



## UvA-DARE (Digital Academic Repository)

### Het voorspellen van problematische opgroei- of opvoedingssituaties

*onderzoek naar de predictieve validiteit van het LIRIK en de ontwikkeling van een Actuarieel Risicotaxatie Instrument voor Jeugdbescherming (ARIJ)*

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

**DOI**

[10.1007/s12453-016-0117-4](https://doi.org/10.1007/s12453-016-0117-4)

**Publication date**

2016

**Document Version**

Final published version

**Published in**

Kind en Adolescent

[Link to publication](#)

**Citation for published version (APA):**

van der Put, C., Assink, M., & Stams, G. J. M. (2016). Het voorspellen van problematische opgroei- of opvoedingssituaties: onderzoek naar de predictieve validiteit van het LIRIK en de ontwikkeling van een Actuarieel Risicotaxatie Instrument voor Jeugdbescherming (ARIJ). *Kind en Adolescent*, 37(3), 133-154. <https://doi.org/10.1007/s12453-016-0117-4>

**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>)*

## Het voorspellen van problematische opgroei- of opvoedingssituaties

Onderzoek naar de predictieve validiteit van het LIRIK en de ontwikkeling van een Actuarieel Risicotaxatie Instrument voor Jeugdbescherming (ARIJ)

Claudia van der Put · Mark Assink · Geert Jan M. Stams

### Samenvatting

Het doel van deze studie was te onderzoeken of terugkeer naar de jeugdbescherming beter voorspeld wordt door het Licht Instrument Risicotaxatie Kindveiligheid (LIRIK) of door een nieuw te ontwikkelen actuarieel instrument. De steekproef bestond uit 3963 kinderen tussen 0 en 18 jaar ( $M_{leeftijd} = 9,1$  jaar,  $SD = 5,33$ ) en hun ouder(s) die in een hulpverleningstraject bij Jeugdbescherming Regio Amsterdam (JBRA) zaten vanwege een problematische opgroei- of opvoedingssituatie. Terugkeer was gedefinieerd als het opnieuw starten van een jeugdbeschermingstraject door JBRA. Het Actuarieel Risicotaxatie Instrument voor Jeugdbescherming (ARIJ) werd ontwikkeld middels een CHAID-analyse. Een voorspelling op basis van het LIRIK bleek niet significant beter dan toeval ( $AUC = .53$ ). De predictieve validiteit van het ARIJ was matig ( $AUC = .63$ ), maar significant beter dan die van het LIRIK. Veel praktijkinstellingen werken inmiddels met het ARIJ en daarom is het belangrijk te onderzoeken hoe de predictieve validiteit van het ARIJ verder kan worden verbeterd.

### Trefwoorden

risicotaxatie-instrument · predictieve validiteit · LIRIK · problematische opgroei- of opvoedingssituaties · kindermishandeling · ARIJ · jeugdbescherming

*Dit artikel is een bewerking van:* Put, C. E. van der, Assink, M., & Stams, G. J. J. M (2016). Predicting relapse of problematic child-rearing situations. *Children and Youth Services Review* 61, 288–295.

Dr. C. van der Put (✉) · Drs. M. Assink · Prof. dr. G.J.M. Stams  
afdeling Forensische Orthopedagogiek, Universiteit van Amsterdam, Nieuwe  
Achtergracht 127, 1018 WS Amsterdam, Nederland  
e-mail: c.e.vanderput@uva.nl

## Predicting problematic child-rearing situations

A study on the predictive validity of the light instrument for risk assessment of child maltreatment (LIRIK) and the development of an actuarial risk assessment instrument for youth protection (ARIJ)

### Abstract

The aim of this study was to examine whether return to youth protection is better predicted by using the light instrument for risk assessment of child maltreatment (Licht Instrument Risicotaxatie Kindveiligheid, LIRIK) or by using a newly-developed actuarial instrument. The sample consisted of 3963 Dutch children aged between 0 and 18 years ( $M_{\text{age}} = 9.12$  years,  $SD = 5.35$ ) and their parent(s) who were under the supervision of a Dutch child welfare agency (Jeugdbescherming Regio Amsterdam, JBRA) because of problematic child-rearing situations. Relapse was defined as resuming treatment by the JBRA as a result of newly-substantiated problematic child-rearing situations. The actuarial risk assessment instrument for youth protection (Actuarieel Risicotaxatie Instrument voor Jeugdbescherming, ARIJ) was developed by means of a CHAID analysis. The LIRIK performed no better than chance ( $AUC = 0.53$ ). The predictive validity of the ARIJ was moderate ( $AUC = 0.63$ ), but performed significantly better than the LIRIK. Further research should focus on how to improve the predictive validity of the ARIJ, particularly since the ARIJ has been implemented by multiple youth care organizations.

### Keywords

Risk assessment instrument · Predictive validity · LIRIK · Problematic child-rearing situations · Child maltreatment · ARIJ · Youth protection

## Inleiding

Professionals in de jeugdbescherming moeten vaak zeer moeilijke beslissingen nemen over de vraag of, en hoe, moet worden ingegrepen om kinderen te beschermen (Arad-Davidson en Benbenishty 2008; Baird en Wagner 2000; DePanfilis en Girvin 2005; Dorsey et al. 2008; Pfister en Böhm 2008). Het inschatten van het risico op kindermishandeling is hierbij cruciaal. In de afgelopen jaren heeft een verschuiving plaatsgevonden van het ongestructureerd klinisch inschatten van het risico op kindermishandeling naar het wijdverbreide gebruik van gestructureerde risicotaxatie-instrumenten (Price-Robertson en Bromfield 2011). Ondanks deze verschuiving staat de ontwikkeling en validatie van risicotaxatie-instrumenten voor kindermishandeling nog in de kinderschoenen. Risicotaxatie-instrumenten worden vaak grootschalig geïmplementeerd zonder eerst empirisch onderzoek uit te voeren, en dat geldt zowel in Nederland (Berge 2008) als internationaal (Knoke en Trocmé 2005). In Nederland wordt grootschalig gebruik gemaakt van het Licht Instrument Risicotaxatie Kindveiligheid (LIRIK; Berge en Eijgenraam 2009). Om-

dat er bij de start van deze studie nog geen validatieonderzoek naar het LIRIK was uitgevoerd, was het eerste doel van deze studie te onderzoeken hoe goed terugkeer naar de jeugdbescherming kan worden voorspeld door het LIRIK. Een tweede doel was te onderzoeken of terugkeer naar de jeugdbescherming beter voorspeld wordt door een nieuw te ontwikkelen actuarieel instrument dan door het LIRIK.

In de literatuur worden twee belangrijke benaderingen voor risicotaxatie onderscheiden: instrumenten gebaseerd op het klinisch oordeel en actuariële instrumenten.

Het belangrijkste verschil tussen klinische en actuariële instrumenten is dat bij klinische instrumenten de professional zelf bepaalt hoe zwaar hij of zij risicofactoren weegt en combineert tot een eindoordeel, terwijl bij actuariële instrumenten risicofactoren op een vaststaande manier worden gewogen en gecombineerd tot een eindoordeel (Dawes et al. 1989). Klinische methoden kunnen verder worden verdeeld in: (a) het ongestructureerd klinisch oordeel (inschatting op basis van kennis, ervaring en expertise van de professional), (b) *consensus-based* instrumenten (instrumenten gebaseerd op meningen van experts, maar zonder wetenschappelijke basis) en (c) het gestructureerd klinisch oordeel (instrumenten met wetenschappelijke basis, maar waarin de uiteindelijke beslissing aan de professional wordt overgelaten). Het LIRIK behoort tot deze derde categorie.

Klinische methoden worden het meest toegepast in de praktijk van de jeugdbescherming, zowel in Nederland als internationaal. Dit is opmerkelijk omdat onderzoeken laten zien dat klinische instrumenten voor kindermishandeling (die vaak breed worden ingezet) matig presteren en soms zelfs niet beter dan 'toeval' (bijv. Baird en Wagner 2000; Barber et al. 2008; Camasso en Jagannathan 2000; D'Andrade et al. 2005; Dorsey et al. 2008). Hierbij dient te worden opgemerkt dat in validatieonderzoek naar risicotaxatie-instrumenten voor kindermishandeling vaak minimaal wordt gerapporteerd over prestatie-indicatoren, zoals sensitiviteit, specificiteit of AUC-waardes. Over het algemeen wordt alleen gerapporteerd over de samenhang tussen de risico-inschatting en de uitkomstmaat, waarbij verschillen worden onderzocht tussen de risicogroepen in de mate waarin kindermishandeling voorkomt.

Voorbeelden van onderzoeken waaruit blijkt dat klinische instrumenten niet beter voorspellen dan toeval zijn onderzoeken naar het Ontario Risk Assessment Model (ORAM) en de Washington State Risk Assessment Matrix (WARM; Baird en Wagner 2000; Barber et al. 2008). Met de ORAM worden 22 risicofactoren gemeten, die betrekking hebben op de opvoeder, het kind of het gezin. Professionals worden gevraagd de risicofactoren te scoren op een 5-puntsschaal en daarna een algemene inschatting te geven op basis van hun klinisch oordeel (eveneens op een 5-puntsschaal). Onderzoek van Barber en collega's (2008) laat zien dat deze klinische inschatting niet samenhangt met latere meldingen van kindermishandeling: er was geen verschil in het aantal meldingen van kindermishandeling tussen de groepen die werden ingeschat als laag risico en de groepen die werden ingeschat als hoog risico. De WARM bestaat uit 37 items verdeeld over 7 domeinen. Deze items worden gescoord op een 5-puntsschaal waarna op basis van deze items een klinische inschatting wordt gemaakt van het risiconiveau. Baird en Wagner (2000)

lieten zien dat ook voor de WARM geldt dat er geen significant verschil bestaat tussen de groepen met een laag, gemiddeld en hoog risico, in de mate waarin in de toekomst daadwerkelijk kindermishandeling voorkomt. Het gebruik van dergelijke instrumenten leidt tot veel verkeerde beslissingen, hetgeen kan resulteren in onterechte uithuisplaatsingen of (herhaalde) kindermishandeling.

In de praktijk van de jeugdbescherming wordt nog weinig gebruikgemaakt van actuariële risicotaxatie-instrumenten, terwijl onderzoek laat zien dat het risico op kindermishandeling beter kan worden ingeschat met actuariële instrumenten dan met klinische instrumenten (Baird en Wagner 2000; D'Andrade et al. 2005; Dawes et al. 1989; Grove en Meehl 1996; Leschied et al. 2003). Ook in andere vakgebieden, zoals de psychologie en de criminologie, is overweldigend empirisch bewijs dat actuariële instrumenten beter presteren dan klinische instrumenten in het voorspellen van diverse uitkomsten (zie bijv. Aegisdóttir et al. 2006; Baird en Wagner 2000; D'Andrade et al. 2005; Dawes et al. 1989; Grove en Meehl 1996; Grove et al. 2000; Hanson en Morton-Bourgon 2009; Hilton et al. 2006; Kaufmann en Wittmann 2016; Leschied et al. 2003; Meehl 1954, 1986). In deze vakgebieden speelt de discussie tussen klinische versus actuariële instrumenten dan ook niet meer en zijn actuariële instrumenten over het algemeen de norm.

Er zijn twee belangrijke verklaringen voor het feit dat actuariële instrumenten beter presteren dan klinische instrumenten. Ten eerste bevatten actuariële instrumenten alleen risicofactoren die daadwerkelijk een voorspellende waarde hebben. Bovendien is statistisch bepaald hoe deze factoren moeten worden gewogen, gebaseerd op de unieke bijdrage van deze factoren aan de voorspelling van de betreffende uitkomstmaat, zodat sprake is van een optimaal scoringssysteem (Dawes et al. 1989). Dit illustreert waarom het zo moeilijk is voor professionals om een juiste inschatting te maken gebaseerd op hun klinisch oordeel. Het is namelijk niet alleen heel lastig om de belangrijkste factoren te selecteren en op de juiste manier te wegen, maar ook om rekening te houden met complexe interacties tussen factoren (Dawes 1994; Dawes et al. 1989). Ten tweede zijn actuariële instrumenten betrouwbaarder dan klinische instrumenten en dat maakt de predictieve validiteit beter. Deze hogere betrouwbaarheid komt doordat professionals allemaal dezelfde objectieve scoringsregels toepassen, onafhankelijk van hun ervaring, terwijl bij klinische instrumenten factoren op een subjectieve manier worden gescoord (bijv. Dawes et al. 1989; Gambrill en Shlonsky 2000).

In Nederland zijn geen actuariële instrumenten beschikbaar waarmee het risico op kindermishandeling kan worden ingeschat. Omdat het gebruik van actuariële instrumenten veelbelovend is, was het doel van deze studie om een actuariel risicotaxatie-instrument te ontwikkelen, gebaseerd op variabelen gemeten met het LIRIK. Dit instrument wordt verder aangeduid als het Actuariel Risicotaxatie Instrument voor Jeugdbescherming (ARIJ). Twee belangrijke verschillen tussen het ARIJ en het LIRIK zijn: (1) het ARIJ bevat alleen risicofactoren die daadwerkelijk een voorspellende waarde hebben, en (2) de factoren in het ARIJ worden op een vaststaande (empirisch bepaalde) manier gewogen en gecombineerd tot een risico-inschatting, terwijl met het LIRIK een professional zelf bepaalt hoe zwaar hij factoren weegt en combineert tot een eindoordeel. Onderzocht werd wat de

predictieve validiteit van het LIRIK en het ARIJ is en in hoeverre deze significant van elkaar verschillen. In dit onderzoek stonden de volgende onderzoeksvragen centraal:

1. Wat is de sterkte van de samenhang van individuele risicofactoren zoals gemeten met het LIRIK en terugkeer naar de jeugdbescherming?
2. Wat is de predictieve validiteit van het LIRIK in de voorspelling van terugkeer naar de jeugdbescherming?
3. Wat is de predictieve validiteit van het nieuw te ontwikkelen actuariële risicotaxatie-instrument (ARIJ) in de voorspelling van terugkeer naar de jeugdbescherming?
4. In hoeverre zijn er verschillen in de predictieve validiteit van het LIRIK en het ARIJ?

---

## Methode

### Onderzoeksgroep

Voor het onderzoek is gebruikgemaakt van exemplaren van het LIRIK die zijn ingevuld door professionals van Jeugdbescherming Regio Amsterdam (JBRA) aan het begin van een hulpverleningstraject voor gezinnen waarbij sprake was van een problematische opgroei- of opvoedingssituatie. Dit betekent dat sprake was van een serieuze bedreiging van de ontwikkeling van het kind door ofwel een onveilige thuissituatie ofwel doordat de ouders niet in staat waren te voorzien in de basisbehoeftes van een kind.

De onderzoeksgroep is als volgt geselecteerd: eerst zijn alle LIRIK's geselecteerd die van 2011 tot en met 2013 ingevoerd waren in het IJ (Informatie Jeugdzorg)-systeem, wat het primaire ICT-systeem van JBRA was waarin informatie van kinderen en hun gezinnen werd opgeslagen, waaronder gezinsrapportages, behandelinformatie en ingevulde LIRIK's. Daarna zijn LIRIK's geselecteerd van casussen waarbij het hulpverleningstraject was afgesloten voor 31 dec. 2013, zodat voor elke casus een terugkeerperiode beschikbaar was van minimaal tien maanden. Vervolgens zijn voogdijzaken uit de steekproef gehaald (dit betrof 1 % van de steekproef) evenals LIRIK's waarbij sprake was van een uithuisplaatsing in de terugkeerperiode. De totale steekproef die overbleef bestond uit  $N = 3963$  kinderen; 2185 jongens (55,5 % van de steekproef) en 1751 meisjes (44,5 % van de steekproef). De leeftijd bij aanmelding van de kinderen varieerde van 0 tot 18, met een gemiddelde leeftijd van 9,1 jaar ( $SD = 5,33$ ). 79,1 % van de kinderen was geboren in Nederland en 20,9 % in het buitenland. Gemiddeld hadden deze kinderen 1,56 ouders/opvoeders ( $N = 6133$  ouders, 63 % moeders, 37 % vaders). Van de moeders was 39 % geboren in Nederland en van de vaders 33 %.

## Ethische toetsing

Formele ethische toetsing was niet vereist omdat: (a) in deze studie gegevens geheel anoniem werden verkregen, geanalyseerd en gerapporteerd, (b) de onderzoeksresultaten alleen betrekking hebben op groepen en niet herleidbaar zijn naar individuen, en (c) bij de start van de hulpverlening bij JBRA aan cliënten wordt gemeld dat hun gegevens worden bewaard en eventueel gebruikt voor onderzoek wanneer zij hiermee instemmen. Hierdoor voldoet het onderzoek aan de regels van de ethische commissie van UvA-FMG.

## Instrumenten

### Licht Instrument Risicotaxatie Kindveiligheid (LIRIK)

Het LIRIK (Berge en Eijgenraam 2009) is in 2007 in het kader van het traject Verbetering Indicatiestelling Bureaus Jeugdzorg ontwikkeld door het Nederlands Jeugdinstituut, in samenwerking met enkele bureaus jeugdzorg. Met behulp van het LIRIK moet de professional snel kunnen nagaan of er mogelijk sprake is van kindermishandeling en wat het risico daarop is in de toekomst. Het LIRIK is gebaseerd op gestructureerde professionele oordeelsvorming. Dit wil zeggen dat het instrument de professional helpt om op grond van de beschikbare informatie systematisch de voorliggende casus te beoordelen (Berge en Eijgenraam 2009). Het LIRIK biedt op gestructureerde wijze relevante aandachtspunten om tot een expliciet en weloverwogen oordeel ten aanzien van de veiligheid van het kind te komen. Het professionele oordeel van de hulpverlener staat daarbij centraal. Het LIRIK is dus niet gebaseerd op een algoritme om tot een objectief oordeel te komen, maar laat de hulpverlener zelf tot een eindoordeel komen door de zorgen, problemen en sterke kanten tegen elkaar af te wegen.

De met het LIRIK gemeten variabelen zijn gebaseerd op resultaten van literatuuronderzoek naar risico- en beschermende factoren voor kindermishandeling en Nederlandstalige literatuur over veiligheidsindicatoren (Ten Berge, 2008). Het LIRIK is primair bedoeld voor professionals in de jeugdzorg die op basis van hun opleiding en werkervaring deskundig zijn in het beoordelen van (ernstige) opvoedings- en ontwikkelingsproblematiek. Voorafgaand aan het gebruik is een korte training vereist (circa 2 uur) in het doel, het gebruik, de mogelijkheden en de beperkingen van het instrument. Het invullen van het LIRIK kost naar schatting 10 tot 15 minuten (Berge 2008).

Het LIRIK bestaat uit twee onderdelen: (1) het vaststellen van de kindveiligheid (veiligheidstaxatie) en (2) het maken van een inschatting van het risico op kindermishandeling in de nabije toekomst (risicotaxatie).

*Veiligheidstaxatie:* In deel 1 van het LIRIK staat de beoordeling van de huidige situatie van het kind centraal. Dit onderdeel bevat circa 23 items die met 'ja'/'nee'/'onbekend' of 'zorgen'/'geen zorgen'/'onbekend' worden gescoord. Het

gaat hierbij primair om het onderkennen van een vermoeden van kindermishandeling. Het LIRIK helpt de professional om systematisch te beoordelen of er sprake is van (a) bedreigend handelen of niet handelen van de ouder(s), (b) kindsignalen die wijzen op een bedreigde ontwikkeling (o. a. onvoldoende psychosociaal functioneren, letsel) of (c) risicofactoren bij ouder(s), jeugdige, gezin en/of omgeving die een indicatie zijn voor kindermishandeling. Voorbeelden van items zijn: ‘Zijn er concrete aanwijzingen voor lichamelijk of psychisch geweld?’, ‘Zijn er zorgen over opvoeding en verzorging wat betreft regels en grenzen stellen?’, ‘Vertoont de jeugdige internaliserende problemen?’, ‘Heeft de opvoeder eerder een kind mishandeld, verwaarloosd, misbruikt?’, ‘Is er sprake van veel conflicten of huiselijk geweld?’. Op grond van een afweging van de gesignaleerde zorgen en risicofactoren concludeert de professional of er sprake is van mogelijke of feitelijke kindermishandeling of dat er geen aanwijzingen zijn voor kindermishandeling.

*Risicotaxatie:* In deel 2 van het LIRIK taxeert de professional wat voor het betreffende kind het risico is op kindermishandeling in de toekomst. Dit wordt gedaan op basis van de signalen en risicofactoren die in deel 1 geconstateerd zijn. De professional wordt gevraagd om aan te geven wat de aard van de risico’s is, op welke termijn deze zich kunnen voordoen, hoe groot de kans daarop is en hoe ernstig dan de gevolgen voor het kind zijn. Daarnaast geeft de professional aan of er beschermende factoren zijn die de gesignaleerde risico’s zouden kunnen compenseren, ofwel het effect van deze risico’s zouden kunnen neutraliseren. In totaal bevat dit onderdeel 25 aanvullende items, waarvan 18 beschermende factoren. Voorbeelden van beschermende factoren bij de ouders zijn: gevoel van competentie, draagkracht, positief zelfbeeld, ondersteunende partner. Voorbeelden van beschermende factoren bij het kind en in de omgeving zijn: sociaal vaardig, aantrekkelijk uiterlijk, positief zelfbeeld, steun (in-) formeel netwerk. Ten slotte volgt een beoordeling van het algemene risico voor het kind. Dit algemene risico wordt onderverdeeld in: (a) risico’s voor de jeugdige op dit moment en (b) risico’s voor de jeugdige bij voorziene veranderingen in de nabije toekomst. Deze risicoschatting wordt op een 4-puntsschaal weergegeven, namelijk het risico voor de jeugdige is ‘zeer ernstig’, ‘ernstig’, ‘matig ernstig’ of ‘gering’.

Een eerste onderzoek naar de betrouwbaarheid en predictieve validiteit van het LIRIK laat zien dat deze laag zijn (Bartelink et al. 2015). In het huidige onderzoek werd de predictieve validiteit onderzocht van de individuele items uit het veiligheidstaxatie-onderdeel, de individuele items uit het risicotaxatie-onderdeel en de algemene risicotaxatie-score (d.w.z., de klinische inschatting van het risico). Samenvattend werd dus onderzocht hoe goed kindermishandeling *in de toekomst* wordt voorspeld met het LIRIK. Hoe goed de *huidige* veiligheidssituatie in kaart wordt gebracht met het LIRIK werd in deze studie niet onderzocht.



## Actuarieel Risicotaxatie Instrument Jeugdbescherming (ARIJ)

Het ARIJ bestaat uit vier onderdelen: (a) een veiligheidstaxatie, (b) een risicotaxatie, (c) een dynamisch risicoprofiel en (d) een aantal experimentele items. Deze onderdelen worden achtereenvolgens besproken.

Met de veiligheidstaxatie wordt de huidige onveiligheid van het kind vastgesteld. Hierbij gaat het om de vraag of onmiddellijk moet worden ingegrepen om het kind veilig te stellen. Dit onderdeel bestaat uit 8 items. Voorbeelden van items van de veiligheidstaxatie zijn ‘er is sprake van (acute dreiging van) fysiek geweld’, ‘er is sprake van (acute dreiging van) seksueel misbruik binnen het gezin, door derden, in de vorm van kinderprostitutie (o. a. loverboy) of anderszins’, en ‘er is geen opvoeder die voldoende beschikbaar is’. Indien één van de items met ‘ja’ wordt beantwoord, dan geeft het instrument aan dat het kind direct moet worden veiliggesteld. Indien ten minste één van de items met ‘onbekend’ wordt beantwoord, dan geeft het instrument aan dat de veiligheid nader in kaart moet worden gebracht. In andere gevallen geeft het instrument aan dat er geen zorgen zijn over de huidige veiligheid. In geval het kind meteen moet worden veiliggesteld, wordt direct een plan opgesteld ten aanzien van de in te zetten middelen, zoals het inschakelen van het sociaal netwerk, de veroorzaker van de onveiligheid uit huis halen of (diverse vormen van) uithuisplaatsing van het kind.

Met de risicotaxatie wordt het risico op toekomstige onveiligheid ingeschat. De ontwikkeling van dit onderdeel wordt in de resultatenparagraaf van dit artikel beschreven.

Met het dynamisch risicoprofiel kan vooruitgang worden gemonitord en wordt aangegeven in hoeverre en hoe het risico op toekomstige onveiligheid kan worden verlaagd. Dit onderdeel is alleen gebaseerd op dynamische (veranderbare) risicofactoren. Tot slot is een aantal experimentele items aan het ARIJ toegevoegd voor onderzoeksdoeleinden die nog niet meewegen in de risico-inschatting, maar waarvan uit verschillende reviews is gebleken dat deze sterk samenhangen met toekomstige onveiligheid. Voorbeelden van deze experimentele items zijn: ‘opvoeder(s) reguleren eigen boosheid onvoldoende’, ‘opvoeder(s) beschouwen de gesignaleerde opvoedproblemen als minder ernstig dan de hulpverlener’, ‘opvoeder(s) zijn onvoldoende bereid/in staat tot verandering van hun (opvoed-)gedrag’ en ‘opvoeders ervaren het kind als het probleem’. In toekomstig onderzoek kan worden onderzocht of het meewegen van deze items leidt tot een verbetering van de predictieve validiteit van het ARIJ.

Een eerste onderzoek naar de betrouwbaarheid van het ARIJ laat zien dat deze voldoende tot goed is (Busschers et al. 2015). In het huidige onderzoek werd de predictieve validiteit onderzocht van het risicotaxatie-onderdeel. Onderzocht werd dus hoe goed kindermishandeling *in de toekomst* kan worden voorspeld met het ARIJ. Hoe goed de *huidige* veiligheidssituatie in kaart wordt gebracht met het ARIJ werd niet onderzocht.

## Uitkomstmaat terugkeer

Terugkeer naar de jeugdbescherming is gedefinieerd als het opnieuw starten van een hulpverleningstraject door JBRA vanwege een problematische opgroei- of opvoedingssituatie (dichotome variabele: 1 = ‘ja’/0 = ‘nee’). Dit betekent dat er sprake is van een serieuze bedreiging van de ontwikkeling van het kind door ofwel een onveilige thuissituatie ofwel door ouders die niet in staat zijn te voorzien in de basisbehoeftes van een kind. De periode waarin terugkeer is gemeten varieerde van 10 maanden tot 3 jaar (gemiddelde periode is 19 maanden) na beëindiging van het eerste hulpverleningstraject. Terugkeer in de follow-upperiode is afhankelijk van de uitkomst van het LIRIK, omdat in geval van een onveilige situatie en/of hoog risico de kans het grootst is dat kinderen uit huis worden geplaatst, waardoor juist minder terugkeer is in de hoogste risicogroep. Daarom is gecorrigeerd voor de perioden dat het kind niet aanwezig was in het gezin gedurende de follow-upperiode (uithuisplaatsingen).

## Analyses

Om de sterkte van de samenhang tussen de individuele items van het LIRIK en terugkeer te onderzoeken, werden punt-biseriële ( $r_{pb}$ ) correlaties berekend. De predictieve validiteit van het klinisch oordeel van het LIRIK (‘zeer ernstig risico’, ‘ernstig risico’, ‘matig ernstig risico’, ‘gering risico’) werd onderzocht aan de hand van de Area Under the receiver-operating-characteristic Curve (AUC). De AUC wordt beschouwd als een geschikte maat om de predictieve validiteit te bepalen (Rice en Harris 2005). De AUC geeft globaal aan welk percentage correcte classificaties het instrument overall zal opleveren (Hanley en McNeil 1982). Bij een waarde van .50 voorspelt een instrument niet beter dan op basis van toeval het geval zou zijn. Bij een waarde van 1.00 is sprake van een perfect positieve voorspelling en bij een waarde van .00 is sprake van een perfect negatieve voorspelling. AUC-waarden vanaf .639 corresponderen met een medium effect ( $d = .50$ ) en AUC-waarden vanaf .714 corresponderen met een groot effect ( $d = .80$ ; Rice en Harris 2005).

De actuariële risicoclassificatie werd ontwikkeld op basis van een Chi-squared Automatic Interaction Detector (CHAID)-analyse. CHAID is een classificatietechniek waarmee interactie-effecten tussen onafhankelijke variabelen worden opgespoord, waardoor combinaties van risicofactoren geïdentificeerd kunnen worden die leiden tot een hoog of laag risico. Deze techniek is daarom zeer geschikt om zicht te krijgen op profielen van kinderen met een hoge respectievelijk lage kans op terugkeer (Thomas en Leese 2003).

Als afhankelijke variabele in de CHAID-analyse werd de gedichotomiseerde variabele terugkeer (0 = geen terugkeer in de terugkeerperiode en 1 = terugkeer in de terugkeerperiode) ingevoerd. De individuele risicofactoren uit het LIRIK die significant samenhangen met terugkeer (zie tab. 1) werden ingevoerd als onafhankelijke variabelen in de CHAID-analyse. Daarnaast werden de individuele

**Tabel 1** Punt-biseriële correlaties tussen LIRIK-items en terugkeer naar jeugdbescherming.

items	$r_{pb}$
1a de directe veiligheid van het kind wordt bedreigd door ouders	
1a1 – ernstige kindermishandeling	.01
1a2 – onvoldoende bescherming	.05**
1a3 – onvoldoende basiszorg	.03+
1a4 – opvoeder niet beschikbaar	.02
1a5 – ouders belemmeren zicht op/toegang tot jeugdige	.05**
1b directe veiligheid bedreigd door jeugdige zelf	.01
1c directe veiligheid bedreigd door ander gezinslid	.04*
2 concrete aanwijzingen voor	
2a – lichamelijk of psychisch geweld	.03+
2b – verwaarlozing	.04*
2c – seksueel misbruik	.02
2d – getuige van huiselijk geweld	.08***
3 zorgen over opvoeding en verzorging	
3a – bescherming en veiligheid	.06***
3b – basale verzorging	.04*
3c – emotionele warmte (ondersteuning)	.01
3d – regels en grenzen	.04*
3e – stimulering	.02
3f – stabiliteit	.08***
4 zorgen over de jeugdige	
4a – psychosociaal functioneren	.01
4b – lichamelijke gezondheid	-.03
4c – vaardigheden en cognitieve ontwikkeling	.01
5 risicofactoren bij ouder(s)	
5a – eerder een kind mishandeld, verwaarloosd of misbruikt	.01
5b – ontoereikende pedagogische kennis en/of vaardigheden	.01
5c – problemen in de ouder-kind-interactie	.01
5d – minimaliseren/ontkennen aangetoonde kindermishandeling	.01
5e – negatieve houding t.a.v. kind	-.01
5f – psychiatrische problematiek	.06**
5g – verslavingsproblematiek	.04+
5h – verstandelijke beperking	.03+
5i – fysieke beschikbaarheid	.04*
5j – emotionele beschikbaarheid	.02
5k – op jonge leeftijd ouder geworden	.01
5l – zelf slachtoffer van kindermishandeling	.05*
5m – eerder geweld gebruikt tegen personen	.08***
5n – problematische partnerrelatie	.07***

**Tabel 1** Punt-biseriële correlaties tussen LIRIK-items en terugkeer naar jeugdbescherming. (Vervolg)

items	$r_{pb}$	
6	risicofactoren bij jeugdige(n)	
6a	– jong kind (<5 jaar)	.07***
6b	– belaste voorgeschiedenis (b.v. prematuur)	.01
6c	– (ernstige) ziekte of handicap	.02
6d	– gedrags- en/of ontwikkelingsproblemen	.01
6 <sup>e</sup>	– moeilijk temperament	.01
6f	– ongewenst	.02
7	risicofactoren in gezin en in omgeving	
7a	– eenoudergezin, stiefgezin, groot gezin	.03 <sup>+</sup>
7b	– veel conflicten/huiselijk geweld	.09***
7c	– instabiel, ongeregeld leven	.01
7d	– materiële/financiële problemen (werkloosheid, huisvesting)	.06**
7 <sup>e</sup>	– ingrijpende levensgebeurtenissen	.02
7f	– sociaal isolement/sociaal conflict	.02
8	beschermende factoren bij ouders	
8a	– gevoel van competentie, draagkracht	–.01
8b	– positief zelfbeeld	.00
8c	– ondersteunende partner	–.03*
8d	– kan eigen jeugdervaringen hanteren	.03
8e	– positieve jeugdervaringen	.00
8f	– kan steun vragen/profiteren van	.01
8g	– emotionele beschikbaarheid	.01
8h	– flexibiliteit	–.01
8i	– bereid en in staat om te veranderen	.03*
9	beschermende factoren bij jeugdige	
9a	– sociaal vaardig	.01
9b	– positief zelfbeeld	–.01
9c	– bovengemiddelde intelligentie	–.01
9d	– aantrekkelijk uiterlijk	.05**
9e	– goede relatie met belangrijke volwassene(n)	.04*
9f	– ego-veerkracht (stressresistentie)	–.00
9g	– bereid en in staat om te veranderen	–.03 <sup>+</sup>
10	beschermende factoren in gezin en in omgeving	
10a	– steun informeel netwerk	.02
10b	– steun formeel netwerk	.04*
11	conclusies risico's voor de jeugdige	
11a	– risico's op dit moment (zeer groot, groot, reëel, klein)	.04*
11b	– risico's bij voorziene veranderingen in de nabije toekomst (zeer groot, groot, reëel, klein)	.03

<sup>+</sup> $p < .10$ ; \* $p < .05$ ; \*\* $p < .01$ ; \*\*\* $p < .01$

risicofactoren van het LIRIK, die ofwel significant samenhangen met terugkeer (zie tab. 1) ofwel sterk empirisch onderbouwd zijn in andere onderzoeken, bij elkaar opgeteld in een somvariabele genaamd ‘totaal aantal risicofactoren’. Dit betreffen de volgende risicofactoren: concrete aanwijzingen voor verwaarlozing; lichamelijk of psychisch geweld; seksueel misbruik of getuige van huiselijk geweld; zorgen over bescherming en veiligheid; zorgen over de basale verzorging; zorgen over regels en grenzen; zorgen over stabiliteit; ouder(s) hebben eerder een kind mishandeld, verwaarloosd of misbruikt; ouder(s) met psychiatrische problemen; ouder(s) met verslavingsproblematiek; ouder(s) met een verstandelijke beperking; ouder(s) zijn niet fysiek of emotioneel beschikbaar voor het kind; ouder(s) zelf slachtoffer van kindermishandeling; ouder(s) hebben eerder geweld gebruikt tegen personen; problematische partnerrelatie; jong kind (<5 jaar); veel conflicten/huiselijk geweld; materiële/financiële problemen; en sociaal isolement/sociaal conflict.

In een CHAID-analyse wordt de totale groep gezinnen verdeeld in een aantal subgroepen op basis van de variabelen (risicofactoren) die het sterkst samenhangen met terugkeer. De CHAID-procedure bestaat uit een aantal stappen waarbij in de eerste stap de totale groep wordt opgesplitst in een aantal subgroepen op basis van de variabele die het sterkst samenhangt met terugkeer. In de tweede stap wordt elke subgroep weer opnieuw opgesplitst op basis van de variabele die dan het sterkst samenhangt met terugkeer in de betreffende subgroep. Deze procedure wordt herhaald totdat er in de subgroepen geen variabelen meer zijn die significant samenhangen met terugkeer of totdat de groepen een minimale omvang hebben bereikt. Zo ontstaat een aantal risicogroepen van gezinnen met dezelfde soort risicofactoren en eenzelfde kans op terugkeer.

Om de risico-indeling te kruisvalideren werd gebruikgemaakt van  $k$ -voudige kruisvalidatie, met  $k = 10$ . Bij deze methode wordt de totale steekproef random gesplitst in  $k$  deelsteekproeven van gelijke omvang om de risico-indeling te kunnen valideren. Vervolgens worden  $k$  iteraties van modelbouw en validatie uitgevoerd zodat in elke iteratie een andere deelsteekproef specifiek gebruikt wordt voor validatie, terwijl de overige ( $k-1$ ) deelsteekproeven worden gebruikt voor modelbouw.

Om de predictieve validiteit van de risico-indeling te onderzoeken is de AUC-waarde berekend. Daarnaast zijn de sensitiviteit, specificiteit, het aantal vals positieven en het aantal vals negatieven onderzocht. De sensitiviteit is de kans op een positieve uitslag bij jongeren waarbij daadwerkelijk sprake is van een problematische opgroei- of opvoedingssituatie en de specificiteit is de kans op een negatieve uitslag bij jongeren waarbij in werkelijkheid geen sprake is van een problematische opgroei- of opvoedingssituatie. Een onjuiste beslissing betreft zowel een positieve uitslag bij een jongere waarbij in werkelijkheid geen sprake is van een problematische opgroei- of opvoedingssituatie (vals positief) als een negatieve uitslag bij een jongere waarbij in werkelijkheid wel sprake is van een problematische opgroei- of opvoedingssituatie (vals negatief). De optimale cutoff-score wordt bepaald door zowel rekening te houden met de hoogte van het aantal vals negatieven en vals positieven.

**Tabel 2** AUC-waarden en punt-biseriële correlaties ( $r_{pb}$ ) tussen de algemene risico-inschattingen van het LIRIK en terugkeer naar de jeugdbescherming.

	$r_{pb}$	AUC (95 % C.I.)
algemene risico-inschatting (huidige moment)	.04*	.53 (.50–.56)
algemene risico-inschatting rekening houdend met voorziene veranderingen in de nabije toekomst	.03	.53 (.50–.56)

\*  $p < .05$

Om te testen of de AUC-waarden van het LIRIK en het ARIJ significant van elkaar verschillen werd de methode van DeLong et al. (1988) gebruikt.

## Resultaten

### Samenhang tussen de individuele LIRIK-items en terugkeer

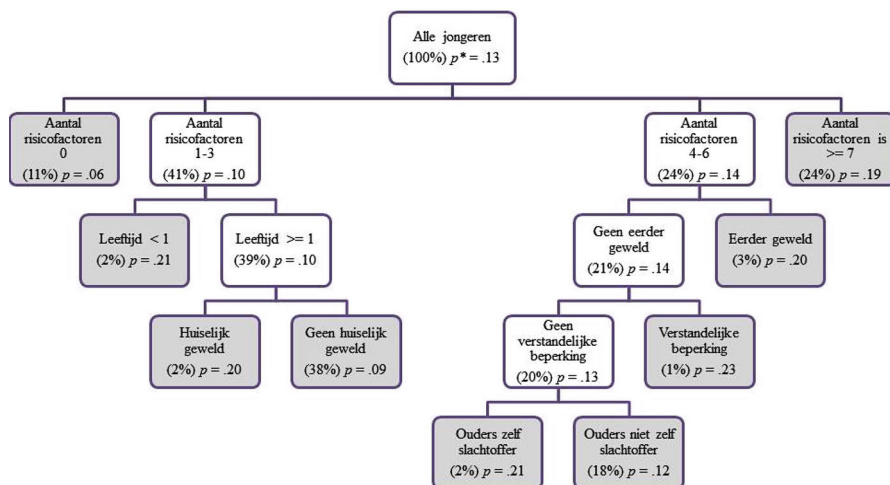
In tab. 1 staan de punt-biseriële correlaties weergegeven tussen de individuele LIRIK-items en terugkeer in de terugkeerperiode. De totale terugkeer was 12,8 % ( $n = 507$ ).

Uit de tabel blijkt dat sprake is van relatief lage correlaties, oftewel een zwakke samenhang, tussen de risicofactoren en terugkeer. Omdat de associaties zijn uitgedrukt in punt-biseriële correlaties (afhankelijke variabele is dichotoom) en omdat de baserate relatief laag was (12,8 % van de steekproef is teruggekeerd), is sprake van een zwakke samenhang vanaf een correlatie van .06 en van een middelsterke samenhang vanaf een correlatie van .16 (berekend met de conversieformule van Rosenthal 1991, uit Rice en Harris 2005).

Risicofactoren die volgens deze criteria zwak samenhangen met terugkeer zijn: getuige van huiselijk geweld; zorgen over bescherming en veiligheid; zorgen over stabiliteit; psychiatrische problematiek bij ouder; ouder heeft eerder geweld gebruikt tegen personen; problematische partnerrelatie; jong kind (<5 jaar); veel conflicten/huiselijk geweld; en materiële/financiële problemen. Er zijn geen beschermende factoren die volgens deze criteria samenhangen met terugkeer.

### Predictieve validiteit van het LIRIK

Om de predictieve validiteit van het LIRIK vast te stellen is de AUC-waarde berekend van beide overall risicoscores van het LIRIK, dus van zowel de risicoscore waarbij wordt uitgegaan van de situatie op het moment waarop het LIRIK wordt ingevuld, als de risicoscore waarbij rekening wordt gehouden met voorziene veranderingen in de nabije toekomst (tab. 2). Uit tab. 2 blijkt dat de inschatting van het risico op terugkeer zoals ingeschat met het LIRIK niet significant beter is dan kans (d.w.z., een willekeurige inschatting).



**Figuur 1** CHAID-output. ( $p$  = kans op terugkeer naar de jeugdbescherming. De grijze vierkanten representeren de risicogroepen die worden onderscheiden.)

## Ontwikkeling ARIJ op basis van CHAID-analyse

Voor de ontwikkeling van het ARIJ is een CHAID-analyse uitgevoerd, met als onafhankelijke variabelen de LIRIK-items die significant samenhangen met terugkeer (zie tab. 1) en de somvariabele ‘totaal aantal risicofactoren’ (zie Methode-paragraaf). Gebaseerd op chi-kwadraattoetsen werd de totale groep ingedeeld in negen verschillende risicogroepen. Fig. 1 geeft de CHAID-output weer (beslisboom) en in tab. 3 worden de verschillende risicogroepen van de CHAID-analyse beschreven. De indeling in risicogroepen is gebaseerd op een combinatie van de volgende zes variabelen: (1) totaal aantal risicofactoren, (2) leeftijd van het kind bij aanvang van de hulpverlening, (3) verstandelijke beperking bij ouder(s), (4) ouder(s) zijn zelf slachtoffer geweest van kindermishandeling, (5) ouder(s) hebben eerder geweld gebruikt tegen personen, en (6) huiselijk geweld. Deze variabelen zijn de sterkste voorspellers voor terugkeer en leveren tevens een unieke bijdrage aan de voorspelling.

De AUC-waarde van het ARIJ was .63 (95 % CI [.60–.66]) en significant hoger dan de AUC-waarde van het LIRIK ( $z = 6,02, p < .001$ ). Tab. 4 laat de sensitiviteit, de specificiteit, het aantal vals positieven en het aantal vals negatieven zien bij de verschillende cut-offscores.

## Discussie

Het eerste doel van deze studie was de predictieve validiteit van het LIRIK te onderzoeken. Het LIRIK is een gestructureerd klinisch instrument dat grootschalig in Nederland wordt gebruikt om het risico op kindermishandeling in te schatten.

**Tabel 3** Risico-indeling op basis van de CHAID-analyse.

		%	<i>p</i>
1	totaal aantal risicofactoren is 4–6; er is sprake van een verstandelijke beperking bij ouder(s); ouder(s) hebben niet eerder geweld tegen personen gebruikt	1,1	.23
2	totaal aantal risicofactoren is 1–3; de leeftijd bij aanvang hulpverlening van het kind is jonger dan 1 jaar	1,9	.21
3	totaal aantal risicofactoren is 4–6; ouders zijn zelf slachtoffer van kindermishandeling; er is <i>geen</i> sprake van een verstandelijke beperking bij ouder(s); ouder(s) hebben <i>niet</i> eerder geweld tegen personen gebruikt	1,6	.21
4	totaal aantal risicofactoren is 1–3; de leeftijd bij aanvang hulpverlening van het kind is 1 jaar of ouder; er zijn concrete aanwijzingen voor recent huiselijk geweld	1,5	.20
5	totaal aantal risicofactoren is 4–6; ouder(s) hebben eerder geweld gebruikt tegen personen	2,9	.20
6	totaal aantal risicofactoren is 7 of meer	23,5	.19
7	totaal aantal risicofactoren is 4–6; er is geen sprake van een verstandelijke beperking bij ouder(s); ouder(s) hebben niet eerder geweld tegen personen gebruikt; ouders zijn niet zelf slachtoffer van kindermishandeling geweest	18,8	.12
8	totaal aantal risicofactoren is 1–3; de leeftijd bij aanvang hulpverlening van de jongere 1 jaar of ouder; er zijn geen concrete aanwijzingen die wijzen op recent huiselijk geweld	37,9	.09
9	totaal aantal risicofactoren is 0	10,8	.06

% = de omvang van de risicogroep als percentage van de totale steekproef ( $n = 3963$ ); *p* is de kans op terugkeer naar jeugdbescherming in de verschillende risicogroepen.

De resultaten lieten zien dat met het LIRIK geen goede inschatting kan worden gemaakt van het terugkeerrisico. De voorspelling van het LIRIK bleek niet beter dan een voorspelling op basis van toeval ( $AUC = .53$ ). Dit is in overeenstemming met eerdere validatieonderzoeken waaruit blijkt dat de meeste klinische risicotaxatie-instrumenten matig presteren (Arad-Davidson en Benbenishty 2008; Baird en Wagner 2000; Barlow et al. 2010; Camasso en Jagannathan 2000; D’Andrade et al. 2005; DePanfilis en Girvin 2005; Dorsey et al. 2008; Knoke en Trocmé 2005; Lyons et al. 1996; Munro 1999; Pfister en Böhm 2008; Wald en Woolverton 1990). Ook hebben eerdere onderzoeken laten zien dat sommige klinische methoden (die breed worden ingezet) niet beter dan toeval presteren, wat betekent dat in veel gevallen een verkeerde inschatting wordt gemaakt (Baird en Wagner 2000; Barber et al. 2008). Het gebruik van het LIRIK om het risico op terugkeer naar de jeugdbescherming in te schatten kan dus tot verkeerde beslissingen leiden, hetgeen kan resulteren in onterechte uithuisplaatsingen of (herhaalde) kindermishandeling.

Veel van de individuele factoren die worden gemeten met het LIRIK bleken niet significant samen te hangen met terugkeer, waaronder de kindfactoren (zoals zorgen t.a.v. psychosociaal functioneren, vaardigheden en cognitieve ontwikkeling van de jeugdige). Ook andere onderzoeken hebben laten zien dat kindfactoren aanzienlijk minder sterk samenhangen met toekomstige kindermishandeling dan ouder- of gezinsfactoren (Cash 2001; Stith et al. 2009). Daarnaast bleek een aantal



**Tabel 4** Sensitiviteit, specificiteit en vals positieven en vals negatieven van het ARIJ.

cut-offscore (>) <sup>a</sup>	sensitiviteit	specificiteit	vals positieven (%)	vals negatieven (%)	totaal aantal foutieve beslissingen (%)
.077	.95	.12	77,1	0,7	77,8
.107	.68	.51	42,7	4,1	46,8
.157	.50	.70	26,2	6,5	32,7
.194	.15	.92	7,2	10,9	18,1
.199	.10	.94	4,8	11,5	16,3
.203	.08	.96	3,6	11,8	15,4
.208	.05	.97	2,4	12,1	14,5
.219	.02	.99	0,9	12,5	13,4

<sup>a</sup>Als een testscore groter is dan de cut-offscore (de kans op terugkeer), wordt het testresultaat als positief beschouwd; een testscore kleiner dan de cut-offscore wordt als negatief beschouwd.

risicofactoren bij ouders niet significant samen te hangen met terugkeer, zoals het eerder hebben mishandeld van een kind, problemen in de ouder-kindinteractie en een negatieve houding ten aanzien van het kind. Uit overzichtsstudies blijkt dat in andere studies wel een samenhang wordt gevonden tussen deze factoren en terugkeer (bijv. Hindley et al. 2006; Stith et al. 2009). Een mogelijke verklaring hiervoor is dat deze concepten in andere studies worden gemeten met behulp van instrumenten die uit meerdere items bestaan, terwijl in het LIRIK deze concepten met één item worden gemeten. Een andere oorzaak kan zijn dat de items niet altijd even zorgvuldig en betrouwbaar worden ingevuld door de professionals van JBRA (zie beperkingen van de studie).

Uit de resultaten bleek ook dat er geen samenhang is tussen de beschermende factoren van het LIRIK en terugkeer. Dit is in overeenstemming met resultaten van recente studies die hebben laten zien dat kinderen uit hoog-risicogezinnen minder profiteren van de eventuele aanwezigheid van beschermende factoren (bijv. Luther en Goldstein 2004; Miller et al. 1999; Vanderbilt-Adriance en Shaw 2008). Dit wijst erop dat interventies gericht op het vermeerderen of versterken van beschermende factoren in hoog-risicogezinnen niet noodzakelijkerwijs effectief zijn en niet hoeven te leiden tot een afname van het terugkeerrisico. Vanderbilt-Adriance en Shaw (2008) concluderen in hun overzichtsstudie dat in hoog-risicogezinnen zowel moet worden gewerkt aan het versterken van beschermende factoren als aan het verminderen van risico's, omdat er grenzen zijn aan de mate waarin risico's kunnen worden verminderd.

Het tweede doel van deze studie was om een actuaireel risicotaxatie-instrument te ontwikkelen (het ARIJ) gebaseerd op variabelen gemeten met het LIRIK. Het ARIJ is ontwikkeld met een CHAID-analyse en bestaat uit een risico-indeling gebaseerd op de volgende variabelen: totaal aantal risicofactoren; ouder(s) hebben eerder geweld gebruikt tegen personen; veel conflicten/huiselijk geweld in het gezin; ouder(s) hebben een verstandelijke beperking; ouder(s) zijn zelf slachtoffer van kindermishandeling geweest; en het kind is jonger dan 1 jaar. Deze variabelen leverden een unieke bijdrage aan de voorspelling van terugkeer.

De predictieve validiteit van het ARIJ bleek significant beter dan die van het LIRIK met een AUC van .63. Deze AUC-waarde is daarmee net iets lager dan de ondergrens van een medium effect (AUC = .639; Rice en Harris 2005). Een gemiddelde AUC-waarde van instrumenten die worden toegepast voor inschatting van het risico op kindermishandeling is niet bekend doordat nog geen meta-analyse is uitgevoerd naar de prestaties van risicotaxatie-instrumenten voor kindermishandeling. Instrumenten die het meest zijn onderzocht op predictieve validiteit zijn de California Family Risk Assessment (CFRA, Baird en Wagner 2000; Dankert en Johnson 2014; Johnson 2011; Johnson et al. 2015; Van der Put et al. 2016) en de Child Abuse Potential Inventory (CAPI, Ayoub en Milner 1985; Chaffin en Valle 2003; Milner et al. 1984; Ondersma et al. 2005). AUC-waarden die zijn gevonden in de studies naar de CFRA variëren van .53 tot .74, met een gemiddelde waarde van .63, en AUC-waarden die zijn gevonden in de studies naar de CAPI variëren van .56 tot .69, met een gemiddelde van .61. Daarnaast zijn op het gebied van (jeugd)delinquentie wel diverse meta-analyses uitgevoerd naar de prestaties van risicotaxatie-instrumenten (Fazel et al. 2012; Olver et al. 2009; Schwalbe 2007). Zo vond Schwalbe (2007) een gemiddelde AUC-waarde van .64, Fazel et al. (2012) vonden een gemiddelde AUC-waarde van .66 en Olver et al. (2009) vonden gemiddelde gewogen correlaties van  $r = .28$  tot  $r = .32$  voor risicotaxatie-instrumenten voor algemene recidive onder jeugdigen. Het ARIJ presteert daarmee vergelijkbaar met andere instrumenten die worden ingezet om het risico op kindermishandeling in te schatten alsmede met instrumenten die worden ingezet om risico's op delinquent gedrag te schatten.

Het ARIJ wordt op dit moment ingezet door JBRA bij alle gezinnen die worden aangemeld en door SPIRIT bij de interventies Spoedhulp en Intensieve Preventieve Aanpak. Op basis van de uitkomst van de veiligheidstaxatie en de risicotaxatie van het ARIJ wordt bepaald welke aanpak het best past. De uitkomst van de veiligheidstaxatie geeft aan in hoeverre een kind direct moet worden veiliggesteld, en op basis daarvan wordt een veiligheidsplan opgesteld ten aanzien van de in te zetten middelen, zoals het inschakelen van het sociaal netwerk, het uit huis halen van de veroorzaker van de onveiligheid, en (diverse vormen van) uithuisplaatsing van het kind. Op basis van de risicotaxatie wordt het risico op toekomstige onveiligheid ingeschat en wordt bepaald welke interventies kunnen worden ingezet om het risico op toekomstige onveiligheid te verkleinen en welke intensiteit gewenst is. In geval van een hoog risico moet een omvattende en intensieve interventie worden ingezet, terwijl bij een laag risico een minder intensieve interventie of geen interventie volstaat. Om het beste afkappunt van het ARIJ te bepalen moet rekening worden gehouden met zowel het aantal vals positieven en vals negatieven alsmede met de consequenties van een vals positief en een vals negatief.

De consequentie van een vals negatief is dat een gezin geen of een niet-intensieve interventie krijgt aangeboden terwijl het kind wel risico op kindermishandeling loopt, en het gevolg van een vals positief is dat een gezin intensieve interventie krijgt aangeboden terwijl het kind in werkelijkheid geen verhoogd risico op kindermishandeling heeft. Om kinderen zo goed mogelijk te beschermen is het belangrijk om het aantal vals negatieven te minimaliseren, maar het is ook

belangrijk dat het aantal vals positieven niet te hoog is, omdat anders kinderen en hun gezinnen onnodig een intensieve behandeling krijgen. Bij een afkappunt van .194 is het percentage vals negatieven 10,9% en het percentage vals positieven 7,2% (totaal aantal foutieve beslissingen is 18,1%). Bij een afkappunt van .157 is het percentage vals negatieven lager (6,5%), maar is het percentage vals positieven hoger (26,6%), evenals het totaal aantal foutieve beslissingen (32,7%). Welk afkappunt de voorkeur geniet is een beleidsbeslissing waarbij ook kosten en capaciteit zullen worden meegewogen.

Bij dit onderzoek hadden we te maken met een aantal beperkingen. In de eerste plaats was de kwaliteit van de ingevulde LIRIK's niet altijd even goed. Het LIRIK wordt ingevuld door gezinsmanagers, maar niet elke gezinsmanager onderkent het belang van het zo goed mogelijk invullen van het instrument. Daardoor zijn de gevonden effecten mogelijk onderschattingen van de werkelijke effecten. Daarnaast zijn voor dit onderzoek alleen de exemplaren van het LIRIK geanalyseerd die digitaal beschikbaar waren in het systeem van JBRA. De verwachting is echter dat deze LIRIK's niet verschillen van de LIRIK's die niet digitaal beschikbaar waren en dat ze dus een goede afspiegeling vormen van de totale populatie van ingevulde LIRIK's bij JBRA. Vanwege financiële beperkingen was het niet mogelijk om LIRIK's uit andere regio's te includeren in de steekproef. Daarom is het niet zonder meer mogelijk om de resultaten van dit onderzoek naar andere steden in Nederland te generaliseren.

Een derde beperking is dat we het ARIJ niet hebben kunnen testen in een nieuwe steekproef en daarom is het belangrijk dat de ecologische validiteit van het ARIJ wordt onderzocht, oftewel de predictieve validiteit bij gebruik in de alledaagse praktijk.

Ten vierde was de uitkomstmaat van dit onderzoek het opnieuw starten van een hulpverleningstraject door JBRA vanwege problematische opgroei- of opvoedings-situaties. Onderzocht is dus hoe goed het LIRIK in staat is te voorspellen welke gezinnen weer terugkomen bij de jeugdbescherming vanwege ernstige bedreiging van de ontwikkeling van een kind door onveiligheid in de thuissituatie of door ouders die niet in staat zijn om te voorzien in de basisbehoeftes van het kind. De predictieve validiteit van zowel het LIRIK als het ARIJ zou ook moeten worden onderzocht voor het voorspellen van (diverse vormen van) kindermishandeling. Tot slot is het belangrijk om het gebruik van het ARIJ te monitoren in de praktijk, omdat eerder onderzoek laat zien dat actuariële methoden niet altijd worden gebruikt zoals bedoeld (DePanfilis en Zuravin 2001; Gillingham 2011; Gillingham en Humphreys 2010).

De resultaten van dit onderzoek zijn van groot belang voor de praktijk van de jeugdbescherming, omdat het ARIJ daar met groot enthousiasme is ontvangen. Zo blijkt uit een interne evaluatie bij JBRA, waar het ARIJ sinds juni 2015 door 270 medewerkers wordt gebruikt, dat het ARIJ daadwerkelijk ondersteunend is in de besluitvorming en het opstellen van een plan van aanpak rondom kindonveiligheid. Hierdoor wordt het ARIJ niet als een verplichting, maar als een zeer zinvolle toevoeging ervaren (i.t.t. het LIRIK). Tevens is door gebruik van het ARIJ de handelingsverlegenheid afgenomen. Voorheen vonden professionals het lastig om

ouders uit te leggen waarom ze het risico op onveiligheid hoog inschatten, terwijl ze nu kunnen wijzen op specifieke risico's voor het kind in bepaalde domeinen. Daarom hebben zeer veel instellingen aangegeven met het ARIJ te willen werken, met name in Amsterdam en Zaanstad, zoals de ouder-kindteams, wijkteams, Samen DOEN-teams, Veilig Thuis, HVO Querido, de Bascule, Polypartners, Altra Jeugdhulp, KOC en Cordaan. De Gemeente Amsterdam heeft inmiddels de wens geuit dat het ARIJ het leidende keteninstrument van jeugdzorginstellingen in de regio Amsterdam wordt, zodat al deze instellingen gebruik gaan maken van hetzelfde instrument en daarmee 'dezelfde taal spreken'.

Let wel, het is belangrijk dat in vervolgonderzoek voortdurend wordt onderzocht hoe de predictieve validiteit van het ARIJ verder kan worden verbeterd. De verwachting is dat de predictieve validiteit van het ARIJ bij een tweede meting hoger zal zijn dan bij de huidige meting. In het huidige onderzoek waren de risicofactoren namelijk gemeten middels het LIRIK, waarvan de betrouwbaarheid laag is (Veenhuizen 2013), terwijl een eerste onderzoek naar de betrouwbaarheid van het ARIJ lijkt te wijzen op voldoende tot goede interbeoordelaarsbetrouwbaarheid (Busschers et al. 2015). De verwachting is dat dit een positief effect heeft op de predictieve validiteit van het ARIJ. Daarnaast zijn aan het ARIJ een aantal experimentele items toegevoegd waarvan in ander onderzoek een sterke samenhang met toekomstige onveiligheid is vastgesteld (Put et al. 2015). Deze items wegen nu nog niet mee in de risico-indeling, maar in een vervolgstudie kan worden onderzocht of de predictieve validiteit van het ARIJ verder kan worden verbeterd door deze experimentele items te laten meewegen.

**Dankbetuiging** De auteurs willen graag Marc Dinkgreve, Inge Busschers en Mirte Forrer bedanken voor hun kritische en waardevolle suggesties ten aanzien van het onderzoek. Daarnaast bedanken we Pro Juventute en Jeugdbescherming Regio Amsterdam (JBRA) voor het financieel mogelijk maken van dit onderzoek alsmede voor het beschikbaar stellen van de data die nodig waren voor dit onderzoek.

---

## Literatuur

- Aegisdóttir, S., White, M. J., Spengler, P. M., Maugherman, A. S., Anderson, L. A., Cook, R. S., Rush, J. D., et al. (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.
- Ayoub, C. C., & Milner, J. S. (1985). Failure to thrive: Parental indicators, types, and outcomes. *Child Abuse & Neglect*, 9(4), 491–499.
- 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. <https://www.gov.uk/government/publications/systematic-review-of-models-of-analysing-significant-harm>. Geraadpleegd op: 11 november 2013.
- Bartelink, C., Kwaadsteniet, L. de, Berge, I. ten, Witteman, C., & Gastel, W. van (2015). *Betrouwbaarheid en validiteit van de LIRIK: Eindrapport LIRIK valideringsonderzoek*. Utrecht: Nederlands Jeugdinstituut.

- Berge, I.J. ten (2008). *Instrumenten voor risicotaxatie in situaties van (vermoedelijke) kindermishandeling*. Utrecht: Nederlands Jeugdinstituut.
- Berge, I.J. ten, & Eijgenraam, K. (2009). *Licht Instrument Risicotaxatie Kindveiligheid (LIRIK) [Check List of Child Safety (CLCS)]*. Utrecht: Nederlands Jeugdinstituut.
- Busschers, I., Forrer, M., & Dinkgreve, M. (2015). *Interbeoordelaarsbetrouwbaarheid Actuarieel Risicotaxatie Instrument voor Jeugdbescherming (ARIJ)*. Amsterdam: JBRA.
- 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.
- Dankert, E. W., & Johnson, K. (2014). Risk assessment validation: a prospective study. NCCD Children's Research Center. [http://www.nccdglobal.org/sites/default/files/publication\\_pdf/risk-assessment-validation.pdf](http://www.nccdglobal.org/sites/default/files/publication_pdf/risk-assessment-validation.pdf). Geraadpleegd op: 15 juli 2016.
- 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, band 243, pag. 1668–1674.
- DeLong, E. R., DeLong, D. M., & Clarke-Pearson, D. L. (1988). Comparing the areas under two or more correlated receiver operating characteristic curves: A nonparametric 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.
- 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.
- D'Andrade, A., Benton, A., & Austin, M. J. (2005). Risk and safety assessment in child welfare: Instrument comparisons. Center for Social Services Research. [http://cssr.berkeley.edu/bassc/public/risk\\_full.pdf](http://cssr.berkeley.edu/bassc/public/risk_full.pdf). Geraadpleegd op: 8 juli 2016.
- 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 24,827 people: systematic review and meta-analysis. *British Medical Journal*, doi:10.1136/bmj.e4692.
- 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), 19–30.
- 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.
- Hilton, N. Z., Harris, G. T., & Rice, M. E. (2006). Sixty-six years of research on the clinical versus actuarial prediction of violence. *The Counseling Psychologist*, 34(3), 400–409.
- Hindley, N., Ramchandani, P. G., & Jones, D. P. (2006). Risk factors for recurrence of maltreatment: a systematic review. *Archives of Disease in Childhood*, 91(9), 744–752.
- 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.

- Johnson, W., Clancy, T., & Bastian, P. (2015). Child abuse/neglect risk assessment under field practice conditions: Tests of external and temporal validity and comparison with heart disease prediction. *Children and Youth Services Review, 56*, 76–85.
- Kaufmann, E., & Wittmann, W.W. (2016). The success of linear bootstrapping models: Decision domain-, expertise-, and criterion-specific meta-analysis. *PLOS ONE, 11*(6), e0157914.
- Knoke, D., & Trocmé, N. (2005). Reviewing the evidence on assessing risk for child abuse and neglect. *Brief Treatment and Crisis Intervention, 5*, 310–327.
- 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.
- 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.
- Milner, J.S., Gold, R.G., Ayoub, C., & Jacewitz, M.M. (1984). Predictive validity of the child abuse potential inventory. *Journal of Consulting and Clinical Psychology, 52*(5), 879.
- 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.
- Ondersma, S.J., Chaffin, M.J., Mullins, S.M., & LeBreton, J.M. (2005). A brief form of the Child Abuse Potential Inventory: Development and validation. *Journal of Clinical Child and Adolescent Psychology, 34*(2), 301–311.
- 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. <https://www3.aifs.gov.au/cfca/sites/default/files/publication-documents/rs24.pdf>. Geraadpleegd op: 15 juli 2016.
- Put, C.E. van der, Assink, M., & Stams, G.J.J.M. (2015). *Actuarieel Risicotaxatie Instrument voor Jeugdbescherming (ARIJ)*. Amsterdam: Universiteit van Amsterdam.
- Put, C.E. van der, Hermanns, J., Gelderen, R.-V.L. van, & Sondejker, F. (2016). Detection of unsafety in families with parental and/or child developmental problems at the start of family support. *BMC Psychiatry, 16*(1), 1.
- 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), 615–620.
- Rosenthal, R. (1991). *Meta-analytic procedures for social research*. Newbury Park: Sage.
- Schwalbe, C.S. (2007). Risk assessment for juvenile justice: a meta-analysis. *Law and Human Behavior, 31*, 449–462.
- Stith, S.M., Liu, T., Davies, L.C., Boykin, E.L., Boykin, E.L., Alder, M.C., Harris, J.M., Dees, J.E.M.E.G., et al. (2009). Risk factors in child maltreatment: A meta-analytic review of the literature. *Aggression and Violent Behavior, 14*(1), 13–29.
- 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*(3), 153–158.
- 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*. Amsterdam: VU Amsterdam. Master's thesis
- Wald, M.S., & Woolverton, M. (1990). Risk assessment: The emperor's new clothes? *Child Welfare, 69*, 483–511.

**Dr. Claudia van der Put** Universitair Docent (UD)

**Drs. Mark Assink** postdoctoraal onderzoeker (postdoc)

**Prof. dr. Geert Jan M. Stams** hoogleraar Forensische Orthopedagogiek