are shown in Table 4.

After receiving the results of the first round, the results of the I-CVI and S-CVI for both the PIPP-R pain assessment scale (Table 5) and its instructions for use (Table 6) were analyzed.

In round 1, several items did not exceed an I-CVI of 0.8, which is considered to be acceptable according to the literature consulted (Lynn, 1986; Polit & Beck, 2006; Yaghmaie, 2003). In the case of the I-CVIs for the instructions for use, step 3 obtained the poorest I-CVI, and none of the ratings were above 0.8. In line with the Delphi technique involving expert opinion on each item, the research

Table 4
Number of Experts Meeting Each Selection Criterion.

Criterion	Score	N - Expert
Doctorate	4 points	6
Doctoral thesis on neonatal pain	1 point	2
Clinical experience with neonates	1 point per year	7
Research projects in neonatal pain	1 point	6
Publications in neonatal pain	1 point	4
Specific training in neonatal pain	2 points	8
Participation in a working group on neonatal pain	1 point	8
Delivery of specific training in neonatal pain	2 points	4

Table 5
Results Obtained from the Survey Conducted Using the LimeSurvey® Platform.

ROUND 1 results: I-CVI for the PIPP-R pain assessment scale					
	Not relevant	Needs revision	Relevant but needs revision	Very relevant	I-CVI
ITEM 1: Changes in heart rate					
• Relevance	1	1	1	7	0.8
• Clarity	1	0	1	8	0.9
• Simplicity	0	0	0	10	1
• Ambiguity	1	0	1	8	0.9
ITEM 2: Decrease in oxygen saturation					
• Relevance	0	1	3	6	0.9
• Clarity	0	1	5	4	0.9
• Simplicity	0	1	4	5	0.9
• Ambiguity	1	2	3	4	0.7
ITEM 3: Brow bulge					
• Relevance	0	1	3	6	0.9
• Clarity	0	1	3	6	0.9
• Simplicity	0	0	4	5	1
• Ambiguity	0	2	3	5	0.8
ITEM 4: Eye squeeze					
• Relevance	0	1	3	6	0.9
• Clarity	1	1	3	5	0.8
• Simplicity	1	0	5	4	0.9
• Ambiguity	1	1	4	4	0.8
ITEM 5: Naso-labial furrow					
• Relevance	2	2	2	4	0.6
• Clarity	1	3	4	2	0.6
• Simplicity	1	2	4	3	0.7
• Ambiguity:	1	2	5	2	0.7
ITEM 6: Corrected gestational age					
• Relevance	0	1	1	8	0.9
• Clarity	0	3	0	7	0.7
• Simplicity	0	2	1	7	0.8
• Ambiguity	0	2	1	7	0.8
ITEM 7: Baseline behavior					
• Relevance	0	1	4	5	0.9
• Clarity	0	3	4	3	0.7
• Simplicity	1	2	4	3	0.7
• Ambiguity	1	2	5	2	0.7

I-CVI = item-content validity index; PIPP-R = Premature Infant Pain Profile-Revised.

team made the following modifications: the display of the 'weeks and days' score in the gestational age item was changed for better visual understanding; a clarification of 'baseline behavior' was added to the instructions for use table to assist in its assessment; the acronyms in the instructions for use table were replaced with their full descriptions for clarity; step 2 of the instructions for use table was reworded for better understanding of the duration of the assessment; steps 3 and 4 in the instructions for use table were restructured for clarity.

The mean I-CVI values for calculating the S-CVI in round 1 are 0.81 (±0.07) for the scale and 0.8 (±0.14) for the instruction table. Values above 0.8 are considered acceptable. Therefore, after implementing the suggested changes, a second round was carried out. The results of the second-round analysis are shown in Tables 7 and 8.

After round 2, all items exceeded an I-CVI of 0.9, with most items scoring 1, which meant full consensus among all experts on both the scale and the instruction table of the PIPP-R scale. As a result, no further rounds were necessary.

The mean I-CVI values for calculating the S-CVIs in round 2 are 0.98 (±0.03) for the PIPP-R scale and 1 for the instruction table. In addition, the kappa index yielded values (Tables 7 and 8) indicating an excellent degree of agreement, beyond random chance.

Discussion

In this study, the Delphi method was used to analyze the content validity of the PIPP-R neonatal pain assessment scale and its instructions for use, after translation and back-translation, for adaptation to the Spanish cultural context (Núñez-López et al., 2022).

The content validity of the Spanish PIPP-R pain assessment scale and its instruction table was confirmed by expert consensus using the Delphi method over two survey rounds.

Please cite this article as: I. Núñez-López, M.-G. Cid-Expósito, R. Abalo et al., Content Validity of the Spanish Adaptation of the Premature Infant Pain Profile Revised, Pain Management Nursing, https://doi.org/10.1016/j.pmn.2023.06.012

Table 6
Results obtained from the survey conducted using the LimeSurvey® platform.

ROUND 1 results: I-CVI for the PIPP-R instructions for use					
	Not relevant	Needs revision	Relevant but needs revision	Very relevant	I-CVI
STEP 1					
• Relevance	0	0	1	9	1
• Clarity	0	0	2	8	1
• Simplicity	0	0	2	8	1
• Ambiguity	0	0	0	10	1
STEP 2					
• Relevance	0	1	2	7	0.9
• Clarity	0	2	2	6	0.8
• Simplicity	0	0	4	6	1
• Ambiguity	0	1	3	6	0.9
STEP 3					
• Relevance	1	1	2	6	0.8
• Clarity	3	2	1	4	0.5
• Simplicity	1	3	2	4	0.6
• Ambiguity	2	2	2	4	0.6
STEP 4					
• Relevance	0	2	1	7	0.8
• Clarity	1	3	0	6	0.6
• Simplicity	0	3	1	6	0.7
• Ambiguity	1	3	0	6	0.6

I-CVI = item-content validity index; PIPP-R = Premature Infant Pain Profile-Revised.

Table 7
Results obtained from the survey conducted using the LimeSurvey® platform.

ROUND 2 results: I-CVI for the PIPP-R pain assessment scale.						
	Not relevant	Needs revision	Relevant but needs revision	Very relevant	I-CVI	k*
ITEM 1: Changes in heart rate						
• Relevance	0	0	0	10	1	1
• Clarity	0	0	1	9	1	1
• Simplicity	0	0	2	8	1	1
• Ambiguity	0	0	2	8	1	1
ITEM 2: Decrease in oxygen saturation						
• Relevance	0	0	2	8	1	1
• Clarity	0	0	3	7	1	1
• Simplicity	0	0	2	8	1	1
• Ambiguity	0	0	3	7	1	1
ITEM 3: Brow bulge						
• Relevance	1	0	2	7	0.9	0.9
• Clarity	1	0	2	7	0.9	0.9
• Simplicity	0	1	3	6	0.9	0.9
• Ambiguity	0	0	3	7	1	1
ITEM 4: Eye squeeze						
• Relevance	0	0	0	10	1	1
• Clarity	0	0	0	10	1	1
• Simplicity	0	0	0	10	1	1
• Ambiguity	0	0	0	10	1	1
ITEM 5: Naso-labial furrow						
• Relevance	0	0	1	9	1	1
• Clarity	0	0	1	0	1	1
• Simplicity	0	0	0	10	1	1
• Ambiguity	0	0	1	9	1	1
ITEM 6: Corrected gestational age						
• Relevance	0	0	1	9	1	1
• Clarity	0	0	2	8	1	1
• Simplicity	0	0	1	9	1	1
• Ambiguity	0	0	3	7	1	1
ITEM 7: Baseline behavior						
• Relevance	0	0	1	9	1	1
• Clarity	0	0	3	7	1	1
• Simplicity	0	0	2	8	1	1
• Ambiguity	0	0	2	8	1	1

I-CVI = item-content validity index; PIPP-R = Premature Infant Pain Profile-Revised.

Please cite this article as: I. Núñez-López, M.-G. Cid-Expósito, R. Abalo et al., Content Validity of the Spanish Adaptation of the Premature Infant Pain Profile Revised, Pain Management Nursing, https://doi.org/10.1016/j.pmn.2023.06.012

Table 8
Results obtained from the survey conducted using the LimeSurvey® platform.

ROUND 2 results: I-CVI for the PIPP-R instructions for use.						
	Not relevant	Needs revision	Relevant but needs revision	Very relevant	I-CVI	k*
STEP 1						
• Relevance	0	0	1	9	1	1
• Clarity	0	0	4	6	1	1
• Simplicity	0	0	2	8	1	1
• Ambiguity	0	0	2	8	1	1
STEP 2						1
• Relevance	0	0	1	9	1	1
• Clarity	0	0	5	5	1	1
• Simplicity	0	0	3	7	1	1
• Ambiguity	0	0	3	7	1	1
STEP 3						1
• Relevance	0	0	1	9	1	1
• Clarity	0	0	1	9	1	1
• Simplicity	0	0	2	8	1	1
• Ambiguity	0	0	2	8	1	1
STEP 4						1
• Relevance	0	0	2	8	1	1
• Clarity	0	0	3	7	1	1
• Simplicity	0	0	2	8	1	1
• Ambiguity	0	0	2	8	1	1

I-CVI = item-content validity index; PIPP-R = Premature Infant Pain Profile-Revised.

In the second and final round of the Delphi method, 98% of the participating experts agreed on the PIPP-R scale in Spanish and 100% of them agreed on its table of instructions for use.

In the validation study of the PIPP-R in its original language (Gibbins et al., 2014), content validity was not explored. It was also not explored in its cross-cultural adaptation to other languages such as Persian (Sadeghi et al., 2017) and four Nordic languages (Olsson et al., 2018). However, the validation and cross-cultural adaptation to Brazilian Portuguese (Bueno et al., 2019) did include this crucial validation phase in its methods and provided results similar to the ones obtained in this study: a CVI of 1 in terms of clarity and relevance in the expert consensus. Nevertheless, the authors did not specify how many rounds were necessary.

Other studies on cross-cultural validation and adaptation of scales—in which a content validity phase was carried out through expert consensus—produced similar results to ours. Examples include the Turkish validation of the Stoma Self-Efficacy Scale (Karaçay et al., 2020), which attained a CVI of 0.96 with the consensus of 10 experts but without specifying the number of rounds required, and the Spanish version of the Practice Environment Scale of the Nursing Work Index (Orts-Cortés et al., 2013), which obtained a CVI of 0.82, stating that future studies are required to improve that score.

Other cross-cultural validation studies of pain scales using a methodology similar to our study include the validation of the King's Parkinson's Disease Pain Scale in Brazilian Portuguese (Regina Coimbra et al., 2021), in which a CVI above 0.9 was attained in three rounds with the consensus of 10 experts, as well as the adaptation of the Revised Nonverbal Pain Scale to Turkish (Erden Çukurova Üniversitesi Sağlık Bilimleri Fakültesi et al., 2019), in which a CVI of 1 was obtained with the consensus of nine experts.

We also found similar methods and results when comparing our study with other validation and cultural adaptation studies of pediatric pain assessment scales, such as in the Brazilian Portuguese adaptation of the COMFORTneo scale (Menegol et al., 2022), with a CVI of 0.99 and a consensus of seven experts.

The final product obtained in the present study is a Spanish-language pain assessment scale for neonatal patients, consisting of seven items and a table of instructions for use with four explanatory steps. This final product will be subjected to a construct validity analysis to obtain a validated scale for measuring neonatal pain in the Spanish health care context.

Limitations

Limitations of the present study include the response times of the experts, which in some cases made it difficult to meet the one-month deadline. It should also be noted that several experts were health professionals specializing in pain but not in neonatal pain, resulting in a limited understanding of neonatal terminology.

Conclusions

The Spanish version of the PIPP-R neonatal pain assessment scale obtained a high degree of content validity through expert consensus using the Delphi technique. Therefore, the items that make up the scale represent the content that the scale is intended to assess, i.e., the assessment of pain in neonatal patients.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors would like to thank the experts for their participation, as well as the Hospital 12 de Octubre Health Research Institute (imas12) for their support and advice. We would also like to thank Pallás, Esther Cabañes, and Patricia Melchor for their valuable feedback.

RICORS was funded by the 2017-2020 Recovery, Transformation, and Resilience Plan, the ISCIII, and the European Union – Next Generation EU funds (ref. RD21/0012/0009).

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.pmn.2023.06.012.

References

Avila-Alvarez, A., Carbajal, R., Courtois, E., Pertega-Diaz, S., Anand, K. J. S., & Muñiz–Garcia, J. (2016). Valoración clínica del dolor en unidades de cuidados intensivos neonatales españolas. *Anales de Pediatria, 85*(4), 181–188.

Bueno, M., Moreno-Ramos, M. C., Forni, E., & Kimura, A. F. (2019). Adaptation and initial validation of the Premature Infant Pain Profile–Revised (PIPP-R) in Brazil. *Pain Management Nursing, 20*(5), 512–515.

Castillo Barrio, B., Rasines Rodríguez, A., Aneiros Suanzes, C., Royuela Vicente, A., Ortiz Movilla, R., & Marín Gabriel, M. A. (2020). A survey of procedural pain assessment and non-pharmacologic analgesic interventions in neonates in Spanish public maternity units. *Journal of Perinatology, 40*(12), 1764–1769.

Cicchetti, D. V., & Sparrow, S. A. (1981). Developing criteria for establishing interrater reliability of specific items: applications to assessment of adaptive behavior. *American Journal of Mental Deficiency, 86*(2), 127–137.

Deville, C. W., & Prometric, S. (1996). An empirical link of content and construct validity evidence. *Appl Psychol Meas, 20*(2), 127–139.

Dorans, N. J., & Lawrence, I. M. (1987). The internal construct validity of the SAT1. *ETS Research Report Series, 1987*(2) i-101.

Erden Çukurova Üniversitesi Sağlık Bilimleri Fakültesi, S., Bölümü, H., Kaya, P., & Erden, S. (2019). PAIN A RI Cross-cultural adaptation, validity and reliability of the Turkish version of Revised Nonverbal Pain Scale Revize Sözel Olmayan Ağrı Ölçeği'nin Türkçe versiyonunun kültürlerarası uyarlaması, geçerliği ve güvenirliği. *Agri, 31*(1), 15–22.

Fleiss, J. (1981). *Statistical methods for rates and proportions* (2nd ed.). New York: John Wiley.

Gibbins, S., Stevens, B. J., Yamada, J., Dionne, K., Campbell-Yeo, M., Lee, G., Caddell, K., Johnston, C., & Taddio, A. (2014). Validation of the Premature Infant Pain Profile-Revised (PIPP-R). *Early Human Development, 90*(4), 189–193.

Giordano, V., Edobor, J., Deindl, P., Wildner, B., Goeral, K., Steinbauer, P., Werther, T., Berger, A., & Olischar, M. (2019). Pain and sedation scales for neonatal and pediatric patients in a preverbal stage of development: A systematic review. *JAMA Pediatrics, 173*(12), 1186–1197.

Hyrkäs, K., Appelqvist-Schmidlechner, K., & Oksa, L. (2003). Validating an instrument for clinical supervision using an expert panel. *International Journal of Nursing Studies, 40*(6), 619–625.

Ilhan, E., Pacey, V., Brown, L., Spence, K., van Ganzewinkel, C.-J., Riddell, R. P., Campbell-Yeo, M., Stevens, B. J., Eriksson, M., Shah, V., Anand, K. J. S., Bellieni, C., Daly, M., Johnston, C., & Hush, J. (2022). What is the definition of acute episodic and chronic pain in critically ill neonates and infants? A global, four-stage consensus and validation study. *BMJ Open, 12*, 55255.

Karaçay, P., To Gluk Yigitoglu, E., Ayis¸, E., & Karada, A. (2020). The validity and reliability of the Stoma Self-Efficacy Scale: A methodological study. *International Journal of Nursing Practice, 26*(6).

Linstone, H. A., & Turrof, M. (1975). The Delphi method: Techniques and applications. *Journal of Marketing Research, 18*(3).

Lynn, M. R. (1986). Determination and quantification of content validity. *Nursing Research, 35*(6), 382–386.

Menegol, N. A., Ribeiro, S. N. S., de Paula, A. C., Montemezzo, D., & Sanada, L. S. (2022). A cross-cultural adaptation and content validity of COMFORTneo scale into Brazilian Portuguese. *Journal of Pain and Symptom Management, 64*(6), e323–e330.

Mokkink, L. B., Terwee, C. B., Patrick, D. L., Alonso, J., Stratford, P. W., Knol, D. L., Bouter, L. M., & de Vet, H. C. W. (2010). The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *Journal of Clinical Epidemiology, 63*(7), 737–745.

Núñez-López, I., Collados-Gómez, L., Abalo, R., Martínez-Pérez, P., Moreno-Vicente, Á., & Cid-Expósito, M. G. (2022). Cultural adaptation and validation of the Premature Infant Pain Profile-Revised (PIPP-R) pain measurement scale: Research protocol. *International Journal of Environmental Research and Public Health, 19*(19), 12338.

Olsson, E., Anderzén-Carlsson, A., Atladóttir, S. M., Axelin, A., Campbell-Yeo, M., Eriksson, M., Kristjánsdóttir, G., Peltonen, E., Stevens, B., Vederhus, B., & Andersen, R. D. (2018). Cultural adaptation and harmonization of four Nordic translations of the revised Premature Infant Pain Profile (PIPP-R). *BMC Pediatrics, 18*(1), 349.

Orts-Cortés, M. I., Moreno-Casbas, T., Squires, A., Fuentelsaz-Gallego, C., Maciá–Soler, L., González-María, E., Gómez-García, T., María González-Martín, A., Gómez-González, B., Quesada-Ramos, C., Candel-Parra, E., Oltra-Rodríguez, E., Terol-García, E., Abad-Corpa, E., Guerra-Bernal, F. J., Lozano-Olea, F., Rodríguez-Escobar, J., Martínez-Gimeno, L., Ferrús-Estopa, L., … Sebastián–Viana, T. (2013). Content validity of the Spanish version of the Practice Environment Scale of the Nursing Work Index. *Applied Nursing Research, 26*(4), e5–e9.

Polit, D. F., & Beck, C. T. (2006). The content validity index: Are you sure you know what's being reported? Critique and recommendations. *Research in Nursing & Health, 29*(5), 489–497.

Polit, D. F., Beck, C. T., & Owen, S. v. (2007). Is the CVI an acceptable indicator of content validity? Appraisal and recommendations. *Research in Nursing & Health, 30*(4), 459–467.

Quatrini Carvalho Passos Guimarães, H. C., Pena, S. B., Lopes, J. de L., Lopes, C. T., & Bottura Leite de Barros, A. L (2016). Experts for validation studies in nursing: New proposal and selection criteria. *International Journal of Nursing Knowledge, 27*(3), 130–135.

Regina Coimbra, M., Megale Almeida-Leite, C., de Faria-Fortini, I., Pereira Christo, P., & Luciana Scalzo, P. (2021). King's Parkinson's Disease Pain Scale (KPPS): Cross–cultural adaptation to Brazilian Portuguese and content validity. *Clinical Neurology and Neurosurgery, 208*, Article 106815.

Sadeghi, A., Rassouli, M., Abolhasan Gharehdaghi, F., Nourian, M., Nasiri, M., Farhadi, R., & Shirinabadi Farahani, A. (2017). Validation of the Persian version of premature infant pain profile-revised in hospitalized infants at the neonatal intensive care units. *Iranian Journal of Pediatrics, 27*(5), e10056.

Stevens, B., Johnston, C., Petryshen, P., & Taddio, A. (1996). Premature Infant Pain Profile: Development and initial validation. *Clinical Journal of Pain, 12*(1), 13–22.

Waltz, C. F., Strikland, O. L., & Lenz, E. R. (2010). *Measurement in nursing and health research* (4th ed.). New York: Springer.

Yaghmaie, F. (2003). Content validity and its estimation. *Journal of Medical Education Spring, 30*(1), 25–27.

Please cite this article as: I. Núñez-López, M.-G. Cid-Expósito, R. Abalo et al., Content Validity of the Spanish Adaptation of the Premature Infant Pain Profile Revised, Pain Management Nursing, https://doi.org/10.1016/j.pmn.2023.06.012