MASTER THESIS

Evaluating content validity of the Distress Thermometer, for Parents of preterm infants.

Patricia van Gestel - 5874874

Research Internship 2: Master Thesis – Final Version

June 25th, 2021

Utrecht University - Master Clinical Health Science - Nursing Science - UMC Utrecht

Dr. J.M. de Man van Ginkel – Lecturer

Dr. A. van den Hoogen - Supervisor

University Medical Centre Utrecht - Division woman & child

Acta Pediatrica

Words total: 3575

Words abstract: 300

Words Dutch Summary: 294

Abstract

Background: Unexpected preterm delivery is a traumatic event, which is associated with parental fear and known to cause significant stress. Parental stress hampers the bonding process between infant and parents and early detection is necessary to minimise disruption of this process. The Distress Thermometer for Parents (DT-P) is a valid measurement instrument that measures parental stress for parents with a chronically ill child. Validating the DT-P for parents of preterm infants could help healthcare professionals' recognise symptoms of distress in parents at an early stage and alert them to parents' need for expert help.

Aim: To evaluate content validity of the DT-P in the context of preterm infants, according to their parents

Method: This content validity study has a cross-sectional design and was conducted by using an online survey. Conform COSMIN criteria, parents rated 42 items of the DT-P on relevance, completeness, and comprehensibility with a four-point ordinal scale. Subsequently, Content Validity Index (CVI) was calculated as Scale-Content Validity Index /Average (S-CVI/Ave) and Item-Content Validity Index (I-CVI).

Results: The S-CVI/Ave on relevance was 0.57. Parents judged the DT-P as irrelevant. Four domains were assessed as important so revision is advised. Comprehensiveness (S-CVI/Ave 0.89) was acceptable and comprehensibility (S-CVI/Ave 0.95) was desirable.

Conclusion: Content validity of the DT-P of preterm infants assessed by parents can be considered insufficient. Although most items were not regarded as relevant it may be interesting to explore the four domains of the relevant items.

Implication of key findings: For clinical practice, this study marks the first step in determining whether validation is possible for this new target group. As the COSMIN criteria indicate, survey research alone is not sufficient. Additional research with in-depth qualitative research is recommended to enrich the collected data.

Keywords: preterm birth, parental stress/ distress, distress thermometer (for parents), content validity, COSMIN

Samenvatting

Achtergrond: Een onverwachte vroeggeboorte is een traumatische gebeurtenis die aanzienlijke stress kan veroorzaken bij ouders. Ouderlijke stress belemmert het hechtingsproces tussen pasgeborene en ouders. Vroegtijdige opsporing van deze ouderlijke stress is noodzakelijk om verstoring van het hechtingsproces tot het minimum te beperken. De Last Thermometer voor Ouders (LTO) is ontwikkeld voor ouders met een chronisch ziek kind. Als deze LTO gevalideerd zou kunnen worden voor ouders van premature pasgeborene, kunnen zorgverleners symptomen van stress bij ouders in een vroeg stadium herkennen en zo ouders vroegtijdig doorverwijzen naar deskundige hulp.

Doel: Het doel is om de content validiteit te beoordelen van de LTO in de context van premature pasgeborene, door hun ouders.

Methode: Dit onderzoek had een cross-sectioneel design en is uitgevoerd middels een online-enquête. Zoals de COSMIN-criteria aanbeveelt, beoordeelden ouders alle items (tweeënveertig) van de LTO op relevantie, volledigheid en begrijpelijkheid op een vier-punts schaal. Daarna werd de content validiteitsindex berekend als Schaal-content validiteitsindex /Average (S-CVI/Ave) en Item-content validiteitsindex (I-CVI).

Resultaten: De S-CVI/Ave op relevantie was 0,57. Ouders beoordeelden de LTO in het algemeen als niet relevant. Toch lijken vier domeinen belangrijk voor hen, mogelijke aanpassing word geadviseerd. De volledigheid (S-CVI/Ave was 0,89) was acceptabel en begrijpelijkheid (S-CVI/Ave was 0,95) wenselijk.

Conclusie: De content validiteit van de LTO van premature kan als onvoldoende worden beschouwd. Hoewel de meeste items niet relevant werden geacht, kan het toch interessant zijn om de vier domeinen waar in de relevante items zich bevinden verder te onderzoeken.

Aanbevelingen: Voor de klinische praktijk betekent dit onderzoek de eerste stap om te kijken of validatie mogelijk is voor deze nieuwe doelgroep. Zoals de COSMIN-criteria aangeven, is aanvullend kwalitatief onderzoeken nodig om de verzamelde gegevens te verrijken met aanvullende informatie.

Sleutelwoorden: vroeggeboorte, ouderlijke stress, last thermometer (voor ouders), content validiteit, COSMIN

Introduction

Worldwide, one in ten babies (10%) is born prematurely[1], defined as birth at a gestational age (GA) from 22 to 37 weeks[2]. In 2018, 162,464 children were born in the Netherlands, of whom 11,535 (7,1%) were born prematurely[3]. Of these preterm infants, 2,275 (1,4%) were born before a GA of 32 weeks[3]. Infants born before 32 weeks' GA are often admitted to a neonatal intensive care unit (NICU)[4].

Giving birth to a healthy infant and the transition to parenthood are major changes in the life of a family[5]. Unexpected preterm delivery and admission of an infant to the NICU is a traumatic event associated with parental fear and known to cause significant stress[5–8]. Seeing the infant at the NICU connected to tubes and surrounded by technological equipment with various unfamiliar alarms, early separation after the delivery, and not being able to touch your infant, are experienced as very stressful[7–9]. In addition to their infants premature medical condition, parents cite the loss of their desired parenting role as the greatest source of stress[10,11]. Other resulting trauma-related symptoms reported by parents are hyperarousal, flashbacks to their baby's admission, and avoiding the NICU[7,10]. Parental stress hampers the bonding process between infant and parents and is associated with a higher risk of acute stress disorder (ASD)[10]. ASD is described as a trauma- and stressor-related disorder during the first month after a potentially traumatic event[7]. In the long term, ASD can develop into posttraumatic stress disorder (PTSD)[6,7]. PTSD can negatively affect the bonding process between parent and infant, the long-term mental health of parents, and the development of the preterm infant[7,8].

Early detection of parental stress is necessary to minimise disruption of the bonding process. An instrument to measure parental distress would be useful to identify which parents need additional support after admission[12]. Various instruments already exist to measure parental stress. For example the Parenting Stress Index - Short Form (PSI-SF)[13] and Parental Stress Scale - Neonatal Intensive Care Unit (PSS-NICU)[9,13] also measures the infants and their environment[9,13]. The PSI-SF and the Nijmeegse Ouderlijke Stress Index – K (NOSI-K) measures stress in parents of children older than age two years[14]. However, these instruments do not focus solely no measuring stress of parents who have preterm infants.

A widely used and implemented instrument to detect stress in parents is the Distress Thermometer (DT)[15–17]. The DT was originally developed in standard adult oncology[15]. The DT has been translated, adapted, and validated[18] for other populations in the Netherlands. Haverman et al. (2013) developed the Distress Thermometer for Parents (DT-P) by adapting the DT for distress in parents of children with a chronic condition[17,19]. During its development, the PSI-SF and the Hospital Anxiety and Depression Scale (HADS) have been used for the validation of the DT-P[19]. Validation was conducted by comparing the thermometer score

rating with the results of the HADS score. The DT-P has six subdomains (practical, family/ social, emotional, physical, cognitive, and parenting) for parents and indicates when it is necessary to refer them to a health care professional which is most appropriate. The DT-P is a well-validated instrument for identifying distress and everyday problems in parents of children with a chronic condition in the Netherlands[17,19,20]. A screening instrument to detect distress specifically in parents of a preterm infant does not yet exist[19].

To adapt the DT-P to the population of parents of preterm infants with a GA from 22 to 37 weeks, the psychometric characteristics must be examined again[21]. A content validity study is recommended to assess this and is described in the COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN) criteria[22,23]. Content validity is described as the degree to which the content of a screening instrument is an adequate reflection of the construct to be measured[24,25] based on 1) relevance (all items should be relevant for the construct of interest within a specific population and context of use); 2) comprehensiveness (no aspects of the construct should be missing); and 3) comprehensibility (the items should be understood as intended)[22]. Content validity of an already existing screening instrument should be assessed by asking patients about relevance, comprehensiveness, and comprehensibility of the items[22,25]. By validating the DT-P for parents of preterm infants, healthcare professionals will be able to recognise symptoms of distress in parents at an early stage, and customised assistance can be offered.

Aim

The study aimed to evaluate content validity of the DT-P, in terms of relevance, comprehensiveness, and comprehensibility, in a population of parents of preterm infants after admission to the NICU.

Method

Design

A cross-sectional design was used. Quantitative data was generated through an online survey. This described a specific health-related event (premature birth) in a specific population (parents with a preterm infant) at a specific time (after visiting the follow-up clinic), without follow-up[26].

Population and domain

Participants were recruited in a university hospital in the Netherlands.

The study's population consisted of parents of preterm infants after their child's admission to the NICU; they were considered experts in the phenomenon being studied[22,27]. Participants were eligible if they could speak, read and understand the Dutch language, were a parent of a preterm infant born before 37 weeks of gestation, and their child had been admitted to the NICU for at least a week[13] and attended a follow-up clinic.

Procedure

Potential participants were invited to participate from February to May 2021.

Ten pilot surveys were sent in February to adjust for any errors and ambiguities.

To ensure that all relevant experiences were captured, participants were recruited through purposive sampling[22].

After discharge from the NICU, parents and their infant visited the follow-up clinic in the children's clinic. Parents filled in the DT-P at the start of their appointment. A social worker checked whether the parents met the inclusion criteria. Next, parents were informed about the study. If the were interested, parents consented to provide their e-mail address to the researcher.

Parents were approached by an e-mail requesting participation in the study. This email contained a link to participate. When they activated the link, the research environment created for them in Castor Electronic Data Capture (EDC) opened and the following fields were shown: 1) patient information letter (PIF) with instructions on completing the survey, as well as the e-mail address and telephone number of the researcher in case there were questions; 2) informed consent (IC); 3) demographic data and the survey. If the parents had not started or completed the procedure within two weeks, they received a reminder by email. All research activities took place in Castor (EDC). Figure 1 shown the flowcharts of the recruitment process

Insert Figure 1

Data collection

Data in this survey research was anonymously collected through Castor (EDC).

To guarantee the quality of the data, the use of a validated data management system was mandatory. University Medical Centre (UMC) Utrecht makes Castor (EDC) available for initiated studies.

Castor (EDC) guarantees the quality of the data and allows researchers to fill the file themselves with the necessary forms. This meant that the researcher built the electronic Case Report Form (eCRF) herself. If the participants completed the procedure, the case was

'locked'. At the end of the study, the investigator selected all 'locked' cases and exported them to Statistical Package for Social Sciences (SPSS) for data analysis. The anonymized data was exported to a secure disk within the UMC Utrecht, division woman and baby (NEO_21-XXX-DtP).

DT-P

As mentioned, the DT-P is a well-validated instrument for identifying distress and everyday problems in parents of children with a chronic condition in the Netherlands[17] [19] [20]. The DT-P consists of three parts: 1) A thermometer ranging from zero (no distress) to 10 (extreme distress) on which parents rate their overall distress of the past week. A score of four or higher indicates clinically increased distress. 2) A problem list that examines 37 everyday problems from the past week, divided into six different problem domains (practical, social, emotional, physical, cognitive, and parenting). Parents can indicate whether they experience this as problematic by choosing 'yes' or 'no' on the problem list. 3) Additional questions regarding; perceived support from the environment, perceived lack of understanding from the environment, parental chronic illness, and whether or not the parent would like to talk to a professional about his or her situation. In daily practice, this means that the validated DT-P of a chronically ill child is already used at an early stage to detect parental stress and to be able to refer parents to expert help[28].

Data analysis

Data analyses were performed in IBM SPSS version 27. When an item was not rated, the Item-Content Validity Index (I-CVI) was calculated by dividing the number of experts with a score of three or higher by the total number of experts who rated that item.

Descriptive statistics (count, mean, and standard deviation) were performed to analyse the demographic characteristics for describing the baseline variables to evaluate the group of interest. The baseline characteristics are as follows;

Infant: gender, age (in weeks), weight, length, type of birth, admission days at the NICU.

Parent: gender, age, marital status, cultural background, educational level, gravidity/ parity, gestational age, causes of premature labour, earlier premature births.

Descriptive statistics (CVI) were performed to analyse the psychometric characteristics of the DT-P.

The COSMIN checklist was applied for structuring this process[23]. In this study, the standards for evaluating the quality of the studies on the content validity of Patient-Reported Outcome Measures (PROMs) parts 2a to 2c were used[22]. Following these parts, experts (parents of preterm infants) were asked about the relevance, comprehensiveness, and comprehensibility of all the DT-P items.

The CVI has been used to calculate the content validity[27,29,30] . CVI summarises the extent to which a panel of experts (parents with preterm infants) agrees on the content validity of an instrument[31]. Experts were asked to evaluate *individual items* as well as the *overall scale* of the instrument[32]. CVI data were analysed as I-CVI and Scale-Content Validity Index/ Average (S-CVI/Ave). This means that for each item, the I-CVI was calculated by dividing the number of experts with a score of three or higher by the total number of experts[33], and for the total scale, the S-CVI/Ave was calculated by dividing the sum of all I-CVIs by the total of items[33]

All items were judged separately on a four-point ordinal scale (1: not relevant; 2: somewhat relevant; 3: quite relevant; 4: highly relevant)[34]. The scale was dichotomised by combining values one and two (not relevant) and three and four (relevant), for calculation relevance, comprehensiveness, and comprehensibility of the DT-P[30]. It should be noted that in the explanation relevance is used, but this procedure also used comprehensiveness and comprehensibility. An I-CVI > 0.78 is considered evidence of a good content validity[27,34], an I-CVI > 0.70 and < 0.78 is considered evidence for revision[27], and an I-CVI < 0.70 is considered evidence for elimination[27]. An S-CVI/Ave > 0.90 is considered desirable[27].

Insert Table 1

Ethical issues

This study has been conducted in accordance with the principles of the 64th Declaration of Helsinki[35] and the law of General Data Protection Regulation (GDPR)[36]. According to the design of this research, the Medical Research Involving Human Subjects(MRIHS)[37,38] does not apply. The Medical Research Ethics Committee (MREC) of the UMC in Utrecht gave approval for conducting this study (21-108/C).

Participants gave informed consent, which was the first question in the survey, before continuing the survey.

Results

Participants

A total of 74 parents were included in this study, of whom thirty-one (42%) responded to the completed survey. Two parents (2%) filled in demographics, but only answered the items for relevance. These studies were included in the analysis because calculations were made by dividing the number of experts with a score of three or higher by the total number of experts who rated that item, which makes the answers valid. Five (7%) parents filled in only the demographic data, and eight (11%) parents filled in the demographic data and started the survey but stopped after one or two grids. These surveys were not included in the analysis. Twenty-eight (38%) surveys were not returned.

According to the COSMIN criteria[21–23], participation between 30 and 50 experts provides an outcome that is representative for the target population within survey research. The findings are classified as adequate as thirty-three surveys were analysed in this study.

The respondents were mainly mothers (91%) whose infants had a GA of around 34 weeks and were admitted to the NICU for an average of 18 days. Infants were divided equally according to gender (boys 49%, girls 51%). Most of the participants were from the Netherlands (97%), and there were no single parents (see Table 2).

Insert Table 2

Content validity results

All content validity (I-CVI and S-CVI/Ave) calculations were from the DT-P (eight domains, 42 items/questions). Parents answered all questions three times on relevance, comprehensiveness, and comprehensibility.

The cut-off point for I-CVI > 0.78, and for S-CVI/Ave > 0.90

Relevance

The S-CVI/Ave was calculated at 0.57.

The I-CVI ranged from 0.26 to 0.88. Four (10%) questions (items 12, 13, 36, and 41) were relevant (I-CVI ≥ 0.78), two (5%) questions (items 19 and 32) need revision (I-CVI 0.70 - 0.78), and 36 (85%) questions should be eliminated (I-CVI < 0.70) from the survey. The relevance of CVI calculations of the DT-P is described in Table 3.

Results translated to the domains (practical-, social-, emotional-, cognitive-, physical-, parenting problems, enough support, and others) , none domain was basically relevant. Only one domain categorised as 'others', on the item 'dealing with staff(41)' could be relevant after revision. If both outcomes, item and domain are next to each other, possible relevance can be seen in the following domains: emotional problems, in the items 'emotions(12)', 'self-confidence(13)', 'recurring events(19)', parenting problems in the item 'child development(32)', and enough support in the item 'environment support(36)'.

The relevant CVI calculations of the DT-P domains are described in Table 4.

Insert Table 3 and 4

Comprehensiveness

The S-CVI/Ave was calculated at 0.89 which represents an acceptable result for comprehensiveness.

The I-CVI ranged from 0.77 to 0.97. Forty-one (98%) questions were complete (I-CVI ≥ 0.78), one (2%) questions needed revision (I-CVI) and none of the questions were incomplete (I-CVI < 0.70). The comprehensiveness CVI calculations of the DT-P are described in Table 3.

All eight (100%) domains were complete (I-CVI $\geq$ 0.78). The comprehensiveness CVI calculation of the DT-P domains is described in Table 4.

Insert Table 3 and 4

Comprehensibility

The S-CVI/Ave was calculated at 0.95 which means a desirable result for comprehensibility. The I-CVI ranged from 0.87 to 1.00. Forty-two (100%) questions are understandable (I-CVI $\geq$ 0.78), and none of the questions need revision or elimination (I-CVI < 0.78). The comprehensibility CVI calculations of the DT-P as described in Table 3.

All eight domains are understandable (I-CVI $\geq$ 0.78). The comprehensibility CVI calculation of the DT-P domains is described in Table 4.

Insert Table 3 and 4

Discussion

The study results showed that the content validity of the DT-P of preterm infants is assessed as insufficient (S-CVI/Ave 0.57), by parents. This means that most of the items (85%) were considered not relevant. Significantly, the relevant items are linked within certain domains. In addition, parents rated the questionnaire as comprehensive (S-CVI/Ave 0.89) and comprehensible (S-CVI/Ave 0.95) in assessing parental stress.

There are no other published validation studies on the adaptation of the DT-P for parents of preterm infants. This makes it difficult to compare the results of the current analysis with other studies. However, extensive use was made of comparable studies[17,19,21,27] to make certain choices within this study.

Because this is a content validity study, the COSMIN[21–23] literature were used in the development of this study, which will ultimately lead the study to a 'very good' or 'adequate' result[22,23]. The COSMIN criteria recommend asking experts, in this study parents of preterm infants, to rate all items of the questionnaire for relevance, completeness, and comprehensibility. COSMIN also recommends applying a mixed-method design for assessing comprehensiveness and comprehensibility. This research focused solely on the quantitative aspect, which is a limitation of this study. This means that the methodology of this study should be labelled with 'adequate' instead of 'very good'.

The study's aim is to validate the DT-P for parents of preterm infants. Haverman et al.[17,19] were responsible for the Dutch validation of the DT-P for a chronically ill child. Their research was included in the preparations of this study. The same type of parents participated in this research as in the research of Haverman et al., although the recruitment process was different. Most parents who participated were women of Dutch nationality with a

secondary vocational educational level who were married or living with their partner. This may have something to do with the fact that women are relatively more burdened with the primary care of the child than men are[39]. Another limitation of this study was that the survey was only available in Dutch which may cause a distorted view of the population who lived in the Netherlands, since non-Dutch speakers were unrepresented.

Finally, this study had a low response rate (44%). According to Haverman et al., this is because care for the child is the first priority for parents. This aligns with unpublished practical data from peer contact groups. Parents of preterm infants indicate that they cannot adequately participate in activities in the first year after delivery. Approaching parents for study proposals at least one year after childbirth is recommended.

Most remarkable in this study is the very low rating that parents give for relevance. This result was unexpected, because the DT-P is a well-validated instrument for measuring stress in parents.

Because their infant had been admitted to the NICU, parents may not have been focused on their own problems and, therefore did not score them in the survey. One parent sent an email about this indicating that the survey was too long, too time-consuming and that the parent was far too busy for this.

Due to the COVID pandemic, the researcher was dependent upon the healthcare worker to recruit parents for this study by asking for their email address. The researcher did not choose to approach parents by telephone; this may have influenced the instructions on how to complete the survey. The only instruction that was given took place via the patient information letter. This may have led to misinterpretation of questions.

Another reason for the low ratings may have been excessively rigid cut-off values. For this study the cut-off values described by Polit and Beck[30,32,40] have been used. However, literature shows that the number of included participants may affect the outcome of the CVI. According to Lynn (1986) having at least nine experts is adequate and Polit and Beck (2007) state that having more than six experts is sufficient[29] with a cut-off value of 0.78. According to the COSMIN criteria for this research, a minimum of 30 participants has been included. However, Lynn et al. and Lawshe[41] recommended a lower cut-off value when the number of inclusions is 30 or higher.

The length of the survey was different than the length that was calculated in the pilot. In the pilot participants needed 30 minutes to complete the survey, while participants in the study had to use 45 to 60 minutes to answer 42 questions. One parent also responded to the researcher by email and had, therefore, stopped completing the survey.

Using a four-point ordinal scale is a strength of the study. With a four-point scale, the participant cannot be neutral and must choose an answer[42]. The chosen method of calculation, dividing the number of experts with a score of three or higher by the total number

of experts who rated that item, also ensures that the outcome is reliable and that missing data does not affect it.

Data collection took place in Castor (EDC), which imposed limitations upon this research. For example, the survey was built in a grid, and if parents forgot to answer some questions, Castor indicated that the survey was complete, so that no reminder was sent to the parents. For this kind of research, another data collection base should probably be selected. For example, Select Survey (SS-net) could be an acceptable option because, it requires participants to answer a question before they can continue to the next one. For calculation of the static analysis, SPSS was used. CVI calculations within SPSS were difficult, as SPSS does not have the ability to perform CVI calculations. Therefore, for this research, a statistician recommended using manual calculations.

The results of this study have shown that a further development of the DT-P is needed. This research used a quantitative approach. It is one-sidedly illuminated, which possibly means that participants may have misinterpreted questions. A mixed-method with in-depth interviews is recommended for gaining more insight into how participants interpret the questions[43]. Additionally, qualitative research can provide rich information about scale items which are rated low on relevance (S-CVI of 0.57).

Further development of the DT-P is important so that healthcare professionals will be able to recognise symptoms of distress at an early stage and, therefore, refer parents to expert help.

Conclusions

Parents of preterm infants have assessed, the DT-P as not valid for them as a screening tool. Due to the quantitative nature of this study, there is no insight into why these parents score in this way. Based on these findings, additional qualitative research is needed to refine the results of this study for possible further development of DT-P for parents of preterm infants.

References

1. Zimmermann et all. EFCNIA A european foundation for the care of newborn infants.

2. Nederlandse vereniging voor Obstetrie en Gynaecologie (NVOG). Richtlijnen Perinataal beleid bij extreme vroeggeboorte. 2010.

3. Perined. Perinatale zorg in Nederland anno 2018. Landelijke perinatale cijfers en duiding.

4. Rijken M. Aanbeveling Landelijke Neonatale Follow-up - NICU follow-up. Nederlandse Vereninging voor Kindergeneeskunde (NVK) 2015.

5. Wraight, C.L. McCoy, J. Meadow W. Beyond stress: describing the experiences of families during neonatal intensive care. Acta Paediatr. 2015;104:1012–7.

6. Leisher BAF, Oopman CHK. Acute Stress Disorder Among Parents of Infants in the Neonatal Intensive Care Nursery. 2006;(June).

7. Vanderbilt, D Bushley, T Young, R Frank D. Acute Posttraumatic Stress Symptoms Among Urban Mothers With Newborns in the Neonatal Intensive Care Unit. J Dev Behav Pediatr. 2009;30(1):50–6.

8. Jubinville, J Newburn-Cook, C Hegadoren, K Lacaze-Masmontiel T. Symptoms of Acute Stress Disorde in Mothers of Prematur Infants. Adv Neonatal Care. 2012;12(4):246–53.

9. Caporali, Camilla Pisoni, Camilla Gasparini, Linda Ballane, Elena Zecca, Marzo Orcesi, Simona Provenzi L. A global perspective on parental stress in the neonatal intensive care unit: a meta-anlytic study. J Perinatol. 2020;40:1739–52.

10. Thaise A, Roque F, Lasiuk GC, Radünz V, Hegadoren K. Scoping Review of the Mental Health of Parents of Infants in the NICU. J Obstet Gynecol Neonatal Nurs [Internet]. 2017;46(4):576–87. Available from: http://dx.doi.org/10.1016/j.jogn.2017.02.005

11. Bouet, K Claudio, N Ramirez, V Garcia-fragoso L. Loss of parental role as a cause of stress in the intensice care unit. Bol Asoc Med P R. 2012;104(1):8–11.

12. Hatzmann J, Maurice-stam H, Heymans HSA, Grootenhuis MA. A predictive model of Health Related Quality of life of parents of chronically ill children : the importance of care-dependency of their child and their support system. BioMed Cent. 2009;9:1–9.

13. Schappin, Renske Wijnrok, Lex Uniken Venema, Monica Jongmans M. Rethinking Stress in parents of Pretem Infants: A Meta-Analysis. PLoS One. 2013;8(2).

14. Barrow SS. Validation of the Last Thermometer voor Ouders for premature infants. 2011.

15. Network NCC. Distress management clinical practice guidelines. J Natl Compr Canc Netw. 2003;1(3):344–74.

16. Bannink M. Instrument voor detecteren psychosociale zorgbehoeften. Pallium.

2010;12(3):14–5.

17. van Oers, H. A. Schepers, S. A. Grootenhuis, M.A. Haverman L. Dutch normative data and psychometric properties for the Distress Thermometer for Parents. Springer. 2017;26:177–82.

18. Tuinman, M.A. Gazendam-Donofrio, S.M. Hoekstra-Weebers JE. Screening and referral for psychosocial distress in oncologic practice. Cancer. 2008;113(4):870–8.

19. Haverman L, Oers HA Van, Limperg PF, Houtzager BA, Huisman J, Darlington A, et al. Development and Validation of the Distress Thermometer for Parents of a Chronically Ill Child. J Pediatr. 2013;163(4):1140–6.

20. Haverman L. Electronic patient and parent reported outcomes in pediatric clinical practice. 2013.

21. Prinsen CBTCAC, Westerman ACMJ, Alonso DLPJ, Bouter LM. COSMIN methodology for evaluating the content validity of patient- reported outcome measures : a Delphi study. Qual Life Res [Internet]. 2018;27(5):1159–70.

22. Chiarotto A, Alonso J, Patrick DL. COSMIN methodology for assessing the content validity of PROMs User manual. 2018.

23. Mokkink LB, Patrick DL, Alonso J, Bouter LM, Terwee CB. COSMIN Study Design checklist for Patient-reported outcome measurement instruments. 2019.

24. Mokkink, Lidwine B. Terwee, C.B. Patrick, D.L. Alonso, J. Stratford, P.W. Knol, D.l. Bouter, L.M. de Vet HC. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. J Clin Epidemiol. 2010;63(7):737–45.

25. Scholtes, V.A. Terwee, C.B. Poolman R w. What makes a measurement instrument valid and reliable? Injury. 2011;42:236–40.

26. National Council for Osteopathic Research. Quantitative Research Methods. 2014.

27. Zamanzadeh V, Ghahramanian A, Rassouli M, Abbaszadeh A, Alavi- H. Design and Implementation Content Validity Study : Development of an instrument for measuring Patient-Centered Communication. Tabriz Univ Med Sci [Internet]. 2015;4(2):165–78.

28. Heverman, L. Oers, H.A. Limperg, P.F. Hijmans, C.T. Schepers S. SNS. et al. Implementation of electronic Patient Report Outcomes in pediatric daily clinical practice. The KLIK experience. Clin Pediatr Pediatr Psychol. 2014;2(1):50–67.

29. Saiful M, Yusoff B. ABC of Content Validation and Content Validity Index Calculation. 2019;(June).

30. Polit, Denise F. Beck CT. The content validity index: Are you sure you know what's being reported? Critique and recommendations. Research in nursing & health. 2006;489–97.

31. Polit, D.F Beck CT. Nursing research. Generatign and assessing evidence for nursing

pratice. 2017. 724 p.

32. Polit, D.F Beck CT. Nursing research. Generating and Assessing Evidence for Nursing Pratice. 2017. 311 p.

33. Rodrigues, I.B. Adachi, J.D. Beattle, K.A. MacDermid JC. Development and validation of a new tool to measure the facilitators, barriers and preferences to exercise in people with osteoporosis. BMC Musculoskelet Disord. 2017;18(540).

34. Pilot, Denise F. Beck CT. Nursing Research. Generating and Assessing Evidence for Nursing Pratice. 2017. 337–339 p.

35. The World Medical Association. WMA DECLARATION OF HELSINKI – ETHICAL PRINCIPLES FOR Scienti c Requirements and Research Protocols. 2013;

36. Groenenweg, WA van de Putte E. Algemene Verordering Gegevensbescherming. Ned Tijdschr Geneeskd. 2018;43:162.

37. S M van VW en. Wet medisch-wetenschappelijk onderzoek met mensen. 2020.

38. Overheid.nl. Wet Medisch-wetenschappelijk Onderzoek met mensen. 2020.

39. Hallers-Haalboom L. Het verschil tussen vaderen en moederen. Universiteit Leiden. 2015. p. https://www.universiteitleiden.nl/nieuws/2015/10/h.

40. Polt, Denise F. Beck CTOS V. Is the CVI an acceptable indicator of content validity? Appraisal and recommendations. Research in nursing & health. 2007;459–67.

41. Lawshe CH. A Quantitative approach to content validity. Pers Psychol. 1975;28:563–75.

42. Lesley, G. Portney M. Foundations of Clinical Research. 2009. 371 p.

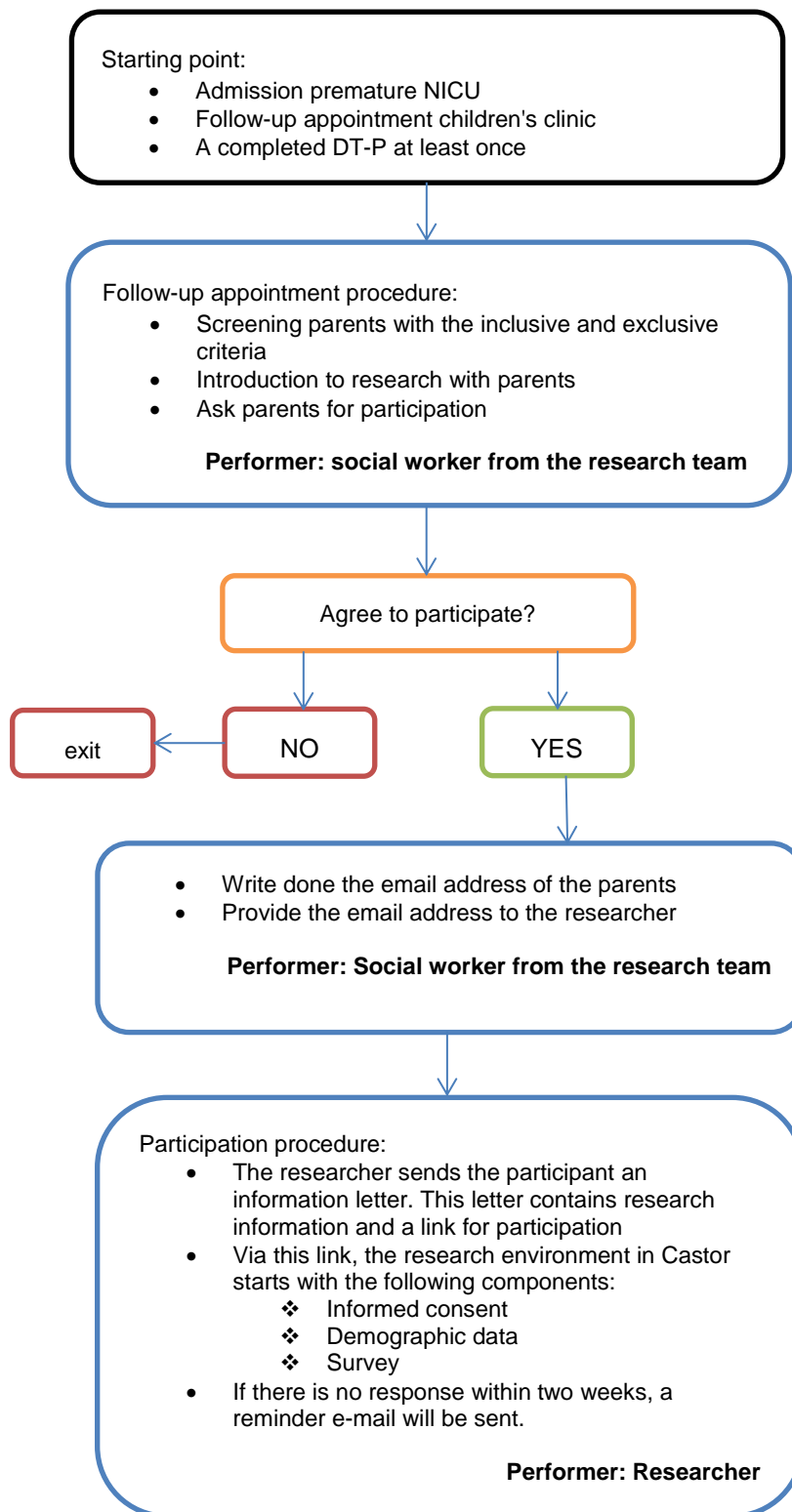43. Polit, Denise F. Beck CT. Nursing research. Generating and assessing evidence for nursing pratice. 2017. 333 p.

Figure 1. Flowcharts of recruitment process.

Table 1. Calculation CVI

I-CVI = the degree of agreement of each item, ranging from zero to one[27].

$$I\text{-}CVI = \frac{\text{Number agreeing}}{\text{Number of experts}}$$

An I-CVI > .78 is considered as evidence of a good content validity[27,34].

An I-CVI > .70 and < .78 is considered as evidence for revision[27].

An I-CVI < .70 is considered as evidence for eliminated[27]

S-CVI/Ave = the proportion of items on an instrument that has been rated three or higher by the content separators[27].

$$S\text{-}CVI/Ave = \frac{\text{Sum of proportion agreeing ratings}}{\text{Number of experts}}$$

An S-CVI/Ave > .90 is considered as desirable[27].

Table 2. Baseline characteristics

Characteristics Child	N=33
Gender, N (%)	
• boy	16 (49)
• girl	17 (51)
Age (weeks), mean (+SD)	34 (12.0)
Weight (grams), mean (+SD)	2039 (1010.0)
Length (cm), mean (+SD)	43 (6.7)
Way of giving birth, N (%)	
• normal delivery	20 (60)
• vacuum extraction	2 (6)
• forcipale extraction	0
• cesarean section	11 (34)
Admission to the NICU (days), mean (+SD)	18 (19.5)

Characteristics Parents	N=33
Gender, N (%)	
• man	3 (9)
• woman	30 (91)
Age (years), mean (+SD)	33 (5.6)
Marital status, N (%)	
• married	19 (58)
• living together	14 (42)
• single	0
• widow/ widower	0
Cultural background, N (%)	
• Dutch	32 (97)
• Not Dutch	1 (3)
Educational level, N (%)	
• Low (primary)	1 (3)
• Median (secondary)	25 (76)
• High (university)	7 (21)
Gravida, mean (+SD)	3 (2.0)
Para, mean (+SD)	2 (1.1)
Gestational age (weeks), mean (+SD)	34 (4.6)
Reason for premature birth, N (%)	
• sick mother	8 (24)
• preterm premature rupture of membranes	8 (24)
• infection	0
• fetal growth restriction	4 (12)
• different	13 (40)
Previously gave birth prematurely, N (%)	
• Yes	4 (12)
• No	29 (88)

Table 3. Calculating CVI (I-CVI & S-CVI) of DT-P of preterm infants.

Items	I-CVI* relevance	I-CVI* comprehensiveness	I-CVI* comprehensibility
Praktische problemen			
Q1 Ervaart u praktische problemen ten aanzien van wonen/ huisvesting?	0.33	0.83	0.90
Q2 Ervaart u praktische problemen ten aanzien van werk/studie?	0.46	0.83	0.97
Q3 Ervaart u praktische financiële/ verzekeringsproblemen?	0.30	0.83	0.90
Q4 Ervaart u praktische problemen met uw huishouden momenteel?	0.42	0.83	0.93
Q5 Ervaart u praktische vervoersproblemen?	0.34	0.93	0.97
Q6 Ervaart u praktische problemen ten aanzien van de zorg voor de kinderen thuis?	0.59	0.86	0.93
Q7 Ervaart u praktische problemen ten aanzien van vrije tijdsbesteding/ ontspanning?	0.39	0.89	0.93
Sociale problemen			
Q8 Ervaart u sociale problemen in omgang met uw (ex) partner?	0.58	0.87	0.90
Q9 Ervaart u sociale problemen in omgang met uw kinderen thuis?	0.64	0.90	0.90
Q10 Ervaart u sociale problemen in omgang met uw familie?	0.49	0.90	0.94
Q11 Ervaart u sociale problemen in omgang met uw vrienden?	0.34	0.90	0.94
Emotionele problemen			
Q12 Heeft u grip op uw emoties?	0.84	0.84	1.00
Q13 Heeft u zelfvertrouwen?	0.81	0.84	0.90
Q14 Heeft u last angsten?	0.64	0.80	0.97
Q15 Heeft u last van stemmingen?	0.52	0.80	0.87
Q16 Ervaart u spanningen?	0.67	0.87	0.97
Q17 Ervaart u eenzaamheid?	0.52	0.81	1.00
Q18 Ervaart u schuldgevoel?	0.61	0.77	0.94
Q19 Ervaart u terugkerende gedachten over bepaalde gebeurtenis(sen)?	0.73	0.83	0.97
Q20 Ervaart u problemen met middelen gebruik (bv alcohol, durgs en/of medicatie)?	0.50	0.96	1.00
Cognitieve problemen			
Q21 Ervaart u concentratie problemen?	0.58	0.90	1.00
Q22 Heeft u last van vergeetachtigheid?	0.55	0.87	1.00
Lichamelijke problemen			
Q23 Ervaart u lichamelijke problemen ten aanzien van eten?	0.27	0.90	1.00
Q24 Ervaart u lichamelijke problemen ten aanzien van verandering in uw gewicht?	0.33	0.90	0.97
Q25 Ervaart u lichamelijke problemen ten aanzien van slapen?	0.58	0.90	0.97
Q26 Ervaart u lichamelijke problemen ten aanzien van vermoeidheid?	0.58	0.94	0.97

Q27 Ervaart u lichamelijke problemen ten aanzien van uw conditie?	0.38	0.89	0.93
Q28 Ervaart u lichamelijke problemen ten aanzien van pijn?	0.37	0.83	0.96
Q29 Ervaart u lichamelijke problemen ten aanzien van seksualiteit?	0.26	0.87	0.91
Opvoedingsproblemen			
Q30 Maakt u zich zorgen over het contact met uw kind?	0.58	0.90	0.90
Q31 Maakt u zich zorgen over de verzorging van uw kind?	0.58	0.90	0.90
Q32 Maakt u zich zorgen over de ontwikkeling van uw kind?	0.73	0.97	1.00
Q33 Maakt u zich zorgen over het opvolgen van adviezen/ behandeling/ medicatie?	0.67	0.90	0.94
Q34 Maakt u zich zorgen over het slapen van uw kind?	0.67	0.94	1.00
Q35 Maakt u zich zorgen over het gedrag/ huilen van uw kind?	0.64	0.93	0.93
Steun van de omgeving			
Q36 Ontvangt u voldoende steun uit uw omgeving?	0.84	0.94	0.97
Q37 Is dit praktische ondersteuning?	0.63	0.88	0.87
Q38 Is dit emotionele ondersteuning?	0.63	0.88	0.90
Q39 Ervaart u onbegrip uit uw omgeving?	0.55	0.88	0.94
Overige			
Q40 Heeft u zelf een (chronische) ziekte?	0.55	0.97	0.97
Q41 Hoe was uw omgang met het medisch personeel tijdens opname?	0.88	0.88	0.90
Q42 Wenst u met een deskundige te praten?	0.67	0.94	1.00
S-CVI/ Ave*	0.57	0.89	0.95

*=Proportion of agreement

Content Validity Index (CVI) is the degree to which an instrument has an appropriate sample of items for construct being measured.
Item-level Content Validity Index (I-CVI) is the proportion of content experts giving item a rating of 3 or 4.
Cut-off point I-CVI: >0.78 evidence of a good content validity (green)
  0.70 – 0.78 evidence for revision (blue)
  <0.70 evidence for elimination (red)
Scale-level Content Validity Index/ Averaging (S-CVI/Ave) is the average of the I-CVIs for all items of the scale.
Cut-off point S-CVI: >0.90 desirable (green)
  0.80 – 0.90 acceptable (blue)

Table 4. CVI results of DT-P within the domains

I-CVI*	relevance	comprehensiveness	comprehensibility
Practical problems	0.40	0.86	0.93
Social problems	0.51	0.89	0.92
Emotional problems	0.65	0.83	0.96
Cognitive problems	0.57	0.89	1.00
Physical problems	0.40	0.89	0.96
Parenting problems	0.65	0.92	0.95
Enough support	0.66	0.90	0.92
Others	0.70	0.93	0.96
S-CVI/Ave*	0.57	0.89	0.95

*=Proportion of agreement

Cut-off point I-CVI: >0.78 evidence of a good content validity (green)
0.70 – 0.78 evidence for revision (blue)
<0.70 evidence for elimination (red)
Cut-off point S-CVI/Ave: >0.90 desirable (green)
0.80 – 0.90 acceptable (blue)