Psychometric qualities of the Dutch version of the Pediatric Inventory for Parents (PIP): a multi-center study

C. M. Jantien Vrijmoet-Wiersma1,*, Josette E. H. M. Hoekstra-Weebers2, W. M. G. Margreet de Peinder3, Hendrik M. Koopman1, Wim J. E. Tissing1, Philip D. A. Treffers4, Marc B. Bierings2, Nathalie C. A. Jansen7, Martha A. Grootenhuis8 and R. Maarten Egeler1

1Pediatric Department, Leiden University Medical Center, Leiden, The Netherlands
2Wenckebach Institute, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands
3Health Psychology, Leiden University Medical Center, Leiden, The Netherlands
4Medical Psychology, Leiden University Medical Center, Leiden, The Netherlands
5University Medical Center Groningen, University of Groningen, Groningen, The Netherlands
6Academic Center for Child and Adolescent Psychiatry Curium, Leiden University Medical Center, Leiden, The Netherlands
7Wilhelmina Children’s Hospital, University Medical Center, Utrecht, The Netherlands
8Academic Medical Center, Amsterdam, The Netherlands

Abstract

Objectives: Diagnosis and treatment of childhood cancer are continuous stressors in the lives of the entire family involved. Disease-related tools for the assessment of parental stress and adaptation are scarce. For that reason, the Pediatric Inventory for Parents (PIP), a disease-related measure, was translated into Dutch and its psychometric qualities were determined to prove its value.

Methods: The PIP and three other measures (State-Trait Anxiety Inventory, General Health Questionnaire and Parenting Stress Index, Short Form) were administered to 174 parents of 107 children diagnosed with cancer in three university medical centers in the Netherlands.

Results: Internal consistency (Crohnbach’s α = 0.94 and 0.95) and test–retest reliability (Pearson’s r between 0.67 and 0.87) of the Dutch PIP total scales are satisfactory. Validity was illustrated by a high correlation between PIP-scores and anxiety and general stress. Confirmatory factor analysis showed acceptable fit to the data for the original four-factor and the one-factor models; the four-factor model showed slightly better fit.

Conclusion: The PIP can be used in clinical practice to assess disease-related parental stress. Further psychometric testing is highly recommended.

Keywords: childhood cancer; parental stress; assessment; oncology; PIP

Introduction

When parents are confronted with a cancer diagnosis in their child, they often report to feel as if their world has fallen apart. Although most parents show remarkable resilience over time, for a subgroup of parents, levels of psychological distress remain high throughout the entire treatment period and thereafter [1–5]. Heightened levels of depression [6,7], anxiety [8–11], stress [10,12], a decreased quality of life [13], marital distress [14,15] and post-traumatic stress symptoms (e.g. [16–18]) have been reported in parents of pediatric cancer patients. This effect has been found to persist in a substantial proportion of the parents, 5 to even 10 years or longer [12,19,20]. Distress levels are highest around and shortly after diagnosis [21]. Parents experiencing most emotional problems at diagnosis and during treatment continue to report high levels of distress, even after treatment ends [12,22]. Mothers tend to report more stress than fathers [7,19] and younger parents and parents of younger children report more stress than parents of older children [23].

To obtain a better understanding of parental stress related to pediatric cancer, multi-dimensional assessment specific to the circumstances of parents of these children is needed [22,24]. In most studies, generic measures of psychological maladjustment have been used to assess parental distress. Significant differences with the reference group were found in some [25], but not in all studies [26].
Some authors argue that these traditional instruments are not sensitive enough to assess emotional and behavioral changes related to medical conditions [27]. However, when no differences are found, one cannot simply conclude that the measures are insufficient; it could also mean that there are in fact no differences between the groups.

Using disease-related measures in combination with generic measures could provide additional information that can be used to study the impact of psychosocial interventions and that will help to guide psychosocial interventions [24]. Several questionnaires have been developed specifically for parents of children with cancer or children undergoing stem cell transplantation [28–33]. The Pediatric Inventory for Parents (PIP) [23] was designed to examine areas of stress and concern in parents of children with a medical illness. It has been proven a reliable instrument for examining parent’s report of stress related to caring for a child with a serious illness, such as cancer [22,23], diabetes [34] and sickle cell disease [35]. One of the assets of the PIP is that parents are asked to rate both the frequency of stressful illness-related events and the difficulty they experience with these events. PIP total scores correlated significantly with a generic measure of state anxiety and parenting stress within a childhood oncology population [23].

The availability of assessment instruments in more than one language is low. This is unfortunate, because it would facilitate international multicenter studies and allow for cross-cultural comparisons. However, the translation in another language and culture is a lengthy and laborious process and is not always carried out adequately and/or documented properly in research articles [36].

The present study evaluates the psychometric qualities of the Dutch PIP, specifically item distribution, test–retest reliability, construct validity (by calculating correlations between the PIP and three other measures) and discriminative validity (i.e. the ability of the PIP to distinguish between known groups). Using confirmatory factor analyses (CFAs), we evaluated the original four-factor model of the PIP by examining the goodness-of-fit to the data. As the total scales are regularly used and previous studies [32] found substantial correlations (ranging from 0.45 to 0.83) between the four subscales, we also evaluated the comparative fit of the four-factor model versus the one-factor model.

Methods
Participants
All parents of children—aged 0 through 18 years—diagnosed with a malignancy 1–18 months ago between January 2005 and February 2007 in three medical centers (Leiden University Medical Center, Leiden, Wilhelmina Children’s Hospital, Utrecht, and University Medical Center, Groningen) were asked by letter to participate in the study. We chose to include parents of children diagnosed between 1 and 18 months ago in order to obtain a sample that would be more or less homogeneous with regards to ‘time since diagnosis’. Parents of deceased children were excluded from the study. In total 268 parents were approached (78 in Leiden, 60 in Utrecht and 130 in Groningen).

Procedure
The study was approved by the Medical Ethics Committee of all three medical centers. The PIP was translated by a team of five persons. An English ‘native speaker’ was asked to make a back-translation, according to the procedure described by van Widenfelt et al. [36]. The author of the PIP provided feedback on the back-translated questionnaire. The translated version was piloted with three couples and adaptations were made if necessary.

Eligible parents received information about the study and an informed consent form. In Leiden, part of the parents (40%) received the questionnaires by mail and part (60%) filled in the questionnaires in the clinic. No differences between the methods were found. Parents were instructed to fill in the questionnaires separately and not to consult each other. In Utrecht and Groningen, questionnaires were mailed to the parents’ homes along with the study information and consent forms.

One week after filling in the PIP and other questionnaires, a random half of the parents (every even numbered returned booklet) received the PIP again, to obtain test–retest reliability data. Parents who did not want to participate in the study were asked to supply demographic and illness-related data.

Measures
Parental disease-related stress was measured using the PIP. The PIP is a 42-item self-report questionnaire that measures parental stress related to the serious illness of the child with respect to (a) Communication with the child and the medical team (9 items), (b) Emotional Distress (15 items), (c) Medical Care (8 items) and (d) Role Function (10 items). Each of the 42 items is rated on two 5-point Likert-type Scales. Parents need to respond to the items twice: the first time to assess the frequency of each stressor; the second time to assess how difficult the issue has been for the parent. Parents are asked to consider last week when responding to each item. Examples of PIP-items: ‘Learning upsetting news’ or ‘Speaking with the doctor’. Higher scores refer to more stress. Adequate internal consistency (α = 0.80–96) and construct validity of the original version of the PIP (scale scores range 42–210) have been reported [23].
Parental anxiety was measured using the State Trait Anxiety Index (STAI), state and trait version, a 40-item questionnaire that measures the respondent's transitory emotional condition of stress and the general inclination toward anxiety. Dutch reference data and information on reliability (Cronbach's α = 0.95 and 0.94) and validity are available [37]. In our study, the α's were 0.95 (state anxiety) and 0.94 (trait anxiety).

Parental psychological distress was assessed using the General Health Questionnaire (GHQ), a 12-item version self-report measure of non-psychotic psychiatric disorders that can be used as a general measure for psychological distress. The psychometric properties of the Dutch version of the scale are reported to be good [38] and the questionnaire has been used frequently in research and patient care [39,40]. In our study, the α was 0.87.

Parental stress associated with raising children (i.e. parenting stress) was assessed using the Parental Stress Index, Short Form (PSI-SF), Dutch version [41]. The PSI-SF is derived from the full 123-item PSI. The PSI-SF is a reliable and valid measure and contains 25 items that are scored on a five-point continuum from strongly agree to strongly disagree. The PSI-SF differentiates well between clinical and non-clinical groups and has been used in various studies [42]. In our study, the α was 0.93.

Demographic and clinical information: Gender, age, marital status, educational level of the parent, gender and age of the target child, the child’s medical diagnosis, current treatment status and the number of weeks since diagnosis were recorded.

Statistical analyses
To assess differences between responders and non-responders, we used independent T-tests and χ² tests for categorical variables. We evaluated the normal distribution of the PIP with the test of normality and we calculated skewness and kurtosis. CFA with weighted least-squares means and variance (WLSMV)-adjusted estimation, applied to the polychoric correlation matrix, was used to evaluate the fit to the data of the original and the modified four- and one-factor models. WLSMV has been shown to perform well with ordinal variables and rather small samples [43]. The Mplus program version 2.02 [44] was used for factor analyses.

The fit of the models was assessed using practical fit indices, the values of which were evaluated according to the guidelines formulated by Marsh et al. [45]. The indices included the normed comparative fit index (CFI), the Tucker–Lewis index (TLI), the root mean square error of approximation (RMSEA) and the standardized root mean square residual (SRMR). RMSEA and SRMR values >0.10 are regarded as indication of bad fit, as were CFI and TLI values <0.85. We used the Cheung and Rensvold [46] CFI-criteria to test the difference between models: changes in CFI (ΔCFI) of –0.01 or less indicate that the hypothesis of equal fit should not be rejected, when ΔCFI lies between –0.01 and –0.02, differences may exist and definite differences between models exist when ΔCFI is greater than –0.02.

To compute test–retest reliability we used Pearson's r. To assess internal consistency of the PIP total and the four domain scales Cronbach's coefficient α was calculated. α values of 0.7 and above were considered adequate. Construct validity was examined by conducting correlation analyses between the PIP and the other psychological measures. The intercorrelations of the PIP subscales were calculated by the use of Pearson correlations. To test the discriminative validity, independent T-tests were used to examine the effect of demographic and illness-related variables on stress reported on the PIP.

Results
Participants
The overall response rate was 66% (72% in Leiden, 65% in Utrecht and 61% in Groningen). Six parents were excluded because of missing data. In total, 174 parents of 107 children participated. Of 15 single-parent families and 25 families, only one parent participated.

Reasons for refusal were the experience of too many stressful events (26%), lack of time (18%), the illness and treatment of the child was considered to have been completed too long ago (16%), too busy with work (15%), language problems (13%) or it was too confronting to the parents (10%). Non-participating parents did not differ from participating parents with regard to age, marital status, educational level, sex and age of the child. However, we did find significant differences with respect to parent gender, parent ethnicity and treatment status. In the non-participating group, the percentage of fathers, non-Dutch parents and parents with a child off treatment was higher (see Table 1).

Results on the outcome measures and demographic characteristics did not significantly differ between the three medical centers; hence, we analyzed all data together.

Item distribution and interscale correlation
First, the test of normality was performed on the different subscales and total scales of the PIP. Both total scales showed normality, the subscales significantly deviated from normality, except for two scales. Kurtosis was found for Communication Frequency (3.25, p<0.001) and Emotional Distress Frequency (2.76, p<0.01). Skewness was found for

the scales Communication Frequency, Emotional Distress Difficulty and Role Function Frequency. Interscale correlations of the PIP-sub scales varied from 0.50 (Medical Care Difficulty with Communication Frequency) to 0.82 (Communication Frequency with Emotional distress Difficulty).

Reliability

Crohnbach’s α’s for the total scales (PIP-Frequency = 0.94, PIP-Difficulty = 0.95) and for the subscales Medical Care, Emotional Distress and Role Function (0.80) were adequate. The α-value for the Communication scale was low for the PIP-Frequency and PIP-Difficulty scales (see Table 2(a)). After deletion of item number 2 Arguing with family members, the α was acceptable (0.65).

When analyzing mothers and fathers separately, we found similar reliability scores (see Table 2(b)). Mean scores between mothers and fathers differed significantly (mothers scoring higher, p < 0.05) on the Frequency scales Communication, Medical Care and Emotional Distress and the Total Frequency score. The Difficulty scale Emotional distress Difficulty
Distress and Total Difficulty scores were also significantly higher for mothers than for fathers.

Test–retest reliability after 14 days was adequate \(0.68 \leq r \leq 0.87\), based on returned questionnaires from 78 parents (33 fathers, 45 mothers) of the 111 parents approached to fill in the PIP twice (70% response rate).

**Confirmatory factor analyses**

Two identical factor models for the PIP-Frequency and the PIP-Difficulty items were (comparatively) evaluated: a four- and a one-factor model. The one-factor model, constituted by all 42 items of the PIP, reflects the possibility that one single latent dimension underlies the items. The factors of the four-factor model represent the original four scales; the factors were allowed to correlate. The fit index values are summarized in Table 3. Results indicated adequate fit of all models on the TLI but a bad fit on the remaining indices for all models except the four-factor model for the Difficulty items.

We evaluated areas of strain in the factor models using the modification indices of ML estimation. Error-correlations >0.100 between item pairs, indicating not modeled minor factors, were found for five item pairs in all models: between item pairs 14 and 16 (indicating a minor factor ‘distress over child suffering’), and between the item pairs 24–29, 24–39, 26–29, 29–36 (indicating a minor factor ‘worrying about the child’s future’). After adding the five error-correlations to the models, good fit on TLI and (nearly) acceptable fit on the remaining indices was found for all models (see Table 3). The differences in fit to the data between the adjusted four- and one-factor models were negligible, as indicated by ΔCFI<0.10. Therefore, the one-factor models, representing the total scales, may be preferred.

**Construct validity**

We calculated correlations between the PIP total scales and the other instruments. For all parents, the PIP-Frequency scale was strongly associated

---

**Table 2.** Means, standard deviations, and internal consistency of the Pediatric Inventory for (a) Parents (N = 174) and for (b) parents, mothers (N = 100) and fathers (N = 74)

<table>
<thead>
<tr>
<th></th>
<th>PIP-F</th>
<th>PIP-D</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M (SD)</td>
<td>M (SD)</td>
</tr>
<tr>
<td></td>
<td>(\alpha)</td>
<td>(\alpha)</td>
</tr>
<tr>
<td>(a) I. Communication</td>
<td>23.2 (4.5)</td>
<td>18.6 (5.4)</td>
</tr>
<tr>
<td>II. Medical care</td>
<td>24.0 (7.2)</td>
<td>16.9 (6.2)</td>
</tr>
<tr>
<td>III. Em. Distress</td>
<td>43.1 (10.1)</td>
<td>44.6 (12.6)</td>
</tr>
<tr>
<td>IV. Role Function</td>
<td>25.1 (7.1)</td>
<td>21.6 (7.4)</td>
</tr>
<tr>
<td>Total</td>
<td>115.4 (26.0)</td>
<td>101.7 (28.5)</td>
</tr>
<tr>
<td></td>
<td>(0.60)</td>
<td>(0.73)</td>
</tr>
<tr>
<td></td>
<td>(0.85)</td>
<td>(0.84)</td>
</tr>
<tr>
<td></td>
<td>(0.88)</td>
<td>(0.91)</td>
</tr>
<tr>
<td></td>
<td>(0.80)</td>
<td>(0.82)</td>
</tr>
<tr>
<td></td>
<td>(0.94)</td>
<td>(0.95)</td>
</tr>
</tbody>
</table>

**Table 3.** Goodness of fit indices for PIP four-factor and one-factor models*

<table>
<thead>
<tr>
<th>Models</th>
<th>(\chi^2)</th>
<th>df</th>
<th>CFI</th>
<th>TLI</th>
<th>RMSEA</th>
<th>SRMR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frequency</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Four-factor model</td>
<td>270.04</td>
<td>88</td>
<td>0.869</td>
<td>0.933</td>
<td>0.109</td>
<td>0.091</td>
</tr>
<tr>
<td>One-factor model</td>
<td>292.91</td>
<td>88</td>
<td>0.852</td>
<td>0.924</td>
<td>0.105</td>
<td>0.091</td>
</tr>
<tr>
<td>Four-factor model with (\theta_{14,16}, \theta_{24,29}, \theta_{24,36}, \theta_{29,36}) free</td>
<td>229.35</td>
<td>88</td>
<td>0.891</td>
<td>0.944</td>
<td>0.099</td>
<td>0.084</td>
</tr>
<tr>
<td>One-factor model with (\theta_{14,16}, \theta_{24,29}, \theta_{24,36}, \theta_{29,36}) free</td>
<td>249.95</td>
<td>88</td>
<td>0.883</td>
<td>0.940</td>
<td>0.103</td>
<td>0.086</td>
</tr>
<tr>
<td>Difficulty</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Four-factor model</td>
<td>235.73</td>
<td>88</td>
<td>0.895</td>
<td>0.956</td>
<td>0.098</td>
<td>0.081</td>
</tr>
<tr>
<td>One-factor model</td>
<td>253.06</td>
<td>87</td>
<td>0.882</td>
<td>0.950</td>
<td>0.105</td>
<td>0.085</td>
</tr>
<tr>
<td>Four-factor model with (\theta_{14,16}, \theta_{24,29}, \theta_{24,36}, \theta_{29,36}) free</td>
<td>204.44</td>
<td>87</td>
<td>0.909</td>
<td>0.962</td>
<td>0.091</td>
<td>0.079</td>
</tr>
<tr>
<td>One-factor model with (\theta_{14,16}, \theta_{24,29}, \theta_{24,36}, \theta_{29,36}) free</td>
<td>227.85</td>
<td>87</td>
<td>0.900</td>
<td>0.957</td>
<td>0.096</td>
<td>0.082</td>
</tr>
</tbody>
</table>

*Weighted least-squares means and variance adjusted estimation were applied to the polychoric correlation matrix; CFI, Comparative Fit Index; TLI, Tucker–Lewis index; RMSEA, Root mean square error of approximation; SRMR, standardized root mean square residual.

---

\(p < 0.05\).
with STAI state and trait \((r = 0.52\) and \(r = 0.55, p < 0.01)\) and with the GHQ \((r = 0.54, p < 0.01)\) and weakly with the PSI-SF \((r = 0.19, p < 0.05)\). The PIP-Difficulty scale was strongly related to STAI state and trait and GHQ \((r = 0.59, r = 0.66\) and \(r = 0.51,\) respectively, \(p < 0.01)\) and weakly with the PSI-SF \((r = 0.24, p < 0.01)\). When analyzing fathers and mothers separately, the correlations were comparable.

**Discriminative validity**

Mothers reported higher scores than fathers (PIP-Frequency, \(p < 0.01, t = 2.84)\). Older fathers (i.e. fathers above the mean age of 41 years at assessment) reported significantly more distress than younger fathers (PIP-Difficulty, \(p < 0.05, t = 2.19)\). Interestingly, older fathers also reported significantly higher state anxiety levels. For mothers, no age effect was found. Parents of younger children (under versus over 115 months) reported higher stress scores than parents of older children (PIP-Difficulty, \(p < 0.05, t = 2.11)\). Parents of children on treatment had significantly higher PIP-scores than parents whose children had completed treatment (PIP-Frequency, \(p < 0.05, t = 2.92)\) and parents of children diagnosed more recently—less than 40 weeks ago—reported more stress (PIP-Frequency, \(p < 0.05, t = 2.49)\) than parents of children who were diagnosed longer ago.

**Item–subscale and item–total correlations**

Item–subscale correlations varied from –0.07 to 0.67 in the Frequency Scale, with a mean of 0.46. For the Difficulty items, item–subscale correlations varied from 0.23 to 0.79 with a mean of 0.53. Lowest item–subscale correlations were found for the scale Communication. Item–total correlations varied from 0.02 to 0.70 for the Frequency items (mean 0.47) and from 0.32 to 0.77 for the Difficulty items (mean 0.55).

**Discussion**

Gaining insight into parents’ stress following pediatric cancer is increasingly important in order to deliver adequate psychosocial care to the entire family. Disease-related measures can add important information about parental adaptation to stressful illness-related situations. Results regarding the Dutch version of the Pediatric Inventory for Parents (PIP), a disease-related measure of parental stress, are satisfactory. We found adequate (test–retest) reliability scores for the PIP total scales and three of the four subscales (Medical Care, Emotional Distress and Role Function). The fourth subscale, Communication, needs improvement. This last finding is not in line with the results from the original study [23]. Cultural differences with regards to communicating with hospital staff and family could perhaps explain part of this difference in results.

PIP-scores correlated strongly with a generic measure of anxiety and general psychological functioning. This means that disease-related distress, although it measures a different construct, can have considerable overlap with general well-being and anxiety. The added value of the PIP, however, is that it asks parents about their stress concerning disease-related situations. Scores on the PIP could be transformed into an individual ‘stress profile’, which could be used to tailor psychosocial support.

The low correlation of PIP-scores with parenting stress scores suggests that stress resulting from difficulties disciplining and setting limits to one’s child (parenting stress) differs from stress associated with having a child with a serious illness (parental stress). In various studies, e.g. [47], the PSI is used as a measure of parental stress instead of stress associated with parenting. This strategy might well result in drawing the wrong conclusions about the stress reactions that parents can have as a result of their child’s illness.

As expected and in line with other research [7,11], mothers showed higher PIP-scores than fathers, parents closer to diagnosis and parents of younger children reported more stress. Older fathers reported higher stress levels. For mothers, no age effect was found. This result is contrary to the findings of the original PIP-study, in which younger parents (both mothers and fathers) reported more stress [23].

CFAs showed that not only the four-factor model, representing the four subscales of the PIP, but also the one-factor model, representing the total scales, showed acceptable fit to the data after three items were dropped from the models and the error-correlation of one item pair was added. The total scales may be regarded as sufficient for practical purposes as the difference in fit between the two models was minor, and very strong correlations were found between the four latent factors.

One of the advantages of the PIP is subdivision of parental stress levels into Frequency and Difficulty scores (although the scales correlate highly and thus outcomes might be partly overlapping), which enables the psychosocial team to target interventions more precisely to the needs of parents in different phases of their child’s treatment. In our study, PIP-Frequency scores discriminated between parents of children on treatment and parents whose children have ended treatment. However, PIP-Difficulty scores for the two groups were equal. This finding may imply that although the frequency of stressful disease-related events is lower in parents of children off treatment, the perceived difficulty of these events remains similar.
Limitations and practical implications

Despite the results of the present study, there are some limitations that need mentioning. First, the age range of the children in our study group varied widely, making comparison of parental stress levels difficult. Being the parent of an ill baby or toddler versus a teenager will render different sources of stress. Second, a substantial proportion of parents refused participation. Reasons for non-participation ranged from being too stressed to considering the treatment of the child to have been completed too long ago or being too busy with other things like work. It is unclear if this caused an under- or overestimation of parental stress levels.

Finally, the procedure of administering the questionnaires was different in the three hospitals. Approaching parents face to face in the clinic or the outpatient’s ward yielded a higher response rate than mailing the questionnaires. However, this did not seem to influence reported levels of stress.

One of the assets of the study is the relatively large, multi-center study group. Furthermore, we managed to include a large percentage of fathers in our study. The PIP could be used in regular patient care to assess all parents of newly diagnosed children at critical time points in therapy: shortly after diagnosis, then again after 4–6 months (usually seen as the ‘stabilization phase’) and by the end of treatment. These time points seem to cover the process of parental stress through the phases of childhood cancer well.

In summary, the Dutch PIP is a reliable and valid assessment tool to gain insight into stress experienced by parents during the course of their child’s cancer treatment. Continuous psychometric testing is recommended in different populations and at different time points.

Acknowledgements

The authors thank the parents for taking the time to complete the questionnaires. Secondly, they express their thanks to the members of the translation team (Natasja van Lang, Bas Berger and Brigit van Widenfelt) and to Arnold Goedhart for his help with the analyses. Finally, they thank Randi Streisand for her consent to translate the PIP in Dutch and for her comments on the manuscript.

References


