UNIVERSITY OF AMSTERDAM

UvA-DARE (Digital Academic Repository)

Predicting relapse of problematic child-rearing situations

van der Put, C.E.; Assink, M.; Stams, G.J.J.M.

DOI
10.1016/j.childyouth.2016.01.002
Publication date
2016
Document Version
Final published version
Published in
Children and Youth Services Review

Link to publication

Citation for published version (APA):
van der Put, C. E., Assink, M., & Stams, G. J. J. M. (2016). Predicting relapse of problematic child-rearing situations. *Children and Youth Services Review*, *61*, 288-295. https://doi.org/10.1016/j.childyouth.2016.01.002

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

UvA-DARE is a service provided by the library of the University of Amsterdam (https://dare.uva.nl)

Download date:17 Jun 2024

Children and Youth Services Review 61 (2016) 288–295

Contents lists available at ScienceDirect

Children and Youth Services Review

journal homepage: www.elsevier.com/locate/childyouth

Predicting relapse of problematic child-rearing situations

Claudia E. van der Put *, Mark Assink, Geert Jan J.M. Stams

Research Institute of Child Development and Education, University of Amsterdam, The Netherlands

ARTICLE INFO

Article history:
Received 22 April 2015
Received in revised form 4 January 2016
Accepted 4 January 2016
Available online 7 January 2016

Keywords:
Risk assessment
Child maltreatment
Structured clinical judgment
Actuarial risk assessment
Child protection

ABSTRACT

Background: The development and evaluation of risk assessment instruments for child maltreatment is still in its infancy, both in the Netherlands and internationally. The aim of this study was to examine the predictive validity of a structured clinical judgement instrument – the Check List of Child Safety (CLCS) – that is widely used in the Netherlands. The second aim was to examine the predictive validity of a newly developed actuarial risk classification that is based on variables measured with the CLCS.
Method: The sample consisted of 3963 Dutch families who were under the supervision of the Dutch Child Welfare Agency (CWA) between 2011 and 2013 because of problematic child-rearing situations. Relapse was defined as restarting treatment by the CWA because of newly substantiated problematic child-rearing situations. The actuarial risk classification was developed by means of a CHAID analysis. The predictive validity of the CLCS and the actuarial risk classification were examined by calculating several performance indicators (sensitivity, specificity, false positives, false negatives and AUC values).
Results: The predictive validity of the CLCS was poor, with a non-significant AUC of .530, meaning that the CLCS performed no better than chance. The predictive validity of the actuarial risk classification was moderate, with a significant AUC of .630.
Conclusion: The actuarial risk classification not only outperformed the CLCS, but is also time-saving in practice since it comprises only variables that are significantly related to relapse.
© 2016 Elsevier Ltd. All rights reserved.

1. Introduction

Child protection workers have to make extremely difficult decisions about whether to intervene in order to safeguard a child's welfare, and how best to do so (Arad-Davidson & Benbenishty, 2008; Baird & Wagner, 2000; DePanfilis & Girvin, 2005; Dorsey, Mustillo, Farmer, & Elbogen, 2008; Pfister & Böhm, 2008). Identifying risks of maltreatment should guide these decisions. In recent years, there has been a shift in the field of child protection from largely unstructured clinical risk assessment to the widespread use of standardized risk assessment instruments (Price-Robertson & Bromfield, 2011). Despite this shift, the development and evaluation of risk assessment instruments in the field of child protection is still in its infancy. Risk assessment instruments are frequently implemented without proper empirical evaluation, both in the Netherlands (Ten Berge, 2008) and internationally (Knoke & Trocmé, 2005), and thus limited knowledge is available about their validity and effectiveness (Barlow, Fisher, & Jones, 2010). Therefore, the first aim of this study was to examine the predictive validity of the risk assessment instrument most widely used in the Netherlands, namely the Check List of Child Safety (CLCS; Ten Berge & Eijgenraam, 2009). The CLCS was implemented with little or no empirical validation and it is therefore important to find out how well this instrument actually performs.

The CLCS is based on clinical judgement, which is one of the major approaches to risk assessment in child protection, besides the actuarial approach. The main difference between the clinical and the actuarial approach is that in clinical approaches, conclusions are based on the judgement of a professional who combines and weighs information in a subjective manner, whereas in actuarial approaches, conclusions are based solely on empirically established relations between risk factors and child maltreatment (Dawes, Faust, & Meehl, 1989). Clinical instruments can be further divided into (a) unaided decision-making based on experience, knowledge and intuition (unstructured clinical judgement), (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 judgement; SCJ).

Clinical approaches are most common in child protection practice, both in the Netherlands and internationally, despite the fact that studies indicate that most clinical methods, including unstructured clinical judgement and consensus based methods, perform questionably in practice (Arad-Davidson & Benbenishty, 2008; Baird & Wagner, 2000; Barlow et al., 2010; Camasso & Jagannathan, 2000; D'Andrade, Benton, & Austin, 2005; DePanfilis & Girvin, 2005; Dorsey et al., 2008; Knoke & Trocmé, 2005; Lyons, Doueck, & Wodarski, 1996; Munro, 1999; Pfister & Böhm, 2008; Wald & Woolverton, 1990). Validation studies have

* Corresponding author at: University of Amsterdam, Department of Forensic Child and Youth Care Sciences, P.O. Box 94208, 1090 GE Amsterdam, The Netherlands.
E-mail address: C.E.vanderPut@uva.nl (C.E. van der Put).

http://dx.doi.org/10.1016/j.childyouth.2016.01.002
0190-7409/© 2016 Elsevier Ltd. All rights reserved.

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). This leads to inappropriate decisions, resulting in the overuse of out-of-home placements or in possibly repeated maltreatment (Baird & Wagner, 2000; Gambrill & Shlonsky, 2000).

It is therefore surprising that actuarial methods are used so little in child welfare. Actuarial methods remain controversial in child welfare, despite the fact that studies have consistently shown that most actuarial methods perform moderate, whereas most clinical instruments perform poor (e.g., Baird & Wagner, 2000; Barber et al., 2008; Camasso & Jagannathan, 2000; Chaffin & Valle, 2003; Chan, 2012; Coohey, Johnson, Renner, & Easton, 2013; D'Andrade et al., 2005; Dorsey et al., 2008; Johnson, 2011; Van der Put, Hermanns, Rijn-Van Gelderen, & Sondeijker, in press). Also in other disciplines, such as criminal justice and psychology, the evidence is clear that actuarial methods outperform clinical judgement in predicting outcomes of interest (Aegisdóttir et al., 2006; Meehl, 1954, 1986; Dawes et al., 1989; Grove & Meehl, 1996; Grove, Zald, Lebow, Snitz, & Nelson, 2000; Hanson & Morton-Bourgon, 2009). Within these fields, actuarial methods are therefore frequently used.

There are two main reasons why actuarial instruments generally outperform clinical methods. First, the mathematical features of actuarial methods ensure that only variables with predictive value are part of the instrument and that these variables are weighed in accordance with their independent contribution to the outcome of interest (Dawes et al., 1989). This illustrates why it is very difficult for professionals to accurately predict an outcome of interest using their clinical judgement, because they are not able to select the most important factors or to properly weigh the observed risk factors (Dawes, 1994; Dawes et al., 1989). Second, the reliability of actuarial instruments is higher and hence the prediction is more accurate, because risk factors are scored according to a fixed algorithm, meaning that professionals use the same objective scoring rules, whereas in clinical methods the scoring of risk factors is done in a subjective way (e.g. Dawes et al., 1989; Gambrill & Shlonsky, 2000).

Given the above, it is remarkable that the actuarial versus clinical prediction controversy still exists within child welfare. One explanation for its persistence might be the confusion about the precise purpose to be served by risk assessment instruments. Assessment in child protection includes two distinct purposes: (a) to predict future child maltreatment (*risk assessment*) in order to establish intervention urgency and intensity, and (b) to identify targets of interventions in order to contribute to individualized case planning (*needs assessment*). It is generally understood that actuarial methods perform better than clinical judgement in risk assessment, but this is not clear for needs assessment. Needs assessment instruments are most often developed through expert consensus and are usually not subjected to empirical validation (Schwalbe, 2008).

The controversy between actuarial and clinical prediction seems therefore mainly to relate to the needs assessment function of instruments and concerns the question which method is better at identifying targets of interventions. Shlonsky and Wagner (2005) assume that actuarial methods are not suitable for needs assessment because these instruments do not identify the full range of risk factors relevant to intervention planning. Most actuarial instruments currently in use in child welfare are brief instruments based on multivariate statistical techniques and consist mainly of static risk factors. These instruments are therefore not suitable for needs assessment. However, in the field of juvenile justice, several actuarial risk instruments have been developed that are suitable for both risk and needs assessment (Schwalbe, 2008). As Schwalbe (2008, p. 1461) stated: "Actuarial risk assessment instruments can be constructed with a broad array of dynamic risk factors that could ground clinical hypotheses and could identify targets of intervention to reduce risk".

In the Netherlands, no actuarial risk assessment instruments for child maltreatment are available. Because the use of actuarial methods by child protection services may be promising, the second aim of this study was to examine the predictive validity of a newly developed actuarial risk classification that is based on variables measured with the CLCS.

2. Method

2.1. Sample

For the present study, we used copies of the CLCS completed in the period January 2011–December 2013. These CLCSs were completed for families who were under the supervision of the Dutch Child Welfare Agency (CWA) because of problematic child-rearing situations, which means that there was a serious threat to the development of the child from an abusive or unsafe domestic situation and/or from parents or caregivers who were failing to provide the child's basic needs.

The CLCSs were retrieved from the system of the CWA. We then only selected the CLCSs of cases for which the treatment was ended before January 2014. Relapse was measured from treatment ending until October 2014, so that the period in which relapse was measured was at least 10 months for each case. This procedure resulted in a sample of 3963 families with at least one child aged 0–18 years (M age = 9.19, SD = 5.32). The families all lived in one of the following cities in the Amsterdam Metropolitan Area: Aalsmeer (2%), Amsterdam (53%), Amstelveen (6%), Haarlemmermeer (10%), Purmerend (9%), Uithoorn (2%), Zaanstad (9%) and other small cities (9%).

2.2. Instruments

2.2.1. Check List of Child Safety

The CLCS — which is known as LIRIK (Licht Instrument Risicotaxatie Kindveiligheid; Ten Berge & Eijgenraam, 2009) in the Netherlands – is a checklist in which professional judgement is central (i.e. it is an SCJ instrument). The CLCS helps the professional to evaluate a case systematically based on the available information. It is not a structured questionnaire and provides no formulae or criteria to form an objective judgement. This means that the final assessment of child safety is based on a subjective clinical judgement.

The CLCS is based both on scientific literature on safety indicators and on risk and protective factors for child maltreatment (Ten Berge & Eijgenraam, 2009). The CLCS was designed for professionals who are experts in the evaluation of (severe) parenting and developmental problems. Before professionals can use the CLCS, they must be briefed on its purpose, use, possibilities and limitations. Completing the CLCS takes about 10–15 min. A first study into the reliability of the CLCS showed that the interrater reliability was low to moderate (Veenhuizen, 2013).

The CLCS consists of two parts: (1) the identification of suspected child maltreatment at present (safety assessment) and (2) the assessment of the risk of child maltreatment in the near future (risk assessment).

2.2.1.1. Safety assessment. The safety assessment part involves assessing the child's current situation. This relates primarily to the substantiation of suspected child maltreatment. The CLCS helps the professional to systematically assess whether (a) parents act in a threatening manner or fail to act (see Table 1, items 1–5), (b) the child shows signs of developmental problems (including inadequate psychosocial functioning and injury; see Table 1, item 6), or (c) there are risk factors in the parent(s), the child, the family and/or the environment that are indicative of child maltreatment (see Table 1, items 7–9). Based on an assessment of the identified concerns and risk factors, the professional concludes whether or not there is actual child maltreatment.

Table 1
Point-biserial correlations between CLCS items and relapse.

Items		r_{pb}
1a	The safety of the child is threatened by the parents	
1a1	Serious child maltreatment	.01
1a2	Insufficient protection of the child	.05**
1a3	Insufficient basic care	.03+
1a4	Parent(s) unavailable	.02
1a5	Parents obstruct access to the child	.05**
1b	Safety threatened by child himself	.01
1c	Safety threatened by another family member	.04*
2	Concrete indications of	
2a	Physical or psychological violence	.03+
2b	Neglect	.04*
2c	Sexual abuse	.02
2d	Witness of domestic violence	.08***
3	Concerns about parenting and care	
3a	Protection and safety	.06***
3b	Basic care	.04*
3c	Emotional warmth (support)	.01
3d	Rules and boundaries	.04*
3e	Stimulation	.02
3f	Stability	.08***
4	Concerns about the child regarding:	
4a	Psychosocial functioning	.01
4b	Physical health	−.03
4c	Skills and cognitive development	.01
5	Risk factors: parents	
5a	Previously abused and/or neglected children	.01
5b	Inadequate pedagogical knowledge and/or skills	.01
5c	Problems in the parent–child interaction	.01
5d	Minimize/deny proven child maltreatment	.01
5e	Negative attitude towards the child	−.01
5f	Psychiatric problems	.06**
5 g	Addiction problems	.04+
5 h	Intellectual disability	.03+
5i	Physical availability	.04*
5j	Emotional availability	.02
5 k	Become a parent at a young age	.01
5 l	Parent(s) victim(s) of child maltreatment	.05*
5 m	Previously used violence against persons	.08***
5n	Problematic partner relationship	.07***
6	Risk factors: child	
6a	Young child (<5 years)	.07***
6b	Problems in history (e.g. premature)	.01
6c	(Serious) illness or disability	.02
6d	Behavioural and/or developmental problems	.01
6e	Difficult temperament	.01
6f	Unwanted child	.02
7	Risk factors: family and environment	
7a	Single parent family, step-family, large family	.03+
7b	Many conflicts/domestic violence	.09***
7c	Unstable, disorderly living	.01
7d	Material/financial problems (unemployment, housing)	.06**
7e	Major life events	.02
7f	Social isolation/social conflict	.02
8	Protective factors: parents	
8a	Sense of competence, capacity	−.01
8b	Positive self-image	.00
8c	Supporting partner	−.03*
8d	Can use their own childhood experiences	.03
8e	Positive childhood experiences	.00
8f	Possible to request support	.01
8 g	Emotional availability	.01
8 h	Flexibility	−.01
8i	Willing and able to change	.03*
9	Protective factors: child	
9a	Social skills	.01
9b	Positive self-image	−.01
9c	Above average intelligence	−.01
9d	Attractive appearance	.05**
9e	Good relationships with key adult(s)	.04*
9f	Ego resilience (stress resistance)	−.00
9 g	Willing and able to change	−.03+
10	Protective factors: family and environment	
10a	Support informal network	.02
10b	Support formal network	.04*
11	Conclusions risks	
11a	Risk at the moment (very large, large, real small)	.04*
11b	Risk in the near future with foreseeable changes (very large, large, real, small)	.03

+ $p < .10$.
* $p < .05$.
** $p < .01$.
*** $p < .01$.

2.2.1.2. Risk assessment. In the risk assessment part, the professional assesses the risk of child maltreatment in the near future based on the signs and risk factors identified in the safety assessment (Table 1, items 1–9). In addition, the professional also identifies protective factors that could neutralize the effect of the risk factors (Table 1, items 10–12). The risk assessment of the CLCS consists of two separate scores: (a) the overall risk assessment score applicable to the present child-rearing situation (current moment) and (b) the overall risk assessment score taking into account expected changes in the near future, such as the birth of a new child in the family or parents who are planning to separate.

In the present study, we examined the predictive validity of the individual items of the safety assessment, the individual items of the risk assessment and the overall risk assessment scores. We did not examine how well the current situation of the child (actual child maltreatment) was assessed by the safety assessment.

2.2.2. Outcome measure

Relapse was defined as the restarting of treatment by the CWA because of problematic child-rearing situations, meaning restarting (forced) care because of a serious threat to the development of the child from an abusive or unsafe domestic situation and/or from parents or caregivers who fail to provide the basic needs of the child. The period in which relapse was measured ranged from 10 to 24 months (M = 19 months) after the end of the first treatment trajectory. Relapse in the follow-up period was influenced by the outcome of the CLCS, because in high-risk cases children are most likely removed from their homes. For this reason, the outcome measure was corrected for periods in which the child was not present at home during the follow-up period. Relapse was quantified as a dichotomous variable (0 = no relapse, 1 = relapse).

2.3. Analyses

Point-biserial correlations (r_{pb}) were calculated to examine the strength of the associations between relapse and each individual item of the CLCS. Cronbach's alphas were calculated for the subscales of the CLCS to examine the internal consistency of the items of the CLCS. The predictive validity of the CLCS and the actuarial risk classification were assessed by calculating the area under the receiver operating characteristic curve (AUC) value. The AUC value indicates the probability that a randomly selected family that will relapse has a higher risk classification than a randomly selected family that will not relapse (Hanley & McNeil, 1982). An AUC value of 0.50 indicates that the instrument performs no better than chance. A value of 1.00 indicates a perfect positive prediction, a value of 0.00 a perfect negative prediction. AUC values of .639 and higher correspond with a medium effect size (d = .50) and AUC values of .714 and higher correspond with a large effect size (d = .80; Rice & Harris, 2005).

The actuarial risk classification was developed by means of a chi-squared automatic interaction detector (CHAID) analysis. CHAID is a decision tree classification method that groups cases into subsets of cases with different levels of risk of relapse on the basis of particular combinations of variables. This method focuses on interactions between variables rather than on main effects of variables in the dataset being

examined. Individual items of the CLCS that were significantly associated with relapse were included as input for the CHAID analysis (see Table 1). In addition, a variable named "total number of risk factors" was created by summing the individual CLCS items that were either significantly related to child maltreatment (see Table 1) or empirically well-supported in other studies. This variable comprised the sum of the following factors: insufficient protection of the child, parents obstruct access to the child, safety threatened by another family member, concrete indications of neglect, witness of domestic violence, concerns about protection and safety, concerns about basic care, concerns about rules and boundaries, concerns about stability, parent(s) who previously abused and/or neglected children, parent(s) with psychiatric problems, parent(s) with addiction problems, parent(s) with intellectual disability, physical availability of parent(s), parent(s) victim(s) of child maltreatment, parent(s) previously used violence against persons, problematic partner relationship, young child (<5 years), many conflicts/domestic violence, material/financing problems, and social isolation/social conflict.

The CHAID algorithm involves dividing the total group of subjects into a number of subgroups on the basis of the independent variables most strongly associated with a specific outcome (in this study, relapse). In the first step of the CHAID procedure, the total group of subjects was divided into a number of subgroups on the basis of the variable most strongly associated with relapse. In the second step, the groups were split again on the basis of the variable that was then most strongly associated with relapse. This procedure was repeated until no variables remained that had a significant association with relapse in the subgroups, or until the groups had reached a minimum size ($n = 45$ in the present study). CHAID is highly appropriate for gaining insight into family profiles with a high or a low risk of relapse, because it identifies groups of cases that share the same risk factors and thus the same risk of relapse (Steadman et al., 2000; Thomas & Leese, 2003). Another advantage of CHAID is that the results are presented graphically and are therefore easy to interpret.

To cross-validate the actuarial risk classification, we used the *k*-fold cross-validation method, choosing $k = 10$. This means that the original sample was randomly partitioned into 10 equally sized subsamples (folds) to validate the risk classification. Ten iterations of training and validation were subsequently performed such that within each iteration a different fold of the data was held out for validation, while the remaining nine folds were used for learning.

To examine the performance of the actuarial risk classification, the sensitivity, specificity, percentage of false positives and percentage of false negatives were examined at various cut-off scores. Sensitivity is the probability of a positive score on the CLCS for children who will actually be reported for child maltreatment, and specificity is the probability of a negative score on the CLCS for children who will not be reported for child maltreatment. A false negative is a negative score on the CLCS whereas in reality the child is at risk of child maltreatment, and a false positive is a positive score on the CLCS whereas in reality the child is not at risk of maltreatment.

The method of Hanley and McNeil (1982) was used to test whether the AUC values of the CLCS and the actuarial risk classification differed significantly.

2.4. Ethical approval

Formal institutional review board (IRB) approval was not required, as this study involved secondary data analysis on de-identified data, which does not pose a threat to the subjects and therefore does not necessitate IRB regulation. Accordingly, this study was ethically conducted based on the rules maintained by the Faculty Ethics Review Board (FMG–UvA) of the University of Amsterdam, the Netherlands.

3. Results

3.1. Relation between individual items of the CLCS and relapse

Table 1 shows the point-biserial correlations (r_{pb}) between relapse and the individual items of the CLCS. The overall relapse rate was 12.7% ($n = 507$). The r_{pb} values for small, medium and large effect sizes for a 12.7% base rate are .06, .16 and .26 respectively, calculated with conversion formulae (after Rosenthal, 1991; Swets, 1986) provided by Rice and Harris (2005).

Risk factors with a small effect on relapse were: the child being a witness of domestic violence, concerns of the professional about protection and safety, concerns of the professional about stability, psychiatric problems of parent(s), parent(s) previously used violence against persons, parent(s) has/have a problematic partner relationship, young child (<5 years), many conflicts/domestic violence, and material/financial problems. There were no risk factors with a medium or large effect on relapse. In addition, none of the protective factors were related to relapse, with all point-biserial correlations smaller than .06.

3.2. Predictive validity of the CLCS

Table 2 shows the Cronbach's alphas for the subscales of the CLCS. The following subscales showed alpha values greater than .70: concerns about parenting and care, risk factors parents, protective factors parents, and protective factors child.

The overall risk assessment score of the CLCS comprises (a) the overall risk assessment score applicable to the present child-rearing situation (current moment) and (b) the overall risk assessment score taking into account expected changes in the near future. Table 3 shows the point-biserial correlations (r_{pb}) and the AUC values predicting relapse for both overall risk assessment scores. The AUC value of the overall risk assessment score at the current moment was .530 (.500–.560) and the AUC value of the overall risk assessment score taking into account expected changes was .531 (.501–.561).

3.3. The development of an actuarial risk classification

A CHAID analysis was performed to develop an actuarial risk classification. The individual items significantly associated with relapse were included as independent variables (see Table 1), including a sum variable of the CLCS items present in each case (see Methods section). Based on chi-squared testing, the total group of cases was divided into nine risk groups. Fig. 1 presents the output of the CHAID analysis (decision tree) and Table 4 a description of the nine risk groups.

The actuarial risk classification (average risk of relapse $p = .13$) was based on a combination of six variables: (1) total number of risk factors, (2) parent(s) previously used violence against persons, (3) many conflicts/domestic violence in the family, (4) parent(s) with intellectual disability, (5) parent(s) victim(s) of child maltreatment, and (6) child younger than 1 year. These variables proved to be the strongest

Table 2
Cronbach's alphas for the subscales of the CLCS.

Item no.		*N* of items	α
1	The safety of the child is threatened by parents	5	.52
2	Concrete indications of child maltreatment	4	.41
3	Concerns about parenting and care	6	.78
4	Concerns about the child	3	.57
5	Risk factors parents	14	.72
6	Risk factors child	7	.23
7	Risk factors family and environment	6	.52
8	Protective factors parents	9	.83
9	Protective factors child	7	.76
10	Protective factors family and environment	2	.59

Table 3
AUC values and point-biserial correlations (r_{pb}) between the CLCS overall risk assessment scores and relapse.

	r_{pb}	AUC (95% C.I.)
Overall risk assessment score at the current moment	.037*	.530 (.500–.560)
Overall risk assessment score in case of expected changes	.029	.531 (.501–.561)

* $p < .05$.

predictors of relapse and made a unique contribution to the prediction of relapse.

First, the total group of children was divided into four subgroups based on the variable total number of risk factors. When no risk factors were present (11% of the sample), the risk was lowest ($p = .06$), and when there were seven or more risk factors present (24% of the sample), the risk was highest ($p = .19$). Second, the subgroups were split again based on the variables that were then most strongly related to relapse. This resulted in nine risk groups (the grey shaded terminal nodes) in which cases have similar scores on the variables and thus a similar risk of relapse.

The AUC value of the actuarial risk classification was .630 (95% CI [.604–.656]), significantly higher than the AUC value of the CLCS ($z = 5.17$, $p < .001$). Table 5 shows the sensitivity, specificity, false positives and false negatives at the different cut-off scores.

4. Discussion

The first aim of this study was to examine the predictive validity of the CLCS, which is a structured clinical judgement risk assessment instrument widely used in the Netherlands for predicting the risk of child maltreatment. The results showed that the CLCS performed no better than chance (AUC = .53). Previous validation studies also showed that many risk assessment instruments for child maltreatment, especially clinical methods, perform questionably (Arad-Davidson & Benbenishty, 2008; Baird & Wagner, 2000; Barlow et al., 2010; Camasso & Jagannathan, 2000; D'Andrade et al., 2005; DePanfilis & Girvin, 2005; Dorsey et al., 2008; Knoke & Trocmé, 2005; Lyons et al., 1996; Munro, 1999; Pfister & Böhm, 2008; Wald & Woolverton, 1990). Using the CLCS to assess the risk of relapse might lead to many inappropriate decisions, which may result in the overuse of out-of-home placements or in possibly repeated maltreatment.

Many of the factors that are part of the CLCS, including all child factors (e.g. concerns of a professional about the psychosocial functioning, skills and cognitive development of the child), were not significantly associated with relapse. Previous studies have also shown that child factors are less strongly associated with child maltreatment compared to parental or family factors (Cash, 2001; Stith et al., 2009). In addition, a number of risk factors in parents were not significantly related to relapse, such as problems in the parent–child interaction and having a negative attitude towards the child. Previous studies found that these factors were related to relapse (Stith et al., 2009).

A possible explanation for these conflicting results is that these factors are often measured using instruments that consist of multiple items, whereas in the CLCS they are measured by single items. Another explanation might be that the items are not always completed very thoroughly and reliably by child protection workers (see the discussion of the limitations of this study). The results also showed that the protective factors were not related to relapse. This is in line with other studies that found that children from high-risk families benefit less from the presence of protective factors (e.g. Luthar & Goldstein, 2004; Miller, Wasserman, Neugebauer, Gorman-Smith, & Kamboukos, 1999; Vanderbilt-Adriance & Shaw, 2008). This indicates that interventions aimed at increasing or strengthening protective factors in high-risk families do not necessarily lead to a decrease in relapse. Vanderbilt-Adriance and Shaw (2008) concluded in their review of studies on resilience that in high-risk families, there should be a focus on both strengthening protective factors and reducing

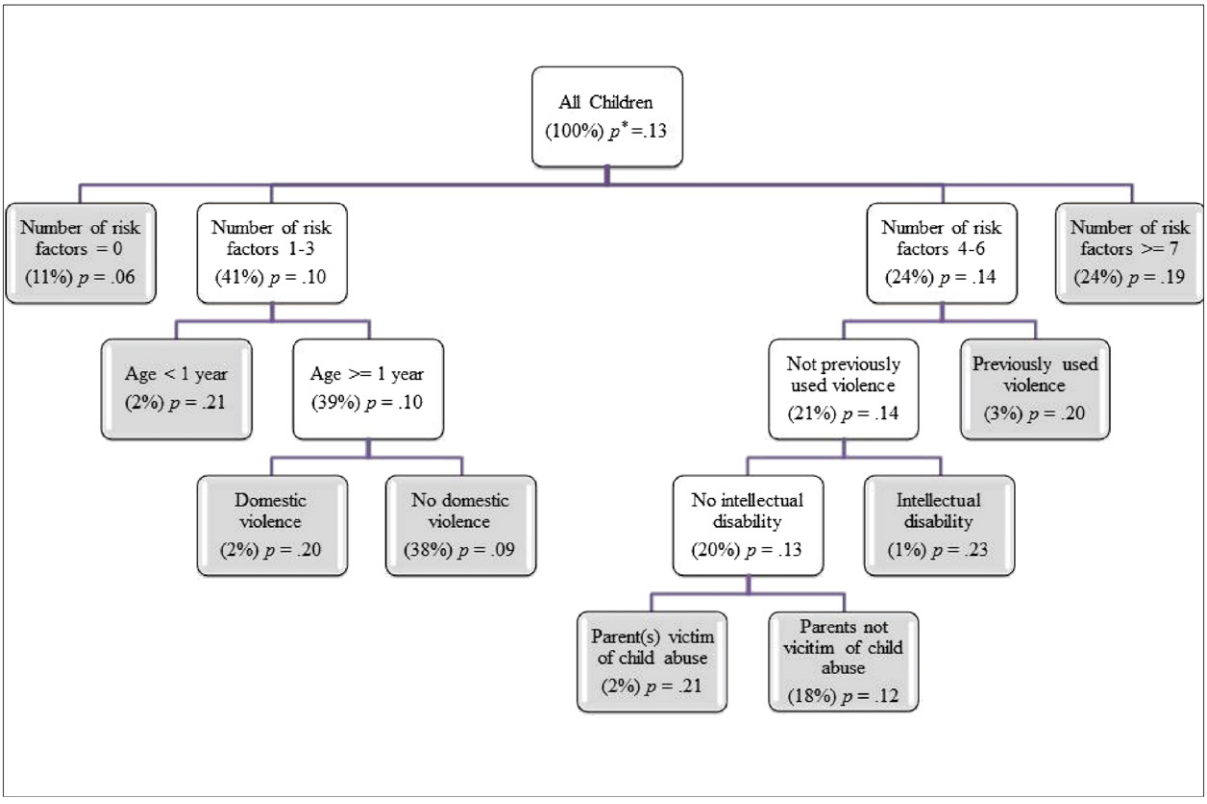

Fig. 1. CHAID output. *Notes.* *p = proportion of relapse. The grey shaded terminal nodes represent the 'risk groups' in which cases have similar scores on the variables and thus a similar risk of relapse.

Table 4
Risk classification CHAID analysis.

		%	p
1	Total number of risk factors is 4–6, parent(s) are intellectually disabled, parents have not previously used violence against persons	1.1%	.23
2	Total number of risk factors is 1–3, the child is younger than 1 year at start of the supervision	1.9%	.21
3	Total number of risk factors is 4–6, parents are themselves victim(s) of child maltreatment, parent(s) are not intellectually disabled, parents have not previously used violence against persons and parents	1.6%	.21
4	Total number of risk factors is 1–3, the child is 1 year or older at start of the supervision and there are concrete indications of recent domestic violence	1.5%	.20
5	Total number of risk factors is 4–6, parent(s) have previously used violence against persons	2.9%	.20
6	Total number of risk factors is 7 or more	23.5%	.19
7	Total number of risk factors is 4–6, parent(s) are not intellectually disabled, parents have not previously used violence against persons and parents are not themselves victim(s) of child maltreatment	18.8%	.12
8	Total number of risk factors is 1–3, the child is 1 year or older at start of the supervision and there are no concrete indications of recent domestic violence	37.9%	.09
9	Total number of risk factors is 0	10.8%	.06

Notes. % = the size of the risk group as a percentage of the total sample ($n = 3963$), p is the probability of relapse in the risk groups.

risks, because there are limits to the amount of risk that can be mitigated.

The second aim of this study was to examine the predictive validity of a newly developed actuarial risk classification that is based on variables measured with the CLCS. CHAID analysis showed that the following six variables were the strongest predictors of relapse and made a unique contribution to the prediction of relapse: total number of risk factors, parent(s) previously used violence against persons, many conflicts/domestic violence in the family, parent(s) with intellectual disability, parent(s) victim(s) of child maltreatment, and the child is younger than 1 year. The variable 'total number of risk factors' consisted of CLCS items that were significantly related to problematic child-rearing situations in the present study or were empirically well-supported in other studies.

The predictive validity of the actuarial risk classification was significantly better than that of the CLCS, with an AUC of. 630. This AUC value is slightly lower than the lower bound of a medium effect size (AUC = .639; Rice & Harris, 2005). The average AUC of risk assessment instruments for child maltreatment is still unknown because no meta-analyses have been conducted in this field. In the field of youth delinquency, however, several meta-analyses on the performance of risk assessment instruments have been conducted (Fazel, Singh, Doll, & Grann, 2012; Olver, Stochdale, & Wormith, 2009; Schwalbe, 2007). Schwalbe (2007) found a mean AUC of .64, Fazel et al. (2012) found a median AUC of .66 and Olver et al. (2009) found mean weighted correlations ranging from $r = .28$ to .32 for youth assessment instruments predicting general recidivism. The AUC value of the present actuarial risk classification is approximately equal to the mean AUC value of instruments used in forensic settings.

Because of the poor performance of the CLCS, the Child Welfare Agency in Amsterdam (and a number of other child welfare institutions) now uses the newly developed actuarial risk classification in assessing the risk of child maltreatment. In addition, a new safety assessment was developed to assess the safety at present (current situation). This newly developed safety assessment consists of eight items, such as 'there is (acute threat of) physical violence within the family', 'there is (acute threat of) sexual abuse or severe neglect', and 'there is no sufficiently available parent/caregiver'. The psychometric properties of this newly developed safety assessment will be examined in future research. The risk and safety assessment together form the Actuarial Risk assessment Instrument Youth Protection (ARIJ).

The ARIJ is used by the Child Welfare Agency (CWA) in Amsterdam in combination with a comprehensive needs assessment. If a child is referred to the CWA, the first step is to perform a safety assessment to determine the current safety of a child. The second step is to assess the risk of child maltreatment in the future using the CLCS. A comprehensive needs assessment is then performed to assess the dynamic risk factors related to child maltreatment that can be addressed in interventions. Based on the outcomes of the safety, risk and needs assessments, the family is referred to an intervention. The intensity of the intervention is based on the risk of child maltreatment: if the risk is high, the intervention should be comprehensive and far-reaching; if the risk is low, a low level of intervention is acceptable. When deciding on the precise cut-off point of the actuarial risk classification, both the number of false positives/false negatives and the consequences of a false positive/false negative must be taken into account.

The consequence of a false negative is that a family might not receive an intensive intervention, while the child is actually at risk of maltreatment, and the consequence of a false positive is that a family receives intensive intervention while in reality the child is not at risk. To protect children as well as possible, it is important to minimize the number of false negatives. However, it is also important to minimize the number of false positives, since false positives mean that children and their families receive unnecessarily intensive intervention. At the cut-off of .194, the percentage of false negative is 10.9% and the percentage of false positives is 7.2% (total percentage of erroneous decisions is 18.1%). At the cut-off of .157, the percentage of false negatives is lower (6.5%) but the percentage of false positives is higher (26.6%), as is the total percentage of erroneous decisions (32.7%).

Some limitations of the study should be mentioned. First, the CLCS was completed by child protection workers who differed in the degree to which they attach importance to completing the instrument as well as possible. The CLCSs were not always completed as thoroughly and reliably as possible by child protection workers, which had an effect on the quality of the data. Therefore, the observed effects might be an underestimation of the actual effects.

Second, only those copies of the CLCSs that could be retrieved from the computer system of the CWA were analysed. We could not use hardcopy CLCSs available in client files due to insufficient financial resources for digitizing these hardcopies during the current research project. Digital CLCFs were introduced at the CWA in 2011. From mid-2011, all CLCFs were filled out digitally at the CWA. However, in the first half of 2011, some professionals were not yet using the digital version. Therefore, some CLCFs completed in the first half of 2011 are missing from our sample. Nonetheless, it is expected that these missing CLCFs do not

Table 5
Sensitivity, specificity, false positives, fase negatives and total erroneous decisions at different cut-off scores of the Actuarial Risk Classification.

Cut-off score (>)[a]	Sensitivity	Specificity	False positives (%)	False negatives (%)	Total erroneous decisions (%)
.077	.947	.116	77.1	0.7	77.8
.107	.677	.510	42.7	4.1	46.8
.157	.495	.699	26.2	6.5	32.7
.194	.146	.918	7.2	10.9	18.1
.199	.101	.944	4.8	11.5	16.3
.203	.077	.958	3.6	11.8	15.4
.208	.051	.973	2.4	12.1	14.5
.219	.020	.990	0.9	12.5	13.4

[a] If a test score is greater than the cut-off value (i.e. the probability of relapse), the test result is considered positive; otherwise, it is considered negative.

differ systematically from those in our sample, and it is therefore expected that our sample is a good reflection of the total number of completed CLCSs in the Amsterdam region.

Third, we could only use CLCSs that had been completed in the Amsterdam region because limited financial resources prevented us from collecting data at the CWAs of other regions. Therefore, the results cannot automatically be generalized to other regions. Fourth, we did not measure the interrater reliability of the actuarial risk classification in our study, so further research on reliability is necessary. Fifth, we did not test the actuarial risk classification in a new sample, and it is therefore very important that the ecological validity (predictive validity in real life) be examined in future research.

Sixth, the outcome measure of the present study was the restarting of treatment by the CWA because of problematic child-rearing situations (relapse). Thus the predictive validity of the CLCS was examined for predicting relapse because of serious threat to the development of the child from an abusive or unsafe domestic situation and/or from parents or caregivers who fail to provide the basic needs of the child. The predictive validity of the CLCS should also be examined for predicting various forms of child abuse and neglect.

Finally, it is important to monitor the use of the actuarial risk classification in practice, because previous studies have shown that actuarial methods are not always used as intended (DePanfilis & Zuravin, 2001; Gillingham, 2011; Gillingham & Humphreys, 2010).

The results of this study are very important for child protection in the Netherlands. This is the first study in which the predictive validity of the CLCS was examined, and the results show that this widely used instrument performed no better than chance in predicting relapse, which may lead to many inappropriate decisions by child protection services. Our study found that the predictive validity of the newly developed actuarial risk classification was significantly better than that of the CLCS. This actuarial risk classification is also time-saving in practice, since it is considerably shorter than the CLCS because only variables that were significantly related to relapse were included. This is important, because actuarial methods (simple risk assessment procedures with known predictive attributes) are of great value in the context of increased demands placed on child protection services in an era of declining resources. In addition, shorter instruments motivate professionals to complete the instrument as well as possible.

Acknowledgements

The authors would like to thank Marc Dinkgreve, Inge Busschers and Mirte Forrer for their helpful comments and suggestions regarding the study. In addition, we would like to thank the Dutch Child Welfare Agency in Amsterdam (JBRA) for providing the opportunity to use the data.

References

Aegisdóttir, S., White, M. J., Spengler, P. M., Maugherman, A. S., Anderson, L. A., Cook, R. S., ... Rush, J. D. (2006). The meta-analysis of clinical judgment project: fifty-six years of accumulated research on clinical versus statistical prediction. *The Counseling Psychologist*, *34*(3), 341–382.

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.

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).

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.

Chaffin, M., & Valle, L. A. (2003). Dynamic prediction characteristics of the child abuse potential inventory. *Child Abuse & Neglect*, *27*, 463–481.

Chan, K. L. (2012). Evaluating the risk of child abuse: the child abuse risk assessment scale (CARAS). *Journal of Interpersonal Violence*, *27*, 951–973.

Coohey, C., Johnson, K., Renner, L. M., & Easton, S. D. (2013). Actuarial risk assessment in child protective services: construction methodology and performance criteria. *Children and Youth Services Review*, *35*, 151–161.

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. (1994). *House of cards: psychology and psychotherapy built on myth.* New York: Free Press.

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

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

DePanfilis, D., & Zuravin, S. (2001). Assessing risk to determine the need for services. *Children and Youth Services Review*, *23*, 3–20.

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.

Fazel, S., Singh, J. P., Doll, H., & Grann, M. (2012). Use of risk assessment instruments to predict violence and antisocial behavior in 73 samples involving 24827 people: systematic review and meta-analysis. *British Medical Journal.* http://dx.doi.org/10.1136/bmj.e4692 (published online).

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

Gillingham, P. (2011). Decision-making tools and the development of expertise in child protection practitioners: are we 'just breeding workers who are good at ticking boxes'? *Child & Family Social Work*, *16*(4), 412–421.

Gillingham, P., & Humphreys, C. (2010). Child protection practitioners and decision-making tools: observations and reflections from the front line. *British Journal of Social Work*, *40*(8), 2598–2616.

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.

Grove, W. M., Zald, D. H., Lebow, B. S., Snitz, B. E., & Nelson, C. (2000). Clinical versus mechanical prediction: a meta-analysis. *Psychological Assessment*, *12*(1), 19.

Hanley, J. A., & McNeil, B. J. (1982). The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology*, *143*, 29–36.

Hanson, R. K., & Morton-Bourgon, K. E. (2009). The accuracy of recidivism risk assessments for sexual offenders: a meta-analysis of 118 prediction studies. *Psychological Assessment*, *21*, 1–21.

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.

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

Luthar, S. S., & Goldstein, A. (2004). Children's exposure to community violence: implications for understanding risk and resilience. *Journal of Clinical Child and Adolescent Psychology*, *33*(3), 499–505.

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.

Meehl, P. E. (1954). *Clinical versus statistical prediction: A theoretical analysis and a review of the evidence.* Minneapolis: University of Minnesota Press.

Meehl, P. E. (1986). Causes and effects of my disturbing little book. *Journal of Personality Assessment*, *50*, 370–375.

Miller, L. S., Wasserman, G. A., Neugebauer, R., Gorman-Smith, D., & Kamboukos, D. (1999). Witnessed community violence and antisocial behavior in high-risk, urban boys. *Journal of Clinical Child Psychology*, *28*, 2–11.

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

Olver, M. E., Stochdale, K. C., & Wormith, J. S. (2009). Risk assessment with young offenders: a meta-analysis of three assessment measures. *Criminal Justice and Behavior*, *36*, 329–353.

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

Price-Robertson, R., & Bromfield, L. (2011). Risk assessment in child protection. *Resource sheet, National Child Protection Clearinghouse.* Australian Institute of Family Studies (Retrieved from < https://www3.aifs.gov.au/cfca/sites/default/files/publication-documents/rs24.pdf>.

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.

Rosenthal, R. (1991). *Meta-analytic procedures for social research.* Newbury Park, CA: Sage.

Schwalbe, C. S. (2007). Risk assessment for juvenile justice: a meta-analysis. *Law and Human Behavior*, *31*, 449–462.

Schwalbe, C. S. (2008). Strengthening the integration of actuarial risk assessment with clinical judgment in an evidence based practice framework. *Children and Youth Services Review*, *30*(12), 1458–1464.

Shlonsky, A., & Wagner, D. (2005). The next step: integrating actuarial risk assessment and clinical judgement into an evidence-based practice framework in CPS case management. *Children and Youth Services Review*, *27*, 409–417.

Steadman, H. J., Silver, E., Monahan, J., Appelbaum, P. S., Clark Robbins, P., Mulvey, E. P., ... Banks, S. (2000). A classification tree approach to the development of actuarial violence risk assessment tools. *Law and Human Behavior*, *24*(1), 83.

Stith, S. M., Liu, T., Davies, L. C., Boykin, E. L., Alder, M. C., Harris, J. M., ... Dees, J. E. M. E. G. (2009). Risk factors in child maltreatment: a meta-analytic review of the literature. *Aggression and Violent Behavior*, *14*(1), 13–29.

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

Ten Berge, I. J. (2008). *Instrumenten voor risicotaxatie in situaties van (vermoedelijke) kindermishandeling.* Utrecht: Nederlands Jeugdinstituut.

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

Thomas, S. D., & Leese, M. (2003). Invited editorial. A green-fingered approach can improve the clinical utility of violence risk assessment tools. *Criminal Behaviour and Mental Health*, *13*. (pp. 153–158). Whurr Publishers Ltd.

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

Vanderbilt-Adriance, E., & Shaw, D. S. (2008). Protective factors and the development of resilience in the context of neighborhood disadvantage. *Journal of Abnormal Child Psychology*, *36*(6), 887–901.

Veenhuizen, H. P. (2013). *Risicotaxatie kindermishandeling (LIRIK): Het effect van een hulpschema als modererende factor.* (Master's thesis) VU Amsterdam.

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