Screening for potential child maltreatment in parents of a newborn baby

The predictive validity of an instrument for early identification of parents at risk for child abuse and neglect (IPARAN)

van der Put, C.E.; Bouwmeester-Landweer, M.B.R.; Landsmeer-Beker, E.A.; Wit, J.M.; Dekker, F.W.; Kousemaker, N. Pieter J.; Baartman, H.E.M.

Link to publication

CrossMark

Screening for potential child maltreatment in parents of a newborn baby: The predictive validity of an Instrument for early identification of Parents At Risk for child Abuse and Neglect (IPARAN)

Claudia E. van der Put[a,*], Merian B.R. Bouwmeester-Landweer[b,1], Eleonore A. Landsmeer-Beker[b,2], Jan M. Wit[b], Friedo W. Dekker[c], N. Pieter J. Kousemaker[d], Herman E.M. Baartman[e]

[a] Research Institute of Child Development and Education, University of Amsterdam, The Netherlands
[b] Department of Pediatrics, Leiden University Medical Center, Leiden, The Netherlands
[c] Department of Clinical Epidemiology, Leiden University Medical Center, Leiden, The Netherlands
[d] Leiden University, Leiden, The Netherlands
[e] Department of Orthopedagogics, Free University, Amsterdam, The Netherlands

ARTICLE INFO

ABSTRACT

For preventive purposes it is important to be able to identify families with a high risk of child maltreatment at an early stage. Therefore we developed an actuarial instrument for screening families with a newborn baby, the Instrument for identification of Parents At Risk for child Abuse and Neglect (IPARAN). The aim of this study was to assess the predictive validity of the IPARAN and to examine whether combining actuarial and clinical methods leads to an improvement of the predictive validity. We examined the predictive validity by calculating several performance indicators (i.e., sensitivity, specificity and the Area Under the receiver operating characteristic Curve [AUC]) in a sample of 4692 Dutch families with newborns. The outcome measure was a report of child maltreatment at Child Protection Services during a follow-up of 3 years. For 17 children (.4%) a report of maltreatment was registered. The predictive validity of the IPARAN was significantly better than chance (AUC = .700, 95% CI [.567–.832]), in contrast to a low value for clinical judgement of nurses of the Youth Health Care Centers (AUC = .591, 95% CI [.422–.759]). The combination of the IPARAN and clinical judgement resulted in the highest predictive validity (AUC = .720, 95% CI [.593–.847]), however, the difference between the methods did not reach statistical significance. The good predictive validity of the IPARAN in combination with clinical judgment of the nurse enables professionals to assess risks at an early stage and to make referrals to early intervention programs.

Child maltreatment is a serious problem, both internationally and in The Netherlands. In The Netherlands, the annual prevalence rate is estimated at 9.9% based on self-reports (Alink et al., 2011), which is comparable to the prevalence rate in the United States

(Finkelhor, Turner, Ormrod, & Hamby, 2009). Young children have the highest rate of victimization, and child maltreatment at an early age has major negative consequences on the child's socio-emotional development, attachment security, physical growth and development, and quality of life (Eigsti & Cicchetti, 2004; English, Thompson, Graham, & Briggs, 2005; Scarborough & McRae, 2008; Straus & Smith, 1993). Research suggests that early intervention programs can be effective in preventing child maltreatment (Dombrowski, Timmer, Blacker, & Urquiza, 2005; Gershater-Molko, Lutzker, & Wesch, 2003; Reynolds, Mathieson, & Topitzes, 2009; Saunders & O'Brien, 1997). In order to be able to prevent child maltreatment it is necessary to screen for potential maltreatment in the general population, for example during youth health care visits up to the age of four years or even before birth (prenatally). The availability of reliable and valid screening instruments would enable professionals to assess risks and to make referrals to early intervention programs. However, the development and validation of screening instruments for child maltreatment in the general population is still in its infancy (Nygren, Nelson & Klein, 2004). Only a few instruments are available worldwide and most of them have only been validated to a limited extent.

There are two major approaches to risk assessment in child protection: actuarial and clinical judgment. In actuarial approaches, conclusions are based solely on empirically established relations between data and the outcome of interest, whereas in clinical approaches, conclusions are based on the judgment of a professional who interprets, combines and weighs information in a subjective manner (Dawes, Faust, & Meehl, 1989). Clinical instruments can be further classified into (a) unaided decision-making based on experience, knowledge and intuition (unstructured clinical judgment), (b) tools based on the opinions of experts, but often without an empirical basis (consensus-based instruments), and (c) empirically based tools that leave the final decision-making process to the professional (structured clinical judgment; SCJ).

Worldwide, clinical approaches are more common in child protection practices than actuarial instruments. This is remarkable because international validation studies indicate that the performance of most clinical methods is questionable (Arad-Davidson & Benbenishty, 2008; Baird & Wagner, 2000; Barlow et al., 2010; Camasso & Jagannathan, 2000; D'Andrade, Benton, & Austin, 2005; DePanfilis & Girvin, 2005; Dorsey, Mustillo, Farmer, & Elbogen, 2008; Knoke & Trocmé, 2005; Lyons, Doueck, & Wodarski, 1996; Munro, 1999; Pfister & Böhm, 2008; Wald & Woolverton, 1990). Validation studies have even shown that some widely used clinical instruments perform no better than chance, meaning that in many cases an incorrect judgement is made (Baird & Wagner, 2000; Barber, Shlonsky, Black, Goodman, & Trocmé, 2008; Van der Put, Assink, & Stams, 2016). Moreover, studies comparing different methods have consistently shown that actuarial approaches outperform clinical approaches at estimating risks within different domains such as child welfare, criminal justice, forensic mental health, and clinical psychology (Baird & Wagner, 2000; D'Andrade et al., 2005; Dawes et al., 1989; Grove & Meehl, 1996; Leschied, Chiodo, Whitehead, Hurley, & Marshall, 2003). One of the explanations for the superior predictive performance of actuarial methods is that the reliability of actuarial instruments is higher, because the scoring and combining of risk factors occurs according to a fixed algorithm, meaning that professionals use the same scoring rules, whereas in clinical methods the scoring and combining of risk factors is done in a subjective fashion (e.g., Dawes et al., 1989; Gambrill & Shlonsky, 2000) (Fig. 1).

At the time that this study was started, no screening instruments were available in The Netherlands for assessing the risk of child maltreatment in the general population. Only instruments for assessing the risk of recurrence of child maltreatment were used, namely the Light Instrument Appraisal Child Maltreatment (LIRIK; Ten Berge & Eijgenraam, 2009), the Child Abuse Risk Evaluation-Netherlands (CARE-NL; De Ruiter & De Jong, 2005) and the Delta Safety List (Heinrich & Braak, 2007), which are all clinical instruments. In addition, an actuarial instrument, the California Family Risk Assessment (CFRA), was used in a regional pilot of The Netherlands to assess the risk of child maltreatment in families who were referred to specialized family support because of parenting and/or child developmental problems (Van der Put, Hermanns, Van Rijn-Van Gelderen, & Sondeijker, 2016). The CFRA is an actuarial risk assessment instrument originally designed for assessing the risk of child maltreatment subsequent to receipt of an initial maltreatment report (Johnson, 2011).

The difference between screening (assessing potential for child maltreatment before it occurs) and assessing risk of recurrence of child maltreatment is crucial since the populations assessed and their risk of child maltreatment differ, as do relevant predictive factors (Cash, 2001). Screening aims to assess the risk of child maltreatment in the general population where the risk of child maltreatment is relatively small, whereas risk assessment aims to assess the risk of (repeated) child maltreatment in high-risk groups such as families under the guidance of child protection services. In scientific literature, there is particular emphasis on instruments assessing the risk of recurrence of child maltreatment and much less on screening instruments for assessing the risk of child maltreatment in the general population. This is because assessing the risk of recurrence of child maltreatment is the most commonly employed prognostic process in child welfare services involved in maltreatment.

Because no screening instruments were available, we developed a new actuarial instrument for parents with newborns, the Instrument for Identification of Parents At Risk for child Abuse and Neglect (IPARAN; Bouwmeester-Landweer, 2006; see instruments section). The development of the IPARAN was part of a large Dutch research and intervention project in collaboration with the Dutch Youth Health Care Centers (YHCs). The aim of the project was to provide parents at increased risk of child maltreatment with a preventive intervention by means of home-visits during the first 18 months of life of their newborn baby (Supportive Parenting Intervention). In The Netherlands, YHCs monitor and promote the health and development of children aged 0–4 years, and are the first portal to Child Health Care in The Netherlands. They are comparable to Well Baby Clinics, with the exception that they continue to monitor the child until the age of four years (since 2015 even until the age of 18 years). The nurses and doctors working at these YHCs regularly examine nearly all infants and young children in The Netherlands (95–98% of all infants aged 0–1 years; Burgmeijer, Van Geenhuizen, Filedt kok-Weimar, & De Jager, 1997; Dunnink, 2010). They have a primary function in screening for risk factors regarding the child's physical, intellectual, social and psychological development, including problematic child-rearing situations and the first signs of child maltreatment. The importance of the role of nurses in the assessment of the risk or actual presence of

maltreatment has been stressed in some early publications (e.g. Dean, MacQueen, Mitchell, & Kempe, 1978; Jolly, 1979).

The following research questions were examined in this study: 1) what is the strength of the associations between the individual IPARAN-items and later reports of child maltreatment; 2) what is the predictive validity of the IPARAN; 3) what is the predictive validity of clinical judgment of nurses of the Youth Health Care Centers (YHCs); and 4) what is the predictive validity of the combination of the IPARAN and clinical judgment.

1. Method

1.1. Sample

In The Netherlands, all parents of newborns are routinely seen in the YHCs for health and developmental check-ups. For the present study, all families with newborns in the northern part of the province South Holland (N = 8899) were approached with the IPARAN during a 13-month period starting October 2001. The semi-urban northern part of the province South Holland contained approximately 4% of the Dutch population and provided for 3.9% of the annual births in The Netherlands. A comparison of the regional and national data regarding demographic variables such as single parent households and number of children showed that this sample was representative for the Dutch population (Bouwmeester-Landweer, 2006).

Eventually 4899 out of 8899 families responded (55.1%), 4882 mothers and 4704 fathers. In 95.8% of these families (n = 4692), both father, mother and the nurse completed their part of the questionnaire. Table 1 presents the background characteristics of the families in the sample.

1.2. Measures

1.2.1. The Instrument for Identification of Parents At Risk for child Abuse and Neglect (IPARAN)

The IPARAN is an actuarial instrument for early identification of families at risk for child maltreatment (Bouwmeester-Landweer, 2006). Actuarial means that the overall risk score is calculated based on the scores of identified risk factors. Presently, the IPARAN is used in a large number of YHCs in The Netherlands with the aim to select parents at increased risk of child maltreatment to provide these parents with a preventive intervention by means of home-visits during the first 18 months of life of their newborn baby (Supportive Parenting Intervention).

The IPARAN was based on a review of the available literature on risk factors for child maltreatment in the general population. Following the ecological model introduced by Belsky (Belsky & Vondra, 1989); as well as the concept of parental awareness (parental childrearing attitudes and conviction) introduced by Newberger and elaborated on by Baartman (Baartman, 1996; Newberger, 1980), these factors have been categorized within three domains: 1) the parental developmental history and personality (ontogenic system), including factors related to parental awareness; 2) child and family characteristics (microsystem) and 3) characteristics of the social context (exosystem). Within each domain several items were identified which constituted the basis for the questions as formulated in the questionnaire.

Identified risk factors in the ontogenic system were: parents' own childhood experience of physical (Agathonos-Georgopoulou & Browne, 1997; Egeland, 1988; Sidebotham & Golding, 2001), emotional or sexual abuse (Pringle, 1980; Straus, 1978), both intra- and extra-familial (Egeland, 1988; Oates, Tebbutt, Swanston, Lynch, & O'Toole, 1998; Sidebotham & Golding, 2001), parents having witnessed parental violence during childhood (Kotch et al., 1995), strong belief in physical punishment (Brown, Cohen, Johnson, & Salzinger, 1998; Browne, Davies, & Stratton, 1988), ambivalent feelings about becoming a parent (Browne et al., 1998; Egeland, 1988), ambivalent feelings about parental competence (Cowen, 1992; Pringle, 1980), young parental age (Browne et al., 1998; Sidebotham & Golding, 2001), the experience of any psychiatric or psychological problems (i.e. depression, summarized by the term dysphoria) during the last three years (Kotelchuck, 1982; Kotch et al., 1999; Sidebotham & Golding, 2001), substance or alcohol dependency (Kotch et al., 1995; Sidebotham & Golding, 2001; Schumacher, Smith Slep, & Heyman, 2001) and parents' impaired temper inhibition (Browne et al., 1988; Cowen, 1992; Pringle, 1980). Identified risk factors in the microsystem were: the tendency to solve partner-conflicts with physical force (Browne et al., 1998; Cowen, 1992), low spousal support expectancy (Cowen, 1992; Kotelchuck, 1982), single parenthood (Browne et al., 1988; Pringle, 1980), child prematurity and dysmaturity (small for gestational age) (Browne et al., 1988; Sidebotham & Heron, 2003). The identified risk factor in the exosystem was social isolation (Kotelchuck, 1982; Sidebotham, Heron, & Golding, 2002; Schumacher et al., 2001).

The IPARAN is meant to be a short questionnaire that could be answered in relatively short time (less than 7 min). Therefore most

Table 1
Background Characteristics of the families in the sample (n = 4692).

	N	%
Single mother	233	5.0%
Single father	147	3.1%
Duration of pregnancy < 37 weeks	230	4.9%
Birth weight < 2500 g	290	6.2%
Age mother < 18	11	.2%
Age father < 18	1	.0%

Table 2
Items of the IPARAN assessing the likelihood of future child maltreatment.

	Item	Responses (scoring)
1.	I expect that I will receive enough support from my partner in the care of our child(ren)	1. always (0), 2. often (0), 3. sometimes (.5), 4. never (.5)
2.	I am very worried about raising this child	1. always (.5), 2. often (.5), 3. sometimes (0), 4. never (0)
3.	During the last nine months I have felt unhappy about expecting a baby	1. always (.5), 2. often (.5), 3. sometimes (0), 4. never (0)
4.	As a child I generally felt safe in the vicinity of (one of) my parents/carers	1. always (0), 2. often (0), 3. sometimes (1), 4. never (1)
5.	I feel that the relationship between my parents/carers was often violent when I was a child	1. yes (1), 2. no (0)
6.	I feel that my parents/carers hit me too much as a child	1. yes (1), 2. no (0)
7.	Looking back at the last three years, I suffered serious mental problems (for a while)	1. yes (1), 2. no (0)
8.	When I have/had an argument with my partner, we sometimes come/came to blows	Yes (1), No (0)
9.	I should actually drink less alcohol/take fewer drugs	1. yes (1), 2. no (0)
10.	A good upbringing involves smacking children regularly	1. yes (1), 2. no (0)
11.	I was forcefully subjected to an unpleasant sexual experience before I was 16	1. yes (1), 2. no (0)
12.	I feel very accepted by my family	1. always (0), 2. often (.1), 3. sometimes (.2), 4. never (.3)
13.	I can get so angry that I lose control	1. always (1), 2. often (1), 3. sometimes (0), 4. never (0)
14.	I feel very much at home in the area where I live	1. always (0), 2. often (.1), 3. sometimes (.2), 4. never (.3)
15.	I would like to have more people around me on whom I could fall back for support	1. always (.3), 2. often (.2), 3. sometimes (.1), 4. never (0)
16.	I find it easy to ask for help or advice when I need it	1. always (0), 2. often (.1), 3. sometimes (.2), 4. never (.3)

risk factors are targeted with one item only, with the exception of social isolation. This risk factor is covered by four items, addressing the inability to ask for help if needed, the need for more people to rely on, insufficient acceptance in the larger family and a sense of alienation from the neighborhood. The IPARAN is a three-page self-administered questionnaire for parents with 16 questions for mother and father individually. It was tested for comprehensiveness in a small pilot-study, involving 74 families of whom 90.5% responded (Tan, 2001). Based on the results of this pilot the phrasing of several questions was ameliorated and some answering options were changed. Table 2 shows the items, the response categories and the scoring of the IPARAN.

1.2.2. Selection of high risk families based on the IPARAN

For both fathers and mothers, a total score was calculated by adding the scores of the individual risk factors (16 items, either 0, 0.1, 0.2, 0.3, 0.5 or 1 point each; see Table 2). The selection of high risk families was based on a combination of the total score of father and mother. A family was classified as high risk if either parent had a total score of 1 or more. So the total scores of fathers and mothers were not added together.

1.2.3. Selection of high risk families based on clinical judgment

To measure clinical judgment of the nurses of the Youth Health Care Centers, another one-page questionnaire was developed in which the nurse was asked to provide some demographics as well as her subjective risk assessment of each particular family, expressed as 'concern'. This concern was the nurses' subjective estimation of need for support in a family ("yes"/"no").

1.2.4. Selection of high risk families based on the combination of IPARAN and clinical judgement

A family was classified as high risk if either the IPARAN classified the family as high risk (total score of 1 or more), or the nurse classified the family as high risk (based on clinical judgment).

1.3. Outcome measure

Verified reports of child maltreatment in the family at the Dutch Child Protection Services (CPS) during a follow-up period of 3 years after completing the IPARAN. CPS has a national network of hotlines where citizens and people who work professionally with children can discuss possible cases of child maltreatment and file a report if necessary.

1.4. Procedure

All 83 nurses employed by three YHC-organizations in the region agreed to participate, both by filling in the one-page questionnaire in which the nurse was asked to provide her risk assessment of each particular family and by explaining the project to parents when needed. All nurses in the project region were female and had followed a training program on child maltreatment within two years prior to the start of the project. To ensure a homogeneous presentation of the project a specific training program was

provided to all nurses, addressing ways to explain the project and the questionnaire to parents and targeting communication skills through role-play. During this training reasons for the nurses' 'concern' were stipulated, such as problematic family interactions, lack of hygiene or safety in the house, parental unsteadiness or parental distrust of (mental) health care institutions.

The IPARAN was sent to all families with newborns within no more than five days after birth. In The Netherlands, YHCs provide one standard home-visit to parents of newborns around the 14th day after birth. This allowed five to eight days for parents to fill out their questionnaire. During the standard home-visit the visiting nurse could answer questions about the questionnaire, complete her own questionnaire and return them both to the study coordinator. Nurses were encouraged to fill out their questionnaire in the presence of the family. In families with multiple children the visiting nurse was allowed to include possible knowledge of the family history from prior visits in her 'concern'. Parents who needed more time to fill out their questionnaire could take it to the YHC at their first visit, approximately one month after the birth of their child, or submit it by mail. To increase response-rate both nurses and parents received written reminders to send in their questionnaire. In case of a nurses' response indicating concern about a non-respondent family the nurse was asked to request the parents (again) to send in the questionnaire.

The Ethics Committee of the Leiden University Medical Center (LUMC) approved this study. In the letter accompanying the questionnaire it was explained that the program concerned parenting support specifically aimed at families facing various difficulties. It was emphasized that any response was appreciated, even from families who felt the program was not applicable to them. Furthermore it was stressed that participation was strictly voluntary.

1.5. Analyses

To examine the strength of the associations between the individual items of the IPARAN and future reports of maltreatment, point-biserial correlations (r_{pb}) were calculated, separately for fathers and mothers.

The predictive validity of the clinical judgment of the nurses and the IPARAN was assessed by examining the Area Under the receiver operating characteristics Curve (AUC) and the sensitivity and specificity at various cut-off scores. The AUC value indicates the probability that a randomly selected family that will be reported for suspected child maltreatment within a follow-up period of 3 years has a higher risk classification than a randomly selected family that will not be reported for suspected child maltreatment within three years (Hanley & McNeil, 1983). An AUC value of .50 indicates that the instrument performs no better than chance. A value of 1.00 indicates a perfect positive prediction, a value of .00 a perfect negative prediction. AUC values of .639–0.714 correspond with a medium effect size (.50 < d < .80) and AUC values of .714 and higher correspond with a large effect size (d > .80; Rice & Harris, 2005).

Sensitivity is the probability of a positive result for children who will actually be reported for child maltreatment, and specificity is the probability of a negative result for children who will not be reported for child maltreatment. Ideally, both sensitivity and specificity are 100%.

The method of DeLong, DeLong and Clarke-Pearson (1988) was used to test whether the AUC values of the IPARAN and the clinical judgment of the nurse differed significantly.

2. Results

2.1. Association between the individual items of the IPARAN and future reports of child maltreatment

Table 3 shows the point-biserial correlations (r_{pb}) between the individual items of the IPARAN and reports of maltreatment during the follow-up period of 3 years after completing the IPARAN. The overall rate of reports of child maltreatment was .4% (n = 17). The r_{pb} values for small, medium and large effect sizes for a 50% base rate are .10, .24 and .37 respectively (Rice and Harris, 2005). For base rates other than 50%, r_{pb} values for small, medium, and large effects can be calculated using the conversion formulae (after Rosental, 1991; Swets, 1986) provided by Rice and Harris (2005). Using that formulae, we found the r_{pb} values for small, medium and large effect sizes for a .4% base rate to be .01, .03 and .05 respectively.

Seven items were significantly related to reports of maltreatment in both fathers and mothers, i.e. ambivalence about parental competence, parents' emotional maltreatment during their childhood, witness of parental violence during their childhood, physical maltreatment during childhood, spousal violence, impaired inhibition with regards to temper and needing more people to rely on. Four items were significantly related to reports of maltreatment only in fathers, i.e., addiction to alcohol/drugs, sexual abuse before the age of 16, insufficiently accepted by family and being unable to ask for help if needed. Finally, two items were significantly related to future reports of maltreatment only in mothers, i.e. ambivalence during pregnancy and dysphoria.

2.2. Predictive validity of the IPARAN and nurses' concern

Table 4 shows the sensitivity, specificity and AUC values of a) clinical judgement of the nurse, b) the IPARAN and c) the combination of the IPARAN and clinical judgment. The AUC value of the IPARAN is .700, 95% CI [.567–.832], significantly higher than chance (.50). In contrast, the AUC value of the clinical assessment of the nurse is .591, 95% CI [.422–.759], not significantly higher than chance. The combination of the IPARAN and clinical judgement resulted in an AUC value of .720, 95% CI [.593–.847]). The difference between the three AUC values did not reach statistical significance. Fig. 1. shows the ROC curves for: a) the IPARAN, b) clinical judgement of the nurse, and c) the combination of IPARAN and clinical judgement of the nurse.

Table 3

Point-biserial correlations between IPARAN-items and future reports of child maltreatment, separately for mothers and fathers.

Item		Mother (n = 4882)	Father (n = 4704)
1.	Insufficient support expected from spouse	.028	.015
2.	Ambivalent about parental competence	.034[*]	.046[**]
3.	Ambivalence during pregnancy	.044[**]	.012
4.	Emotional maltreatment during childhood	.033[*]	.043[**]
5.	Witness of parental violence during childhood	.063[***]	.031[*]
6.	Physical maltreatment during childhood	.038[*]	.060[***]
7.	Dysphoria	.030[*]	.013
8.	Spousal violence	.151[***]	.113[***]
9.	Addiction to alcohol/drugs	− .004	.063[***]
10.	Belief in physical punishment	.022	.020
11.	Sexual abuse before the age of 16	.014	.034[*]
12.	Insufficiently accepted by family	.010	.033[*]
13.	Impaired inhibition with regards to temper	.106[***]	.094[***]
14.	Alienated from neighborhood	.013	.027
15.	Needing more people to rely on	.066[***]	.077[***]
16.	Unable to ask for help if needed	.015	.035[*]

* $p < .05$.
** $p < .01$.
*** $p < .001$.

3. Discussion

The aims of this study were to examine a) the predictive validity of the IPARAN and b) whether the combination of actuarial assessment (IPARAN) and clinical assessment of nurses of the Youth Health Care Centers (YHCs) leads to an improvement of the predictive validity. IPARAN-items most strongly related to future reports of child maltreatment were spousal violence, impaired inhibition with regards to temper, needing more people to rely on and parents' own experiences with child maltreatment. Cash (2001) found that risk factors for initial maltreatment are associated with parental depression, substance abuse, unemployment, social isolation, increased stress, unrealistic expectations of the child and parents' history of being abused. In the present study, parental depression was only related to future reports of child maltreatment in mothers. Substance abuse was only related to future reports of child maltreatment in fathers. Some of the measured items of the IPARAN were not or only weakly associated with future reports of child maltreatment, such as insufficient support expected from spouse, belief in physical punishment, and alienation from neighborhood. In future research, with a longer follow-up period and a more comprehensive outcome measure, it should be examined whether these items can be removed from the IPARAN without compromising on predictive validity.

The predictive validity of the IPARAN was acceptable with an AUC of .700, 95% CI [.567–.832], whereas the predictive validity of the clinical judgement of the nurses was poor, with a non-significant AUC of .591, 95% CI [.422–.759], meaning that clinical judgment performed no better than chance. Especially the sensitivity of nurses' clinical judgement was very low with a value of .214 meaning that only 21.4% of the reports of child maltreatment was predicted accurately. The results of this study with regard to the poor predictive power of clinical judgment is in line with results of previous findings. Unaided clinical judgement in relation to the assessment of risk of child maltreatment is widely recognised to be flawed (Arad-Dvaidson & Benbenishty, 2008; DePanfilis & Grivin, 2005; Munro, 1999; Pfister & Böhm, 2008). The combination of the IPARAN and clinical judgement resulted in the highest predictive validity (AUC = .720, 95% CI [.593–.847]), however, the difference between the methods did not reach statistical significance.

Some other limitations of the study need to be mentioned. Most limitations were related to the outcome measure of our study, which consisted of reports of child maltreatment at CPS during a follow-up period of 3 years after completing the IPARAN. First, possibly partly because of the relatively short follow-up period, there were reports of child maltreatment in only .4% (n = 17) of the families. A second limitation of the outcome measure is that not every case of child maltreatment is reported to CPS. The Netherlands' Prevalence study on Maltreatment of children and youth estimated the number of reported cases of child maltreatment to CPS to be around 20% of the total number of cases of child maltreatment (Alink et al., 2011). Unfortunately, we did not have the possibility to ask multiple resources, such as police officers, teachers, day-care workers, and social workers at the Child Welfare Agency, about suspected child maltreatment during the follow-up period. Therefore, the percentage of children suspected for child maltreatment as obtained in our study may well be an underestimation of the actual rates. Third, some forms of maltreatment, such as emotional

Table 4

Sensitivity, specificity and AUC values for clinical judgement of the nurse, IPARAN and the combination of IPARAN and clinical judgement.

	Sensitivity	Specificity	AUC	95% CI
Clinical judgment nurse	.214	.967	.591	(.422–.759)
IPARAN	.611	.786	.700	(.567–.832)
Combination of IPARAN and clinical judgment	.667	.774	.720	(.593–.847)
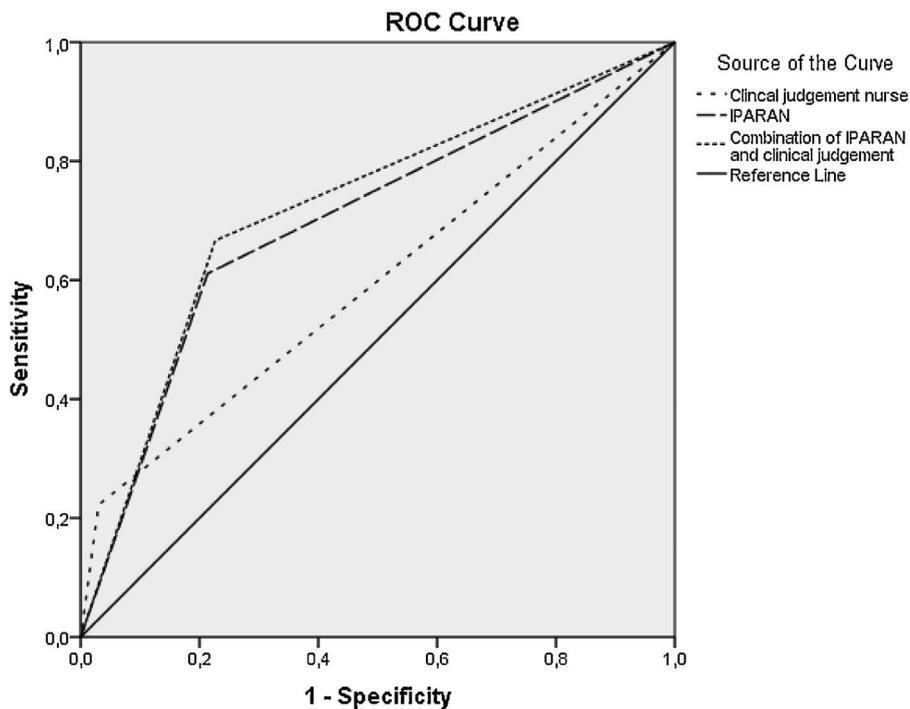
Fig. 1. Shows the ROC curves for: a) the IPARAN, b) clinical judgement of the nurse, and c) the combination of IPARAN and clinical judgement of the nurse.

abuse and neglect, are less visible and therefore more difficult to detect (Glaser, 2002). As a result, the number of cases of emotional abuse and neglect may be underreported. To summarize, the abovementioned limitations of our outcome measure may have influenced the results of our study and, therefore, the predictive validity of the IPARAN should be examined with a longer follow-up period and a more comprehensive outcome measure. Related to this, both sensitivity (.611) and specificity (.786) of the IPARAN might be considered low for a screening tool for the general population. However, these values are related to the specific outcome measure that is used in this study (the occurrence of *a report* of child maltreatment at CPS within a period of *three years* after completing the instrument). As indicated above, it is estimated that only about 20% of the total number of cases of child maltreatment is reported to CPS (Alink et al., 2011). In addition, the aim of the IPARAN is not only to predict actual cases of child maltreatment, but also the problems and situations that precede child maltreatment (i.e. risk factors regarding the child's physical, intellectual, social and psychological development, including problematic child-rearing situations and the first signs of child maltreatment). Therefore, we expect the measured values of both sensitivity and specificity to be an underestimation of the intended performance of the IPARAN.

Despite these limitations, the results of this study may be helpful for future studies and programs on screening for potential maltreatment in the general population, aimed at preventing child maltreatment at an early stage. The good predictive validity of the IPARAN in combination with clinical judgment of the nurse enables professionals to assess risks at an early stage and to make referrals to early intervention programs such as the Supportive Parenting intervention program. Future research should examine whether the IPARAN can be shortened without compromising on predictive validity which is of importance to save time and costs and to make the IPARAN more widely applicable.

References

Agathonos-Georgopoulou, H., & Browne, K. D. (1997). The prediction of child maltreatment in Greek families. *Child Abuse & Neglect, 21*, 721–735.

Alink, L., IJzendoorn van, M., Bakermans-Kranenburg, M., Pannebakker, F., Vogels, A., & Euser, S. (2014). *Kindermishandeling 2010. De tweede nationale prevalentiestudie mishandeling van kinderen en jeugdigen (NPM-2010)*. Leiden: Leiden University/TNO.

Arad-Davidson, B., & Benbenishty, R. (2008). The role of workers' attitudes and parent and child wishes in child protection workers' assessments and recommendation regarding removal and reunification. *Children and Youth Services Review, 30*, 107–121.

Baartman, H. E. M. (1996). *Opvoeden kan zeer doen, over oorzaken van kindermishandeling [Childrearing can be painful, about the causes of child maltreatment]*. Utrecht: SWP.

Baird, C., & Wagner, D. (2000). The relative validity of actuarial and consensus based risk assessment systems. *Children and Youth Services Review, 22*, 839–871.

Barber, J. G., Shlonsky, A., Black, T., Goodman, D., & Trocmé, N. (2008). Reliability and predictive validity of a consensus-based risk assessment tool. *Journal of Public Child Welfare, 2*(2), 173–195.

Barlow, J., Fisher, J. D., & Jones, D. (2010). *Systematic review of models of analyzing significant harm. Department of Education.* Retrieved on November 11, 2013 from https://www.gov.uk/government/publications/systematic-review-of-models-of-analysing-significant-harm.

Belsky, J., & Vondra, J. (1989). Lessons from child abuse: the determinants of parenting. In D. Cicchetti, & V. Carlson (Eds.), *Child Maltreatment, theory and research on the causes and consequences of child abuse and neglect*. Cambridge: Cambridge University Press.

Bouwmeester-Landweer, M. B. R. (2006). *Early home visitation in families at risk for child Maltreatment (Doctoral dissertation)*. Leiden University Retrieved from https://openaccess.leidenuniv.nl/handle/1887/4396.

Brown, J., Cohen, P., Johnson, J. G., & Salzinger, S. (1998). A longitudinal analysis of risk factors for child maltreatment: Findings of a 17-year prospective study of officially recorded and self-reported child abuse and neglect. *Child Abuse & Neglect, 22*(11), 1065–1078.

Browne, K. D., Davies, C., & Stratton, P. (1988). *Early prediction and prevention of child abuse*. Chichester: John Wiley & Sons.

Burgmeijer, R. F. J., Van Geenhuizen, Y. M., Filedt kok-Weimar, T., & De Jager, A. M. (1997). *Op weg naar volwassenheid. Evaluatie jeugdgezondheidszorg 1996*. Leiden: Maarsen: TNO preventie en Gezondheid en KPMG.

Camasso, M., & Jagannathan, R. (2000). Modeling the reliability and predictive validity of risk assessment in child protective services. *Children and Youth Services Review, 22*, 873–896.

Cash, S. J. (2001). Risk assessment in child welfare: The art and science. *Children and Youth Services Review, 23*(11), 811–830.

Cowen, P. S. (1992). Abuse protection: child. In J. C. M. McCloskey, & G. M. Bulechek (Eds.), *Nursing Interventions; effective nursing treatment*. Philadelphia: Saunders.

D'Andrade, A., Benton, A., & Austin, M. J. (2005). *Risk and safety assessment in child welfare: Instrument comparisons*. Center for Social Services Research Retrieved on November 11, 2013 from http://cssr.berkeley.edu/bassc/public/risk_full.pdf.

Dawes, R. M., Faust, D., & Meehl, P. E. (1989). Clinical versus actuarial judgment. *Science, 243*(4899), 1668–1674.

De Ruiter, C., & De Jong, E. M. (2005). *CARE-NL. Richtlijn voor gestructureerde beoordeling van het risico van kindermishandeling*. Utrecht: Corine de Ruiter.

DeLong, E. R., DeLong, D. M., & Clarke-Pearson, D. L. (1988). Comparing the areas under two or more correlated receiver operating characteristic curves: a non-parametric approach. *Biometrics, 44*, 837–845.

DePanfilis, D., & Girvin, H. (2005). Investigating child maltreatment in out-of-home care: Barriers to good decision-making. *Children & Youth Services Review, 27*, 353–374.

Dean, J. G., MacQueen, I. A. G., Mitchell, R. G., & Kempe, C. H. (1978). Health visitor's role in prediction of early childhood injuries and failure to thrive. *Child Abuse & Neglect, 2*, 1–17.

Dombrowski, S. C., Timmer, S. G., Blacker, D. M., & Urquiza, A. J. (2005). A positive behavioural intervention for toddlers: parent-child attunement therapy. *Child Abuse Review, 14*(2), 132–151. http://dx.doi.org/10.1002/car.888.

Dorsey, S., Mustillo, S. A., Farmer, E. M. Z., & Elbogen, E. (2008). Caseworker assessments of risk for recurrent maltreatment: association with case-specific risk factors and re-reports. *Child Abuse & Neglect, 32*, 377–391.

Dunnink, G. (2010). *Standpunt Bereik van de jeugdgezondheidszorg*. Centrum Jeugdgezondheid RIVM-rapport 295001015/2010.

Egeland, B. (1988). Breaking the cycle of abuse: implications for prediction and intervention. In K. D. Browne, C. Davies, & P. Stratton (Eds.), *Early prediction and prevention of child abuse*. Chichester: John Wiley & Sons.

Eigsti, I.-M., & Cicchetti, D. (2004). The impact of child maltreatment on expressive syntax at 60 months. *Developmental Science, 7*(1), 88–102. http://dx.doi.org/10.1111/j.1467-7687.2004.00325.x.

English, D. J., Thompson, R., Graham, J. C., & Briggs, E. C. (2005). Toward a definition of neglect in young children. *Child Maltreatment, 10*(2), 190–206. http://dx.doi.org/10.1177/1077559505275178.

Finkelhor, D., Turner, H., Ormrod, R., & Hamby, S. L. (2009). Violence, abuse, and crime exposure in a national sample of children and youth. *Pediatrics, 124*(5), 1411–1423.

Gambrill, E., & Shlonsky, A. (2000). Risk assessment in context. *Children and Youth Services Review, 22*, 813–837.

Gershater-Molko, R. M., Lutzker, J. R., & Wesch, D. (2003). Project SafeCare: Improving health, safety, and parenting skills in families reported for, and at-risk for child maltreatment. *Journal of Family Violence, 18*(6), 377–386. http://dx.doi.org/10.1023/A:1026219920902.

Glaser, D. (2002). Emotional abuse and neglect (psychological maltreatment): A conceptual framework. *Child Abuse & Neglect, 26*, 697–714.

Grove, W. M., & Meehl, P. E. (1996). Comparative efficiency of informal (subjective, impressionistic) and formal (mechanical, algorithmic) prediction procedures: the clinical-statistical controversy. *Psychology, Public Policy and Law, 2*, 293–323.

Hanley, J. A., & McNeil, B. J. (1983). A method of comparing the areas under receiver operating characteristic curves derived from the same cases. *Radiology, 148*, 839–843.

Heinrich, R., & Braak van den, J. (2007). *De Delta veiligheidslijst. Gestructureerde risicotaxatie en management binnen de Deltamethode*. Amsterdam/Woerden: PI Research/Adviesbureau Van Montfoort.

Johnson, W. L. (2011). The validity and utility of the California Family Risk Assessment under practice conditions in the field: a prospective study. *Child Abuse & Neglect, 35*, 18–28.

Jolly, J. D. (1979). The paediatric nurse − key roles in diagnosis and parent support. *Child Abuse & Neglect, 3*, 803–807.

Knoke, D., & Trocmé, N. (2005). Reviewing the evidence on assessing risk for child abuse and neglect. *Brief Treatment and Crisis Intervention, 5*, 310–327.

Kotch, J. B., Browne, D. C., Ringwalt, C. L., Stewart, P. W., Ruina, E., Holt, K., et al. (1995). Risk of child abuse or neglect in a cohort of low-income children. *Child Abuse & Neglect, 19*, 1115–1130.

Kotelchuck, M. (1982). Child abuse and neglect, prediction and misclassification. In R. H. Starr (Ed.), *Child abuse prediction, policy implications*. Cambridge: Ballinger Publishing Company.

Leschied, A. W., Chiodo, D., Whitehead, P. C., Hurley, D., & Marshall, L. (2003). The empirical basis of risk assessment in child welfare: the accuracy of risk assessment and clinical judgment. *Child Welfare, 82*, 527–540.

Lyons, P., Doueck, H. J., & Wodarski, J. S. (1996). Risk assessment for child protective services: A review of the empirical literature on instrument performance. *Social Work Research, 20*, 143–155.

Munro, E. (1999). Common errors of reasoning in child protection work. *Child Abuse & Neglect, 23*, 745–758.

Newberger, C. M. (1980). The cognitive structure of parenthood; the development of a descriptive measure. In R. L. Selman, & R. Yando (Eds.), *Clinical-developmental psychology. New directions of child development: clinical developmental research, No. 7*. San Francisco: Jossey-Bass.

Nygren, P., Nelson, H. D., & Klein, J. (2004). Screening children for family violence: a review of the evidence for the US Preventive Services Task Force. *The Annals of Family Medicine, 2*(2), 161–169.

Oates, R. K., Tebbutt, J., Swanston, H., Lynch, D. L., & O'Toole, B. I. (1998). Prior childhood sexual abuse in mothers of sexually abused children. *Child Abuse & Neglect, 22*, 1113–1118.

Pfister, H., & Böhm, G. (2008). The multiplicity of emotions: A framework of emotional functions in decision making. *Judgment and Decision Making, 3*, 5–17.

Pringle, M. K. (1980). Towards the prediction of child abuse. In N. Frude (Ed.), *Psychological approaches to Child Abuse*. London: Batsford Academic and Educational Ltd.

Reynolds, A. J., Mathieson, L. C., & Topitzes, J. W. (2009). Do early childhood interventions prevent child maltreatment? *Child Maltreatment, 14*(2), 182–206. http://dx.doi.org/10.1177/1077559508326223.

Rice, M. E., & Harris, G. T. (2005). Comparing effect sizes in follow-up studies: ROC Area, Cohen's d, and r. *Law and Human Behavior, 29*(5), 615.

Saunders, T. R., & O'Brien, D. E. (1997). Child maltreatment: Case studies in assessment and intervention. *Psychotherapy, 34*(4), 498–508.

Scarborough, A. A., & McRae, J. S. (2008). Maltreated infants: Reported eligibility for part C and later school-age special education services. *Topics in Early Childhood Special Education, 28*(2), 75–89. http://dx.doi.org/10.1177/0271121408320349.

Schumacher, J. A., Smith Slep, A. M., & Heyman, R. E. (2001). Risk factors for child neglect. *Aggression and Violent Behaviour, 6*, 231–254.

Sidebotham, P., & Golding, J. (2001). Child maltreatment in the children of the nineties a longitudinal study of parental risk factors. *Child Abuse & Neglect, 25*, 1177–1200.

Sidebotham, P., & Heron, J. (2003). Child maltreatment in the children of the nineties: the role of the child. *Child Abuse & Neglect, 27*, 337–352.

Sidebotham, P., Heron, J., & Golding, J. (2002). Child maltreatment in the Children of the Nineties: deprivation, class: and social networks in a UK sample. *Child Abuse & Neglect, 26*, 1243–1259.

Straus, M. A., & Smith, C. (1993). Family patterns and primary prevention of family violence Trends in Health Care. *Law and Ethics, 8*(2), 17–25.

Straus, M. A. (1978). The social causes of interpersonal violence: the example of family violence Odyssee House Non-violence. In Toronto: Paper read at the APA.

Swets, J. A. (1986). Indices of discrimination or diagnostic accuracy: Their ROCs and implied models. *Psychological Bulletin, 99*, 100–117.

Tan, F. (2001). *Eindverslag Pilot-study project OKé [Final report Pilot Study project OKé]. Unpublished master's thesis*Leiden, The Netherlands: Leiden University Medical Center.

Ten Berge, I. J., & Eijgenraam, K. (2009). *Licht Instrument Risicotaxatie Kindermishandeling (LIRIK) [Check List of Child Safety (CLCS)]*. Utrecht: Nederlands Jeugdinstituut.

Van der Put, C. E., Hermanns, J., Van Rijn-van Gelderen, L., & Sondeijker, F. (2016a). Detection of unsafety in families with parental and/or child developmental problems at the start of family support. *BMC Psychiatry, 16*(1), 1.

Van der Put, C. E., Assink, M., & Stams, G. J. J. M. (2016b). Predicting Relapse of Problematic Child-Rearing Situations. *Children and Youth Services Review, 61*, 288–295.

Wald, M. S., & Woolverton, M. (1990). Risk assessment: The emperor's new clothes? *Child Welfare, 69*, 483–511.